

Einführung in die Computerlinguistik

Dialogsysteme

WS 2010/2011
Manfred Pinkal

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Wozu Dialogsysteme?

„Physische“ Motivation:

- Gerätebedienung für Behinderte: Aufzug, Rollstuhl, Fahrkartenautomaten, ... („barrierefreie Umgebung“)
- Gerätebedienung: Tätigkeiten, bei denen Hände und Augen für andere Aufgaben benötigt werden: Autofahren, ärztliche Diagnose, Steuerung von Instrumenten in der Chirurgie, Wartung und Inventarisierung, ... („hands-free“/„eyes-on“ Anwendungen)
- Telefonie: Anwendung, in denen nur der akustische Kanal zur Verfügung steht (Fahrplan-Auskunft, Kundendienst, Telebanking, ...)

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Der sprechende Fahrstuhl (1)

- User: Fahrstuhl
- System: In welches Stockwerk möchten Sie bitte?
- User: Vier / Vierter Stock
- System: <Kommando-Ausführung>

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Der sprechende Fahrstuhl (5)

- U: Fahrstuhl
- S: In welches Stockwerk möchten Sie bitte?
- U: Zu Professor Möbius.
- S: Ich fahre Sie in den vierten Stock.
- S: <Kommando-Ausführung>

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Wozu Dialogsysteme?

„Semantische“ Motivation:

- Natürliche Sprache ist attraktiv, weil sie das natürliche Kommunikationsmedium des Menschen ist.
- Wir können auf dasselbe Objekt/ dieselbe semantische Information mit unterschiedlichsten sprachliche Mitteln zugreifen.
- Dialogschnittstellen ermöglichen dem Benutzer, seine Absichten ohne Spezialwissen und ohne das Auswendiglernen von Sprachkommandos zu realisieren:

Der Benutzer sagt, *was er will, so, wie er es will (und kann)*.

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Der sprechende Fahrstuhl (2)

- U: Fahrstuhl
- S: In welches Stockwerk möchten Sie bitte?
- U: Vierter Stock
- S: Ich fahre Sie in den vierten Stock.
- S: <Kommando-Ausführung>

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Semantische Motivation

- Sprachdialog für Nicht-Experten zur Interaktion mit komplexen technischen Umgebungen (Auto, Fahrkartenautomat)
- Intelligentes Haus: Unterstützt ältere Menschen dabei, in einer immer technischer werdenden Alltagsumgebung lange selbständig zu bleiben.
- Service-Roboter im privaten Haushalt und in Pflege-Einrichtungen

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Unsichere Spracherkennung

- Automatische Spracherkennung (Automatic speech recognition – ASR) ist nicht hundertprozentig verlässlich. Die Verlässlichkeit ist abhängig von einer Anzahl von Faktoren (Stimme, Sprechweise, physikalische Umgebung, Hintgrundgeräusche, ...).
- Aber: Sprachdialog hat verschiedene Möglichkeiten, die grundsätzliche Erkennungproblematik zu kompensieren.

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Der sprechende Fahrstuhl (3)

- U: Fahrstuhl
- S: In welches Stockwerk möchten Sie bitte?
- U: Vierter Stock
- S: Ich habe Sie nicht verstanden. Bitte sagen Sie die Etagen-Nummer!
- U: Vier
- S: Ich fahre Sie in den vierten Stock.
<Kommando-Ausführung>

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Der sprechende Fahrstuhl (4)

- U: Fahrstuhl
- S: In welches Stockwerk möchten Sie bitte?
- U: Vierter Stock
- S: Ich fahre Sie in den fünften Stock.
- U: Nein, vierter Stock!
- S: Ich fahre Sie in den vierten Stock.
<Kommando-Ausführung>

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Methoden zur Verständnissicherung

- Unterstützung der Spracherkennung durch Information über erwartete Benutzereingabe: kontextspezifische Sprachmodelle, "Erkennungsgammataken".
- Erkennungfehler werden teilweise kompensiert durch:
 - Klärungsfragen (des Systems)
 - Korrekturen (durch den Benutzer)
 - Implizite Verständigungssicherung (engl. "grounding")
- Spracherkennung liefern auch **Konfidenzwerte**, die gezielte Klärungsstrategien erlauben:
 - Akzeptieren
 - Explizite Rückfrage
 - Implizites Grounding

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Kooperative Klärung auf Inhaltsebene

- U: Fahrstuhl
- S: In welches Stockwerk möchten Sie bitte?
- U: Zur Phonetik.
- S: Möchten Sie in den vierten oder in den fünften Stock?
- U: Zu Professor Möbius.
- S: Ich fahre Sie in den vierten Stock.
- S: <Kommando-Ausführung>

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Kooperative Klärung auf Inhaltsebene

- U: Fahrstuhl
- S: In welches Stockwerk möchten Sie bitte?
- U: Zu Professor Seidel
- S: Meinen Sie Raimund Seidel oder Hans-Peter Seidel?
- ...

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Kooperative Klärung auf Inhaltsebene

- U: Fahrstuhl
- S: In welches Stockwerk möchten Sie bitte?
- U: Zu Professor Crocker.
- S: Das Büro von Professor Crocker ist im Gebäude C7 1. (Soll ich Sie ins Erdgeschoss fahren?)
<Kommando-Ausführung>

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Kooperative Klärung auf Inhaltsebene

- Disambiguierung durch Rückfragen/ Klärungsdialoge
- Kooperative Präzisierung unterspezifischer Angaben
- Kooperative Überbeantwortung von Benutzerturns

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Eine weitere Herausforderung

© 1999 Randy Glasbergen.
www.glasbergen.com



" ... Wenn Sie alle Optionen noch einmal hören wollen, sagen Sie 49.
Wenn Sie inzwischen vergessen haben, warum Sie angerufen hatten,
sagen Sie 50. "

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Dialogdesign

- Bei komplexen Aufgabenstellungen haben Benutzer Probleme, den Überblick zu behalten: Beschränktes Gedächtnis, Dialogstruktur ist nicht sichtbar (im Gegensatz zur menübasierten Interaktion am Display)

Aber:

- Der Dialog kann mit sprachlichen Mitteln strukturiert und für den Benutzer nachvollziehbar gestaltet werden: (z.B.: Zerlegung einer Auswahl zwischen vielen Alternativen in eine Folge von Auswahlritten, bei der die Zahl der Alternativen überschaubar ist)
- Der Dialog kann durch Schrift oder Grafik ergänzt oder unterstützt werden. („Multimodaler Dialog“)
- **Dialogdesign** ist eine zentrale Aufgabe.

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UoS Computerlinguistik

Dialogsysteme der zweiten Generation

- Technische Funktionalität **plus** Benutzerzufriedenheit (**Design!**)
- **Inhaltsorientierung**, flexible Eingabe
- **Kooperative Klärung** von Inhalten, kooperative Überbeantwortung
- Interaktive, kontextsensitive **Hilfe**
- **Initiativwechsel**: Flexible Dialogstruktur, Dialoginitiative wechselt zwischen Benutzer und System

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UoS Computerlinguistik

Dialogsysteme der ersten Generation

- „Command and Control“
- Stark eingeschränktes Vokabular (Schlüsselworte), system-initiierte und kontrollierte Dialogführung
- Optimierung auf Robustheit und Effektivität (wegen schwacher Erkennerrqualität)
- Forschungssysteme seit den achtziger, kommerzielle Systeme seit Mitte der neunziger Jahre

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UoS Computerlinguistik

Dialogsysteme der zweiten Generation

- **Adaptivität**:
 - Das System passt sich an Kenntnisse, Fähigkeiten und Präferenzen des Benutzers an.
 - Adaption an Kontext und Dialogsituation (im Fahrzeug: Ablenkung durch Verkehrsgeschehen!)
 - Anpassung unter anderem in Äußerungslänge und Sprechtempo, Wortwahl, Syntax, Wahl der Ausgabe-Modalität.

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UoS Computerlinguistik

NaDia I



Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik



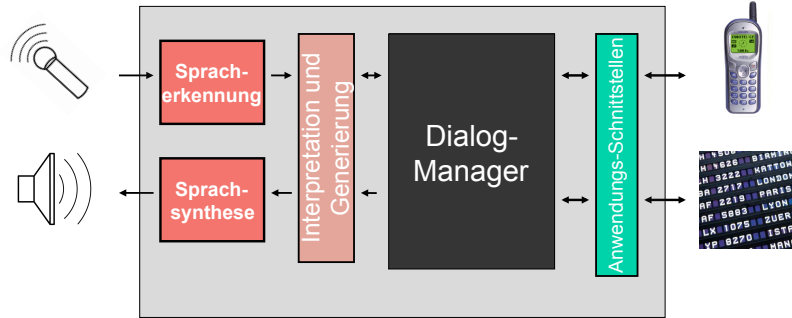
Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

TALK: Dialog im Fahrzeug



Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Architektur von Dialogsystemen



Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Sprachsynthese

- Ansätze:
 - Einzellautsynthese (unmöglich)
 - Voraufgenommene Sprache (unpraktisch)
 - Diphon-Synthese
 - Wortkonkatenation
 - Unit Selection

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Linguistische Verarbeitung: Analyse

- Interpretation:
 - Schlüsselwort- / Schlüsselphrasenerkennung oder:
 - Grammatik in Form von endlichen Automaten/ sehr flache CFGs/ BNF
 - Interpretationsgrammatiken: Aufbau von einfachen Merkmalsstrukturen/ Records mithilfe [semantischer Tags](#)

- Ein Beispiel:

Regel: S -> ich möchte [speech_act = request]
 SONG_TITLE ([song_title=SONG_TITLE]) von \$GARB
 ARTIST ([artist=ARTIST] \$GARB*

Eingabe: Ich möchte Yesterday von den Beatles hören

Ausgabe: [speech_act = request, song_title=yesterday,
 artist = beatles]

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

Linguistische Verarbeitung: Generierung

Generierung ist typischerweise Template-basiert

- Beispiel-Template

```
Confirmation_question_song:
Sie möchten also ${1} hören
```

- Beispiel-Aufruf:

```
Confirmation_question_song; Yesterday
```

- Ausgabe:

```
Sie möchten also Yesterday hören
```

Vorlesung "Einführung in die CL" 2010/2011 © M. Pinkal UdS Computerlinguistik

