

Einführung in die KI

Prof. Dr. sc. Hans-Dieter Burkhard
Dipl.-Inf. Daniel Göhring
Vorlesung Winter-Semester 2005/06

Sprachverarbeitung
Repräsentation
Verstehen

„da kamen plötzlich drei Mädchen“

Sprachlicher Kommunikationsprozess

„da kamen plötzlich drei Mädchen“

- Kontext-Bezug: wenig isolierte Information
- Partnermodell: Hoffnung auf Verstehen
- Benutzung „vereinbarter“ Formen
- Benutzung gemeinsamer Vorstellungen
 - gemeinsame Umwelt
 - allgemeines Beispiel:
Benutzung räumlicher Konzepte in der Sprache

Im Garten, im Januar, im Zusammenhang, im Prinzip, im Namen,
im Dienst, im Radio, im Elend, im Scherz, im Werden, ...

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

3

Sprachlicher Kommunikationsprozess

Gesamtprozess jeweils durch lokale Prozesse realisiert:

- lokal bzgl. Zeitfenster
- lokal bzgl. Sprachebene
- lokal bzgl. Sender/Empfänger

Robustheit bzgl. lokaler Fehlleistungen

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

4

Sprachlicher Kommunikationsprozess

	akustisch	visuell
aktiv(Synthese)	<i>Sprechen</i>	<i>Schreiben</i>
passiv(Analyse)	<i>Hören</i>	<i>Lesen</i>

Zeitlich - linearer Verlauf

Segmentierung/Strukturierung

Pausen, Betonung, Satzzeichen, Formatierung, Mausclick, Sprech taste

inkrementelle (zeitlich versetzte) Parallelarbeit

- Hören/Verstehen
- Verbalisieren/Sprechen

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

5

Bedeutung der Sprache

- Verständigung
 - einmalig in der Natur hinsichtlich Flexibilität
- Koordination
 - Verhandlungen/Übereinkünfte erzielen
- Übermittlung von Wissen
- Bewahrung von Wissen (Schriftsprache)
- Denken

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

6

In Worte fassen

- Sprache und Denken
- Wörter vs. Konzepte
- Sätze vs. Gedanken
- Verselbständigung von Begriffen

Denn eben wo Begriffe fehlen,
da stellt ein wahres Wort zur
rechten Zeit sich ein.

Sprache

- Exakte Formulierungen sind schwierig.
 - und was denkt der andere dabei?
- Vollständige Beschreibungen sind unmöglich.
 - außer in der Mathematik
- Ökonomie-Prinzipien:
 - nur wesentliches mitteilen, Rest hinzudenken
 - vielfache Verwendbarkeit (Mehrdeutigkeit)

Inkrementelle Verarbeitungsschritte

- Mitteilungswunsch
- Handlungsplanung
- Verbalisierung
- Ansteuerung der Sprechorgane
- aktive Bewegungen der Sprechorgane
- Folge von Lauten
- *akustische Schwingungen: Übertragung*
- Folge von Phonemen
- induzierte Bewegungen der Hörorgane
- Übertragung in das Gehirn
- Sprachliche Gebilde synthetisieren
- Interpretation
- Verstehen

Probleme der
Rücktransformation:
Verrauschte Übertragung
Mehrdeutigkeiten

Erwartungsgesteuerte Verarbeitung

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

9

Flexibilität und Mehrdeutigkeit

- Er liest Kartoffeln.
- Sie liest die Zeitung.

- Sie sah die Fliege, die die Mutter schlug.

- Equals goes it loose.

- Sie geht in die Vorlesung der Professorin,
die sie schon lange nicht besucht hat in ihrer Studenten-Wohnung.

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

10

Korrektheit/Semantik, Robustheit

ich habe fertig

- Ein hessliches Tür.
- Da kann sie geholfen werden.
- Der Brille sind beschlagt.
- Er startet stündlich Kartoffeln in der Fahne mit dem Plasma.

Prozesse bzgl. Sprache

- Analyse (erfassen): „hören/lesen“, verstehen
- Synthese (erzeugen): „sprechen/schreiben“
- Bearbeiten
 - Übersetzen (Analyse + Synthese)
 - Klassifizieren (Analyse)
 - Extrahieren (Analyse + Synthese)
- Management
 - Strukturieren
 - Dokumentieren
 - Speichern/Suchen

Assistenz für Mensch-zu-Mensch

- Übersetzung

- beschränkte Domänen, z.B. technisch (Wetterbericht)
- allgemeine Domänen, z.B. EU-Übersetzungsdienst
- Literatur

technischer Stand:

- geschriebene Sprache gemäß sprachlicher Oberflächenform
- statistische Verfahren
- gesprochene Sprache: z.B. VERBMOBIL

- Hilfe für Behinderte

- Hilfsmittel für Sprachunterricht

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

13

Übersetzungen (Rückübersetzung aus dem Englischen)

Sie haben eine Vortragzeit von 10 Minuten für Ihr Thema. Erwarteten Sie nicht in den Nachfragen dieser Zeit der Testkommission. Stellen Sie sich zum Beispiel, Sie sind mit Ihrem Vortrag das Problem andere zu beschreiben vor Informatiker, damit diese ein Urteil

- bilden können, das es das Frage analysiert (Klassifikation erklären von Bezeichnungen, von Problemen...).
- , das es abhängig nach Frage bespricht, welche, die Lösung betraf, die Theorie, das Verfahren...
- , welches ihm die Antwort gibt, wenn die Aufgabe ein Frage
- enthält, das Sie Probleme anruft, Nachteile, Vorteile, Alternativen...

Sprechen Sie offenbar und offenbar. Verwenden Sie die üblichen technischen Bezeichnungen. Verwenden Sie Skizzen, Formeln, etc. für Erklärung. Betrachten Sie sich, wie Sie gegenüber von (angenommen anderem Informatikerin) Ihnen verstehen können gut. Verwenden Sie Ihre Zeit und halten Sie Sie für Grenzzeit.

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

14

Übersetzungen

Ursprünglicher deutscher Text

Sie haben für Ihr Thema eine Vortragszeit von 10 Minuten.
Erwarten Sie in dieser Zeit keine Nachfragen der Prüfungskommission.

Stellen Sie sich zum Beispiel vor, Sie sollen mit Ihrem Vortrag das Problem anderen Informatikern erläutern, so dass diese sich ein Urteil bilden können.

- Analysieren Sie die Fragestellung (Einordnung Klärung von Begriffen, Probleme, ...).
- Diskutieren Sie je nach Fragestellung die betreffende Lösung, die Theorie, das Verfahren, ...
- Geben Sie die Antwort, wenn die Aufgabe eine Frage beinhaltet.
- Nennen Sie Probleme, Nachteile, Vorteile, Alternativen, ...

Sprechen Sie klar und deutlich. Verwenden Sie die üblichen Fachbegriffe. Nutzen Sie Skizzen, Formeln, usw. zur Erläuterung.

Überlegen Sie sich, wie Ihr Gegenüber (eine angenommene andere Informatikerin) Sie am besten verstehen kann.

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

15

Übersetzungen

(Deutsch→Englisch)

They have a lecture time of 10 minutes for your topic. Do not expect in this time demands of the test commission. Imagine for example, you are with your lecture the problem other computer scientists to describe, so that these can form a judgement.

- analyzing it the question (classification clarifying of terms, problems...).
- discussing it depending upon question the solution concerned, the theory, the procedure...
- giving it the answer, if the task contains a question
- calling you problems, disadvantages, advantages, alternatives...

Speak clearly and clearly. Use the usual technical terms. Use sketches, formulas, etc. for explanation. Consider yourselves, how you can understand opposite (accepted other Informatikerin) you best. Use your time and consider you time limit.

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

16

Assistenz: Texterzeugung

Editoren, Textprozessoren

- Erreichter Stand:
 - Silbentrennung, Korrektur, Wörterbuch
 - Vorlagen
 - Gestaltung (LaTeX, Word, etc.)
 - spezialisierte (z.T. personalisierte) Diktiersysteme
- in Arbeit:
 - Grammatik-Korrektur
 - Stilistische Korrektur
 - (multilinguale) automatische Textgenerierung

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

17

Interface Mensch-Maschine

Natürlich-sprachliches Interface (Schrift, gesprochene Sprache)

- Eingabe
 - Analyse (ggf. gesprochener) Sprache
 - Vervollständigung (Texteingabe)
 - Interpretation von Kommandos
 - Modellierung von Nutzerabsichten (Dialog)
- Ausgabe
 - vorgegebene Fragen/Antworten
 - Sprachsynthese

Einfache Modelle

- beschränktes Repertoire
- Skriptbasis
- ELIZA-artig

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

18

Automatische Arbeit

- Indizierung (Suchmaschinen), Retrieval
- Textanalyse
 - Klassifizierung
 - Inhaltsextraktion
- Texterzeugung
 - Dokumentenerzeugung (z.B. Produkt-Beschreibung)
- Sprachsynthese

Informations-/Wissens-Management

Effizienter Umgang mit „Wissen“:

- Erzeugung/Pflege von Dokumentensammlungen
- Strukturierung
- Erfassung von Wissen:
 - Experten-Interview
 - Tätigkeits-Protokolle
 - Automatische Erfassung
- **Auswertung**
- Datenschutz
- Datensicherheit

„Wenn Siemens wüßte,
was Siemens weiß“

Informations-/Wissens-Management

- Auswertung:
 - Informationsfilterung, -klassifikation (Suchmaschinen)
Retrieval einzelner Dokumente aus Sammlung
 - Informationsextraktion, -erschließung
Analyse einzelner Dokumente:
 - Inhalt extrahieren
 - Detektion von Information
 - Repräsentation
 - Data Mining:
Auswertung von Dokument-Sammlungen

Informations-/Wissens-Management

Techniken:

- Maschinelles Lernen
 - Fallbasiertes Schließen
 - Mustererkennung
- (statistische) Auswertung von
- Oberflächlichen-Merkmalen bzgl. Wortformen
 - Oberflächenmerkmalen bzgl. Satzstruktur
 - Oberflächenmerkmalen bzgl. semantischer Ähnlichkeit
- Bessere Ergebnisse erfordern „Verstehen“

Repräsentation

- Wortebene
 - Wortbedeutungen
 - Wortformen
 - Kompositionen
- Satzebene
 - semantische Bezüge
 - strukturelle Bezüge, abgeleitete Wortformen
- Diskursebene
 - strukturelle/semantische Bezüge
 - Referenzen

Repräsentation: Wortbedeutung

- Homonyme: Hahn
- Synonyme: Samstag, Sonnabend
- Ähnlichkeit: Kunde, Käufer
- Polysemie: Schule

- Hintergrund:
 - Assoziationen „Rathaustür“

- Aufbereitung des notwendigen Wissens: Lexikon

Repräsentation: Morphologie

- Beschreibung von
 - Eigenschaften eines „Lexems“
 - Grundform **verstell**
 - phonetische Form
 - Grammatikalische Eigenschaften **reflexiv, transitiv, ...**
 - Grammatikalische Formen **verstellte**
 - Wort-Komposition
 - Bedeutung (Konzept)
 - Eigenschaften einer Wortform
 - grammatische Form: Klassifizierung
 - grammatische Form: Erscheinung

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

25

Wortformen

- Anotationen zu Wörtern bzgl.
 - Casus, Numerus, Flexion etc.
 - Zusammensetzungen
- Reguläre Bildungen:
 - Regeln für Formbildungen
 - Lexikon-Einträge mit Formen
- irreguläre Bildungen:
 - Lexikon-Einträge mit Formen
- Realisierungen
 - Automaten mit Ausgabe: „finite state transducers“
 - Statistische Verfahren/Lernverfahren
(Zusammenhang mit vorangehenden Wörtern)

Analyse-Systeme: Tagger

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

26

Repräsentation: Lexikon

Datenbank für Wortbezogene Informationen

- Grammatische Informationen
- ggf. Übersetzungen
- Wortbedeutung(en)

Allg. z.B.: WordNet

Zusammenhänge zwischen Wörtern

- Semantische Netze
 - Darstellung semantischer Bezüge zwischen Wörtern
- Ontologische Systeme
 - Darstellung hierarchischer Bezüge zwischen Wörtern
 - Auch zur Vorgabe von Bezeichnern für Konzepte (Standardisierung)
- Terminologische Systeme
 - Mengen/Logik-basierte Bezüge

Speziell:
Domänenbezogene Lexika

Repräsentation auf Satzebene

Grammatik: Syntax

- Regeln für Satzbildung
- Darstellung von Strukturen/Abhängigkeiten durch Graphen/Bäume
 - Conceptual dependency theory
 - Konstituentengrammatik (Syntax)
 - Constraintbasierte Grammatiken
- (eingeschränkte) Kontroll-Funktion

Verarbeitung (Parsing) mit Methoden wie im Compilerbau

Repräsentation eines Diskurses

- Sukzessive Erweiterung eines Modells gemäß
 - erfolgter Satzanalysen
 - Referenzen
 - Hintergrundwissen
- Repräsentationformen z.B.
 - Frame
 - Skript
 - Terminologisches System
 - Semantisches Netz
 - Logik

Repräsentation von Bedeutungen

- Formalismen:
 - Prädikatenlogik
kaufen(er,sie,es)
 - Conceptual dependency theory u.a.

Unabhängigkeit von grammatikalischer Form

Er kauft es von ihr.

Sie verkauft es ihm.

Es wird ihm von ihr verkauft.

Es wird von ihr an ihn verkauft.

Von ihr wird es an ihn verkauft.

Von ihm wird es von ihr gekauft.

Repräsentation von Bedeutungen

Sie beginnt um neun im Raum 0.313 .

- Semantik-Konstruktion (Semantisches Potenzial):
BEGIN(person, event, time, location)
1.Variante: person=female, event=?, time∈{9,21}, location= 0.313
2.Variante: person=?, event=female, time∈{9,21}, location= 0.313
- Semantik-Resolution (aktueller semantischer Wert):
 - Auflösung von Mehrdeutigkeiten (Disambiguierung)
 - Referenzierung im Kontext
z.B.: *Wann beginnt die Vorlesung?*2.Variante:
person=?, event=female, time =9, location= 0.313

Repräsentation von Bedeutungen

Wann beginnt die Vorlesung?

Sie beginnt um eins im Raum 0.313 .

person=?, event=female, time =9, location= 0.313

- Semantische Auswertung:
 - Relevante Information extrahieren
 - Benutzung von Weltwissen usw., auch Hörer-abhängigz.B.:
time =Mittwoch, 9.15-10.45
location=Schrödinger Zentrum, Erdgeschoss, Hörsaal 0.313
event =Einführung in die KI

Tatsächlich gewünschte Information: *nach 9.00*

Negation umgangssprachlich

Genauere Analyse mittels Logik, z.B.

She must not smoke.

$\text{must}(\neg \text{smoke})$ *Sie darf nicht rauchen.*

Nicht: $\neg \text{must}(\text{smoke})$ *Sie muß nicht rauchen.*

Nicht nur Peter ist intelligent.

Variante 1: $\neg(\text{nur Peter ist intelligent})$

$\neg(\text{intelligent}(\text{Peter}) \wedge \forall x(x \neq \text{Peter} \rightarrow \neg \text{intelligent}(x)))$

$\neg \text{intelligent}(\text{Peter}) \vee \exists x(x \neq \text{Peter} \wedge \text{intelligent}(x))$

Variante 2: $\text{intelligent}(\text{Peter}) \wedge \exists x(x \neq \text{Peter} \wedge \text{intelligent}(x))$

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

33

Negation umgangssprachlich

Schlafen die Kinder im Zimmer? - Nein.

$\neg \forall x(\text{Kinder}(x) \wedge \text{imZimmer}(x) \rightarrow \text{schlafen}(x))$

– Variante 1: *Die Kinder im Zimmer schlafen nicht.*

$\forall x(\text{Kind}(x) \wedge \text{imZimmer}(x) \rightarrow \neg \text{schlafen}(x))$

– Variante 2: *Im Zimmer schlafen nicht die Kinder.*

$\forall x(\text{schlafen}(x) \wedge \text{imZimmer}(x) \rightarrow \neg \text{Kind}(x))$

– Variante 3: *Die Kinder schlafen nicht im Zimmer.*

$\forall x(\text{Kind}(x) \wedge \text{schlafen}(x) \rightarrow \neg \text{imZimmer}(x))$

Formal nicht die Negation der oberen Formel

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

34

Quantifizierung umgangssprachlich

- *Der Lehrer hat ein Bonbon für alle Kinder.*
- *Der Lehrer hat eine Sammelfahrkarte für alle Kinder.*

$\forall x (\text{Kind}(x) \rightarrow \exists y (\text{Bonbon}(y) \wedge \text{BonbonFür}(x,y)))$
 $\exists y (\text{Sammelfahrkarte}(y) \wedge \forall x (\text{Kind}(x) \rightarrow \text{FahrkarteFür}(x,y)))$

Behandlung von Skopus-Ambiguitäten:

- Unterspezifikation
 - Verarbeitung unter Belassung von Mehrdeutigkeiten
 - ggf. Constraints hinzufügen

Viele Studenten aus einigen Fächern wissen wenig aus der Linguistik.

Sprache als Kommunikationsform

- Eingebunden in weitere Formen: Gestik, Mimik.
- Gestaltung: Wörter, Betonung, Prosodie
- Dimensionen:
 - (Textueller) Inhalt
 - Beabsichtigte Wirkung (Sprechakt)
Ist die Tür offen?
Frage, Aufforderung, ...
 - Übermittlung des eigenen Zustandes
(Ausdruck von Emotionen)
Ärger, Angst, Freude, ...

Sprache als Kommunikationsform

- Basiert auf Übereinkunft/Verhandlung
- Ökonomie:
 - Reduktion auf notwendige Information im Kontext
 - allgemeines Hintergrundwissen
 - Annahmen über Partner
 - Diskurs-Bezüge
- Soziale Aspekte: Offenheit (least commitment)
 - Mehrdeutigkeiten
 - Unsicherheiten
 - Verhandlungsführung

Problem der Rekonstruktion.
„Schlecht gestelltes Problem“:
Es wird nicht alles gesagt.

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

37

Sprache als Kommunikationsform

- Fixieren von Bedeutungen:
 - gleiche/beabsichtigte Interpretation erzwingen
 - Vertragsgestaltung
 - Gesetzliche Regelungen
 - Protokolle von Besprechungen
- Offenheit von Bedeutung nutzen
 - „der Nutzer denkt sich sein Teil“
 - Dialogsysteme
 - Kompromisse in Verhandlungen
 - Zensur umgehen

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

38

Bedeutung erfassen

Koordination mit Sprecher:

- Empfänger agiert gleichfalls
(„Embodiment“: mitsprechen)
- Empfänger reagiert unmittelbar.

Bezugspunkte gemäß

- (Textueller) Inhalt *Ist die Tür offen?*
- Beabsichtigte Wirkung *Frage, Aufforderung, ...*
- Übermittlung des eigenen Zustandes
(Ausdruck von Gefühlen) *Ärger, Angst, Freude, ...*

Sprechakte

- Textueller Inhalt + beabsichtigte Wirkung
 - analoge Unterscheidung: Kommando-Eingabe
- Protokoll zwischen Sender/Empfänger
 - Information
 - mögliche Reaktion:
 - Information übermitteln/verweigern
 - Aufforderung
 - mögliche Reaktion:
 - Akzeptieren/Zurückweisen
 - Ausführen
 - Ergebnis übermitteln

Standardisierungen: FIPA- ACL, KQML

Emotionen

- (gesprochene) Sprache übermittelt Gefühle
- Analyse:
 - Anotation bzgl. der Emotionen
- Synthese:
 - Glaubwürdigkeit: Übereinstimmung mit Gefühlsausdruck
 - Prosodie,
 - ggf. Gestik/Mimik

Bedeutungs-Rekonstruktion

Vollständigkeit?

- Mensch rekonstruiert/erfaßt nicht vollständig
- Primitive Dialogsysteme (shallow systems)
 - Eliza
 - Elektron. Auskunftssysteme (beschränkter Bereich)
- Maschinelle Übersetzung
 - Annahmen über Empfänger (Hintergrundwissen/Diskurswissen)
 - I saw the man on the hill with the glasses.
 - aber: richtige Bezüge herstellen

Bedeutung erfassen

- Übermittelter Text
 - Wörter
 - Wortgruppen
 - Zusammenhänge
 - Syntax
 - Semantik
 - Pragmatik
- Kontext, Hintergrund
- Diskurs

„Schlecht gestelltes Problem“:
Es wird nicht alles gesagt.

Modulares Schichtenmodell

- Phoneme bzw. Buchstaben
- Silben
- Wörter
- Wortgruppen
- Sätze
- Satzgruppen
- Text

Spezielle Methoden auf jeder Ebene

Problem:

Abhängigkeit zwischen Schichten

Hypothesen bilden/evaluieren

Segmentierung des akustischen Eingabestroms

- Unsicherheit bei der Erkennung
- keine korrekte Syntax (aber nicht beliebig)
- Performanz-Probleme (ähm, Korrektur,...)

- Erfassung von Sprachmelodie
- Erfassung von Gefühlsausdrücken

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

45

Erfassen von Wortformen/-bedeutungen

- Tagger
 - Annotationen grammatischer Formen
 - mögliche Beziehungen zu anderen Wörtern
- Lexikalische Auswertung
 - mögliche Wortbedeutungen
- Worthypothesengraph

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

46

Analyse von Satzstrukturen

- Syntax-Analyse
 - Grammatik: Baumstruktur gemäß syntaktischem Aufbau (NP, VP, ...)
- Semantik-Analyse
 - Grammatik: Baumstruktur gemäß semantischer Abhängigkeiten
 - Referenzen herstellen
 - Mehrdeutigkeiten reduzieren
 - Interpretation

Auflösen von Mehrdeutigkeiten

- Nach Möglichkeit auf unteren Ebenen
- Auswertung von Constraints
 - Orthographie
 - Grammatik
 - Kontext
 - Dialog-Verlauf (Historie)
 - Domäne
 - Partnermodell
 - Hintergrundwissen

Unterschiedliches Wissen
verarbeiten und kombinieren:
„Experten“/„Agenten“

Auflösung von Mehrdeutigkeiten

Rechtschreibung

Fälle - Felle

Groß- und Kleinschreibung

Der gefangene Floh. Der Gefangene floh.

Grammatik

Helga fragte Bernd nach dem Buch. Er gab es ihr.

Kontext

Bernd fragte Fritz nach dem Eimer. Er gab ihn an ihn.

Allgemeinwissen

Sie sah die Fliege, die die Mutter schlug.

Betonung

übersetzen

Unterschiedliches Wissen
verarbeiten und kombinieren:
„Experten“/„Agenten“

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

49

„Flache“ Verfahren

- Statistische Verfahren, Wahrscheinlichkeiten
z.B. für Korrektur oder Vervollständigung (Voraussage):

$$P(x_1x_2x_3 = \text{die} \mid x_1x_2 = \text{di})$$

$$P(x_1x_2x_3 = \text{die} \mid x_5x_6x_7x_8 = \text{Frau})$$

(alternativ: in „tieferem“ Modell:

Übereinstimmung gemäß Grammatik-Regeln prüfen)

z.B. Für Dokumentenretrieval:

$$P(\text{Dokument } D \text{ passt auf Anfrage } Q \mid D \text{ enthält Zeichenketten } \dots)$$

(Zeichenkette z.B. n-Gramme)

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

50

„Flache“ Verfahren

- Ähnlichkeit

z.B. für Korrektur:

- Ähnlichkeit von Zeichenketten zu bekannten Wörtern
(Zahl der Abweichungen, Buchstabendreher etc.)

sieben ~ siuben

z.B. Für Dokumentenretrieval:

- Dokument D ähnlich zu Anfrage Q:
Als Maß der Übereinstimmung
bzgl. N-Grammen, Schlüsselwörtern, Satzmustern, ...

„Flache“ Verfahren

- Orientierung an Schlüsselwörtern

- Verknüpfung
 - Gruppierung
 - log. Operatoren: und, oder, nicht(?)
- Ähnlichkeit
 - Synonyme, verwandte Begriffe

- Orientierung an Satzmustern

- Prädikate(Verben)
- Füller

- Beschränkte Domänen

„Flache“ Verfahren

- Vorgegebener Wortschatz („Ontologien“)
 - für spezielle Domäne
 - Terminologie
- Vorgegebene Struktur
 - für Dokumente
 - für Dialog
 - Nachfragen
- Vorgegebene Skripts/Regeln
 - ELIZA, Chatterbots

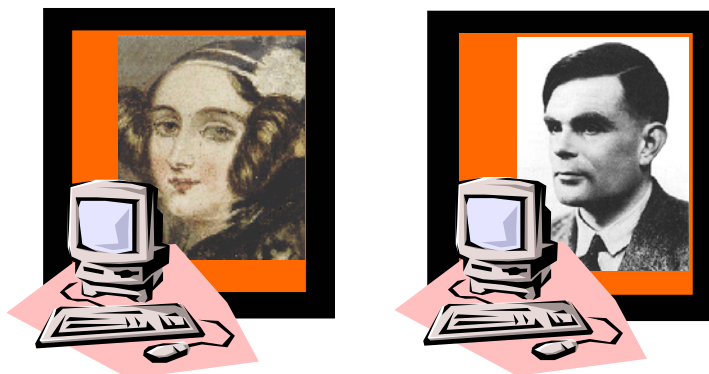
„Tiefe“ Verfahren

- Grammatikalische + inhaltliche Struktur erfassen
 - Komplexe Formalismen notwendig
- Interpretation („Verstehen“)
 - Formalismen für Bedeutungen
 - Formalismen für Hintergrundwissen
 - Partnermodell (Intentionen)
 - Inferenzverfahren
 - ...
 - Bewusstsein? (vgl. Turing-Test)



Wann ist eine Maschine "intelligent"?

- Turing-Test: Wenn Maschine und Mensch nicht unterscheidbar sind



H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

55

„Intelligente“ Maschinen: Eliza

- JF: Die Männer sind alle gleich.
- C: In welcher Weise?
- JF: Sie machen uns immer wegen irgendwas verrückt.
- C: Können Sie ein Beispiel nennen.
- JF: Na ja, die Idee zu diesem Gespräch stammt von meinem Freund.
- C: Die Idee zu diesem Gespräch stammt von Ihrem Freund.
- JF: Er sagt, daß ich oft deprimiert bin.
- C: Es tut mir leid, von Ihnen zu hören, daß Sie deprimiert sind.
- JF: Das stimmt, ich bin unglücklich.
- C: Glauben Sie, daß dieses Gespräch Ihnen hilft, nicht mehr unglücklich zu sein?
- JF: Ich brauche jedenfalls Hilfe, das steht fest.
- C: Was würde es für Sie bedeuten, Hilfe zu bekommen?
- JF: Vielleicht könnte ich lernen, mit meiner Mutter auszukommen.
- C: Erzählen Sie mir mehr über Ihre Familie.
- JF: Meine Mutter erledigt alles für mich.

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

56

„Intelligente“ Maschinen: Eliza

- C: Wer in Ihrer Familie erledigt noch alles für Sie?
- JF: Mein Vater.
- C: Ihr Vater.
- JF: Sie sind ein wenig wie mein Vater.
- C: Welche Ähnlichkeiten sehen Sie?
- JF: Sie sind nicht besonders aggressiv, aber ich glaube, Sie möchten nicht, daß ich das beachte.
- C: Was läßt Sie vermuten, daß ich nicht sehr aggressiv bin?
- JF: Sie streiten nicht mit mir.
- C: Wieso glauben Sie, daß ich nicht mit Ihnen streite.
- JF: Sie haben Angst vor mir.
- C: Macht Ihnen der Gedanke Spaß, daß ich vor Ihnen Angst habe?
- JF: Mein Vater hat vor jedem Angst.
- C: Was fällt Ihnen noch ein, wenn Sie an Ihren Vater denken?
- JF: Angeber.
- C: Hat das etwas damit zu tun, daß die Idee zu diesem Gespräch von Ihrem Freund stammt?

H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

57

„Intelligente“ Maschinen: Eliza

- JF: Sie xxxxx nicht mit mir.
- C: Wieso glauben Sie, daß ich nicht mit Ihnen xxxxx.
- xxxxx = streiten
- JF: Sie streiten nicht mit mir.
- C: Wieso glauben Sie, daß ich nicht mit Ihnen streite.
- JF: Sie xxxxx nicht mit mir.
- C: Möchten Sie, daß ich mit Ihnen xxxxx.
- oder:
- C: Haben Sie erlebt, daß jemand mit Ihnen xxxxx.
- JF: Ich fühle mich heute xxxxx.
- C: Passiert es oft, daß Sie sich xxxxx fühlen?

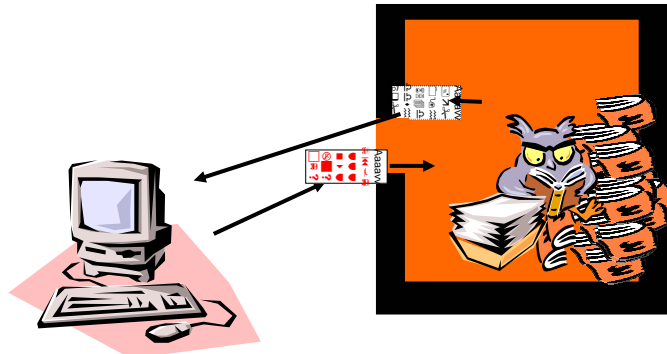
H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

58

Wann ist eine Maschine "intelligent"?

- Kritik am Turing-Test:
- Das chinesische Zimmer (Searle)



H.D.Burkhard, D.Göhring
Winter-Semester 2005/06

Vorlesung Einführung in die KI
Sprachverarbeitung

59