

Einführung in die Computerlinguistik

Verarbeitung gesprochener Sprache

WS 2012/2013

Manfred Pinkal

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Wozu sind Dialogsysteme gut?

"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" 2012/2013 © M. Pinkal UoS Computerlinguistik

Beispieldialog (1)

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

- Etwas pragmatische Terminologie:
 - **Turn**: Einzelner Dialogbeitrag eines Sprechers
 - **Sprechakt/ Dialogakt**: Sprech-/Dialoghandlungstyp, z.B. "Informationsfrage", "Mitteilung"

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Beispieldialog (2)

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

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Gesprochene Sprache

- Für die Interaktion zwischen Menschen und zwischen Mensch und Maschine stehen grundsätzlich unterschiedliche "Modalitäten" zur Verfügung (visuell, akustisch, etc.)
- Gesprochene Sprache
 - hilft in allen Situationen, in denen Hände ("haptische" Modalität) und Augen (visuelle Modalität) nicht zur Verfügung stehen
 - bringt das Problem unzuverlässiger Spracherkennung mit sich.
- Ein wichtiger Aspekt aller Sprachdialogsysteme: Der Umgang mit Erkennenfehlern.
- Zwei wesentliche Strategien:
 - Sichere(re) Spracherkennung durch Einschränkung des Sprachumfangs (handkodierte **Erkennungsgesetze**)
 - Erkennenfehler kompensieren durch **Verständigungssicherung** ("Grounding")

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Klärungsfrage

- U: **Fahrrad**
- S: **In welches Stockwerk möchten Sie bitte?**
- U: **Vierter Stock**
- S: **Ich habe Sie nicht genau verstanden. Wollen Sie in den dritten Stock?**
- U: **Nein, in den vierten!**
- S: **Ich fahre Sie in den vierten Stock.**
<Kommando wird ausgeführt>

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Benutzerkorrektur

- U: **Fahrrad**
- 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 wird ausgeführt>

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Spracherkennung

- Erkennenfehler kompensieren durch Verständigungssicherung ("Grounding")
 - **Klärungsfragen** des Systems (explizites Grounding)
 - **Implizite Verständigungssicherung**, die Gelegenheit zur Korrektur durch den Benutzer gibt (implizites "grounding")
- Dialogstrategien:
 - **Optimistische Strategie**: Verzicht auf Grounding – im Erfolgsfall effizient, bei Missverständnis Effektivitäts- und/oder Effizienzproblem
 - **Vorsichtige Strategie**: Durchgängig (explizites) Grounding – maximiert Effektivität, aber auf Kosten von Effizienz und Benutzerzufriedenheit
 - **Flexible Strategie**: Je nach **Konfidenz** (Wahrscheinlichkeit der besten Erkennungshypothese) kein Grounding/ implizites Grounding/ explizite Rückfrage

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Beispieldialog (5)

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

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik

Wozu sind Dialogsysteme gut?

Die zweite Motivation ist „semantisch“:

- 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 inhaltsorientiert und intuitiv, ohne Spezialwissen und ohne das Auswendiglernen von Sprachkommandos zu realisieren:
Der Benutzer sagt, *was er will, so, wie er es will.*

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik

Wozu sind Dialogsysteme gut?

Intuitive Gerätebedienung:

- Sprachdialog für Nicht-Experten zur Interaktion mit komplexen technischen Umgebungen (Auto, Fahrkartenautomat)
- Intelligentes Haus: Unterstützt ältere Menschen, in einer immer technischer werdenden Alltagsumgebung lange selbständig zu bleiben.

Künstliche Systeme als Partner:

- Service-Roboter im privaten Haushalt und in Pflege-Einrichtungen

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik

Inhaltsorientierter Dialog: Herausforderungen

- Der Benutzer sagt, *was er will, so, wie er es will – und kann.*
- Problem: Benutzeräußerung kann unpräzise, mehrdeutig sein oder auf falschen Voraussetzungen beruhen.
- Semantische Verständigungssicherung:
 - Klärungsdialoge
 - Semantische Korrektur der Benutzereingabe

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik

Klärungsdialog

- 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 wird ausgeführt>

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik

Korrektur des Benutzers

- U: Fahrstuhl
- S: In welches Stockwerk möchten Sie bitte?
- U: Zu Professor Crocker.
- S: Das Büro von Professor Crocker ist nicht in diesem Gebäude.
<Kommando wird ausgeführt>

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik

Kooperative Überbeantwortung

- 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 wird ausgeführt>

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik

Qualitätskriterien für Dialogsysteme

- Effektivität:
 - Der Dialog führt zum (gewünschten) Ergebnis
- Effizienz:
 - Der Dialog führt in angemessener Zeit/ in einer angemessenen Anzahl von Dialogschritten zum Ergebnis
- Benutzerzufriedenheit

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik

Die Deutsche Bahn



Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

NaDia I – Demonstrator für BMW



Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik



Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik



Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Dialogdesign



" ... 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" 2012/2013 © M. Pinkal UoS Computerlinguistik

Command and Control

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

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS 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)

Lösungen:

- **Strukturierung** des Dialogs, z.B.: Zerlegung einer Auswahl zwischen vielen Alternativen in eine Folge von Auswahlritten, bei der die Zahl der Alternativen überschaubar ist
- **Multimodaler Dialog**: Gesprochene Sprache + Haptische Eingabe + visuelle Ausgabe
- **Dialogdesign** ist eine zentrale Aufgabe.

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

"Natürlicher " Sprachdialog

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

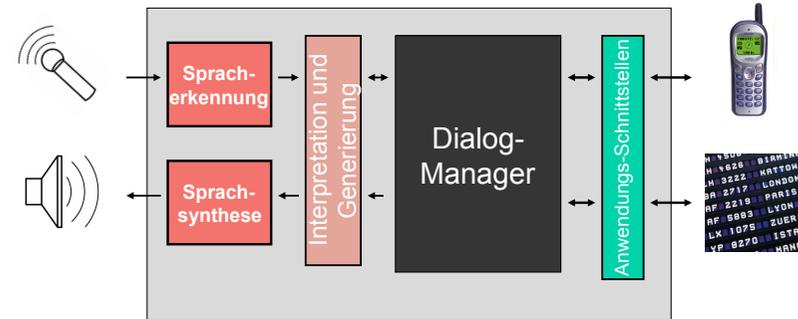
Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Adaptive Dialogsysteme

- **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" 2012/2013 © M. Pinkal UoS Computerlinguistik

Architektur von Dialogsystemen



Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Sprachsynthese

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

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Linguistische Verarbeitung: Analyse

- **Interpretation:**
 - Schlüsselwort- / Schlüsselphrasenerkennung oder:
 - Grammatik in Form von endlichen Automaten/ sehr flache CFGs/ BNF
 - Interpretationsgrammatiken: Aufbau von einfachen Merkmalsstrukturen mithilfe **semantischer Tags**
- **Ein Beispiel:**

```

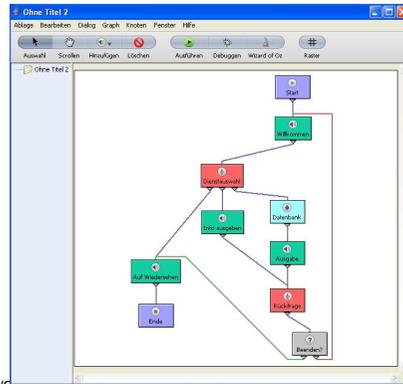
Regel: S -> ich möchte [speech_act = request]
      SONG_TITLE ([song_title=SONG_TITLE]) von $GARB
      ARTIST ([artist=ARTIST] $GARB*
Eingabekette: Ich möchte Yesterday von den Beatles hören
Ausgabe: [speech_act = request, song_title=yesterday,
          artist = beatles]
    
```

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

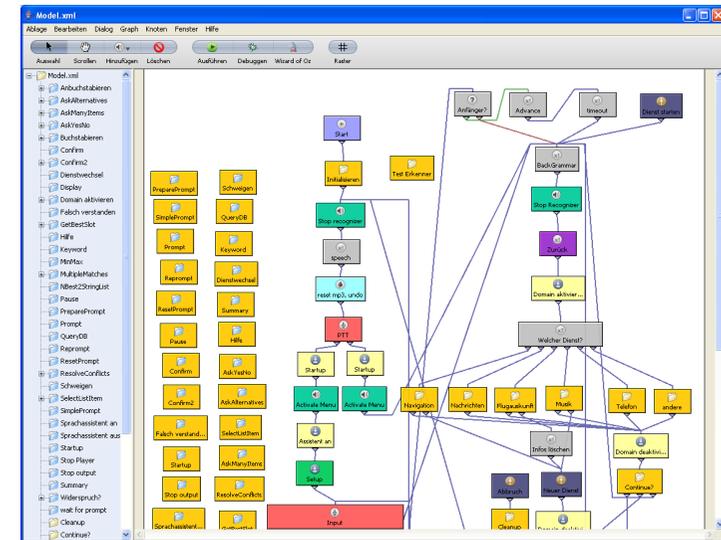
Dialogmodellierung

- Standard-Framework: Modellierung der Abfolge von Dialogakten durch endliche Automaten

- DialogOS-Screenshot:



Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik



Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Wizard-of-Oz Studien

- Ein verborgener menschlicher Dialogpartner ("wizard") simuliert einen Teil eines Dialogsystems.
- Die Versuchspersonen glauben, dass es sich um ein Computersystem handelt.

Vorzüge von WoZ-Systemen

- System-Design kann ohne teure Implementierung früh erprobt und verbessert werden.
- Kontrolle über die Performanz der (simulierten) Spracherkennung
- Datensammlung für
 - das Training von Spracherkennern
 - die qualitative Analyse des Nutzerverhaltens

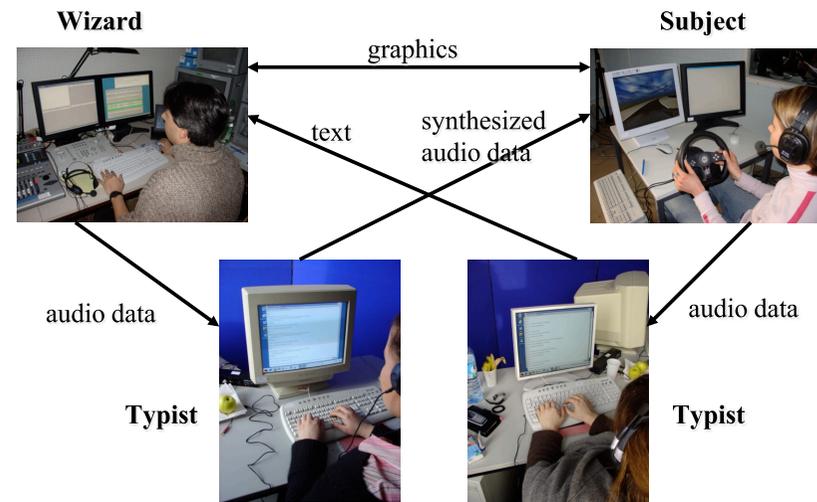
Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UoS Computerlinguistik

Wizard-of-Oz Studien

- Ein überzeugendes WoZ-System benötigt künstliche Sprachausgabe: Entweder vom Wizard durch Menü angestoßen, oder: Eintippen + Sprachsynthese
- Erkennenfehler müssen (und können kontrolliert) simuliert werden: Die Benutzereingaben werden für den Wizard mit Geräuschen überlagert, oder der Wizard erhält getippte Version der Benutzereingabe, die gezielt verfälscht wird.
- Die natürliche kommunikative Kompetenz des Wizards muss eingeschränkt werden: Ausgabemenüs, vorgegebene Ausgabe-Templates
- Effizienz von Computersystemen bei Datenbankzugriff, Berechnungen muss in Form einfach bedienbarer Schnittstellen für den Wizard bereitgestellt werden.

Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik



Vorlesung "Einführung in die CL" 2012/2013 © M. Pinkal UdS Computerlinguistik