

Einführung in die Computerlinguistik

Dialogsysteme

WS 2013/14

Manfred Pinkal

Beispieldialog

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

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, ...)

Beispieldialog

- 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>

Wozu sind Dialogsysteme gut?

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

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.

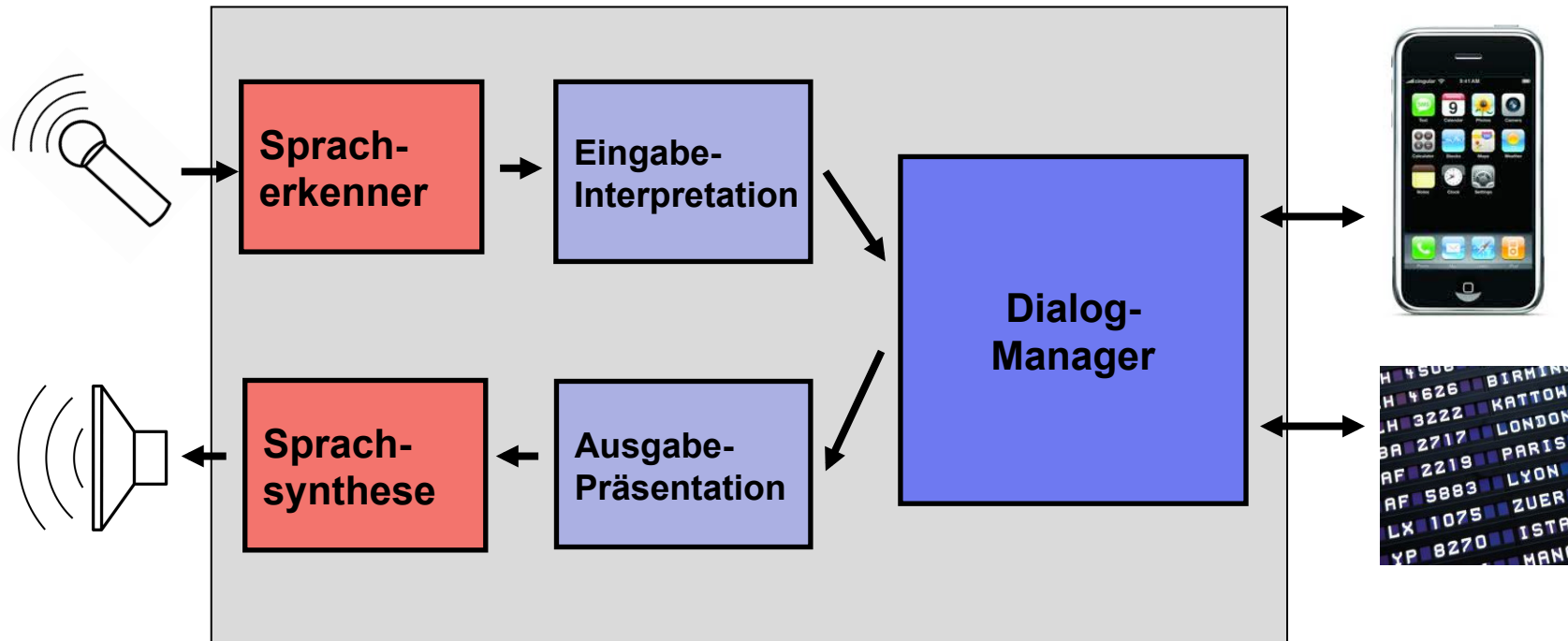
„Companion-Technologie“

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

Etwas pragmatische Terminologie

- **Turn:** Einzelner Dialogbeitrag eines Sprechers
- **Sprechakt:** Sprechhandlungstyp, z.B. "Mitteilung", "Aufforderung", "(Selbst)-Verpflichtung"
- **Dialogakt:** Feinere Klassifikation von Dialoghandlungen, z.B. "Informationsfrage", "Klärungsfrage"

Dialogsysteme: Standardarchitektur



Spracherkenner

- Statistische Sprachmodelle, trainiert auf
 - großen Textkorpora
 - Verschriftungen (Transliterationen) von Dialogkorpora
- Ein Problem ist (immer noch) die Unzuverlässigkeit der Spracherkennung
- Eine Alternative: Handkodierte Sprachmodelle: „Erkennergrammatiken“

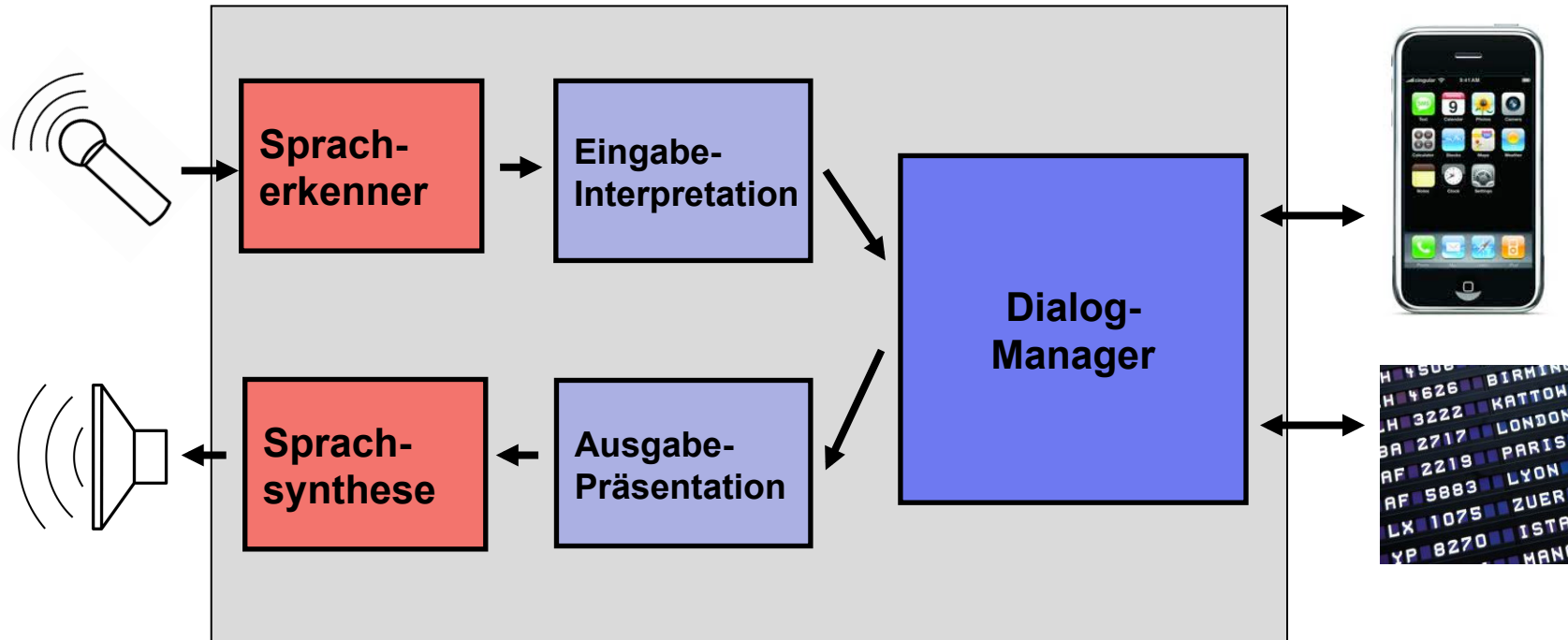
Erkennergrammatiken

- Üblicherweise in BNF (Backus-Naur-Format) geschrieben, einer Variante von CFGs
- Typischerweise extrem flache “semantische Grammatiken” ohne Rekursion
- Können ohne Informationsverlust in endliche Automaten umgewandelt werden.
- Beispiel:
 - `S` → ich möchte `SONG_TITLE` (von `ARTIST`) hören
 - `SONG_TITLE` → yesterday | ...
 - `ARTIST` → (den) beatles | ...

Erkennergrammatiken

- Erkennergrammatiken erlauben die komfortable Spezifikation von Sprachmodellen, die auf eine Domäne und einen bestimmten Dialogzustand zugeschnitten sind.
- Damit wird die Erkennererwartung drastisch reduziert.
 - Beispiel: `$yn_answer = ja | nein`
- Gleichzeitig wird der Sprachumfang für den Erkenner drastisch eingeschränkt.
- Ein Ausweg: Arbeiten mit Wildcards:
 - Regel: `S -> ich möchte $GARB* SONG_TITLE (von) ARTIST $GARB*`
 - `SONG_TITLE` → yesterday | ...
 - `ARTIST` → (\$GARB) beatles | ...
- Relevante Information wird erkannt, aber die Erkennungsqualität sinkt.

Eingabe-Interpretation



Erkennergrammatiken mit Interpretations-Tags

- Beispielgrammatik

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

```
SONG_TITLE → yesterday [Yesterday] | ...
```

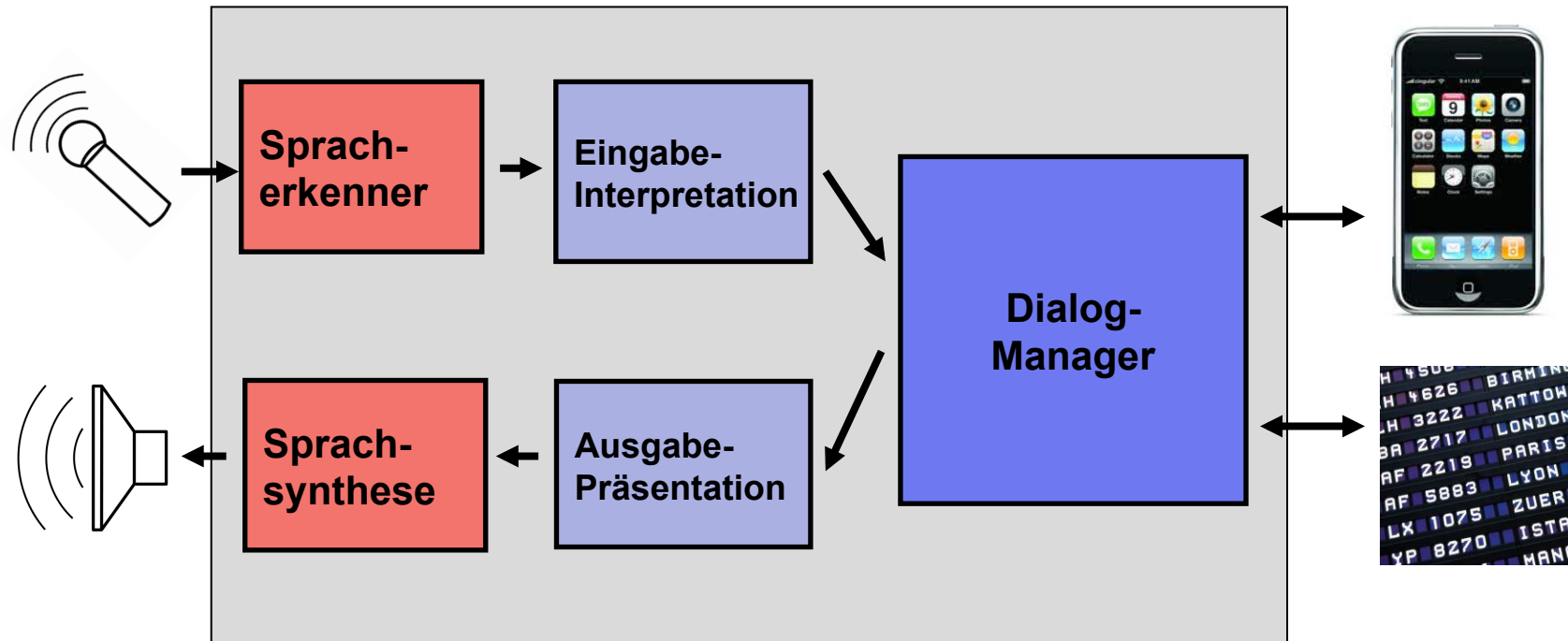
```
ARTIST → ($GARB) beatles [Beatles] | ...
```

- Eingabekette: Ich möchte Yesterday von den Beatles hören
- Ausgabe: [speech_act = request,
song_title=yesterday, artist = beatles]
- Erkennergrammatiken mit Interpretationstags haben zwei Funktionen:
 - Spezifikation des Sprachmodells
 - Erzeugung einer semantischen Repräsentation als Eingabe für den Dialog-Manager

Interpretationsgrammatiken für statistische Sprachmodelle

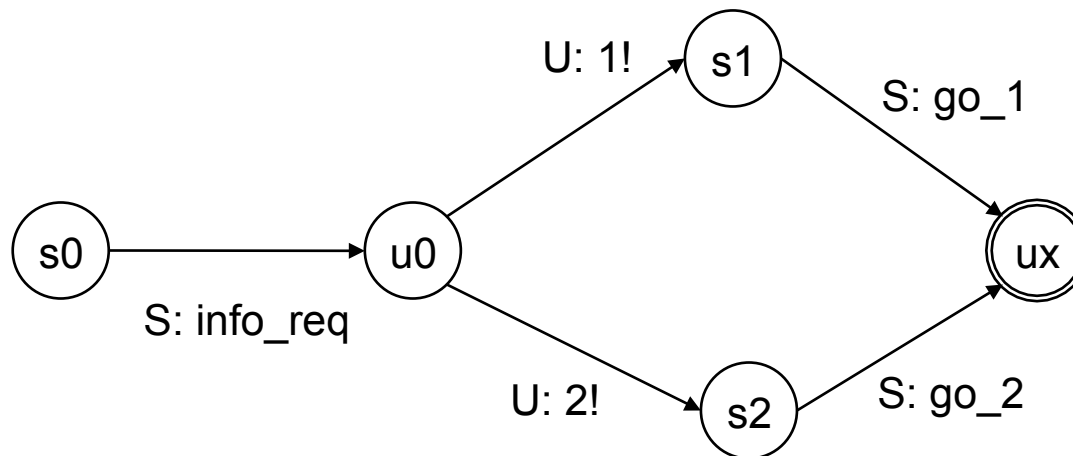
- Interpretationsgrammatiken für statistische Sprachmodelle sehen aus wie erweiterte Erkennergrammatiken.
- Sie funktionieren anders: Sie parsen Wortketten, die der Spracherkenner liefert.
- Problem: Typischerweise tauchen unbekannte Wörter (OOV „Out of vocabulary“) und unbekannte Strukturen auf (OOG: „Out of Grammar“).

Dialog-Management





Dialogmodelle als endliche Automaten



Benutzerkorrektur

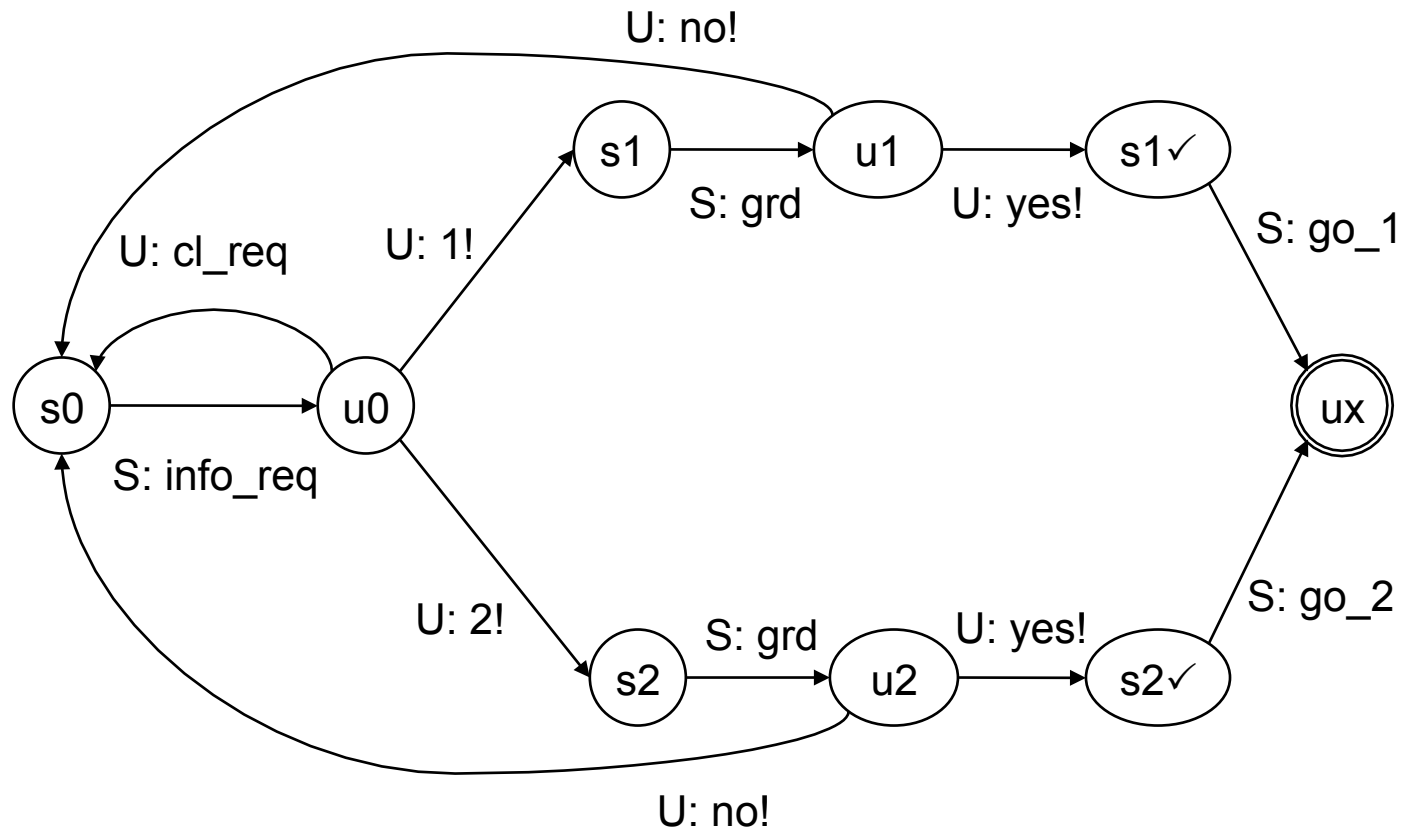
- U: Fahrstuhl
- S: In welches Stockwerk möchten Sie bitte?
- U: Vierter Stock
- S: Ich fahre Sie in den fünften Stock.
- U: Nein, vierter Stock!
- S: Ich fahre Sie in den vierten Stock.
<Kommando wird ausgeführt>

Klärungsfrage

- U: Fahrstuhl
- 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>

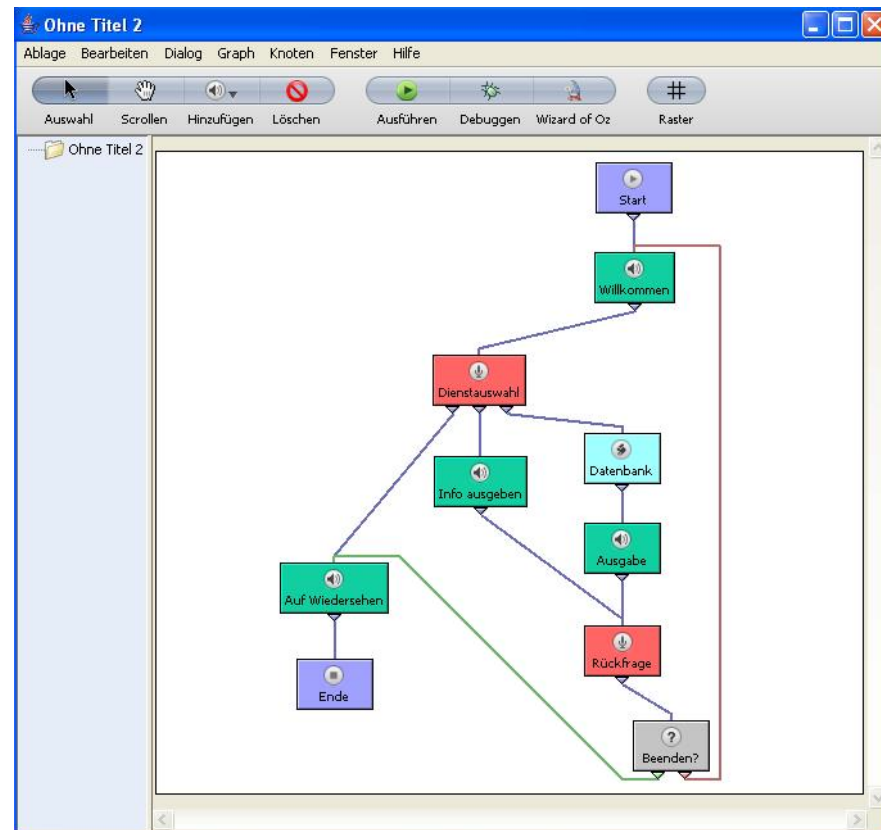


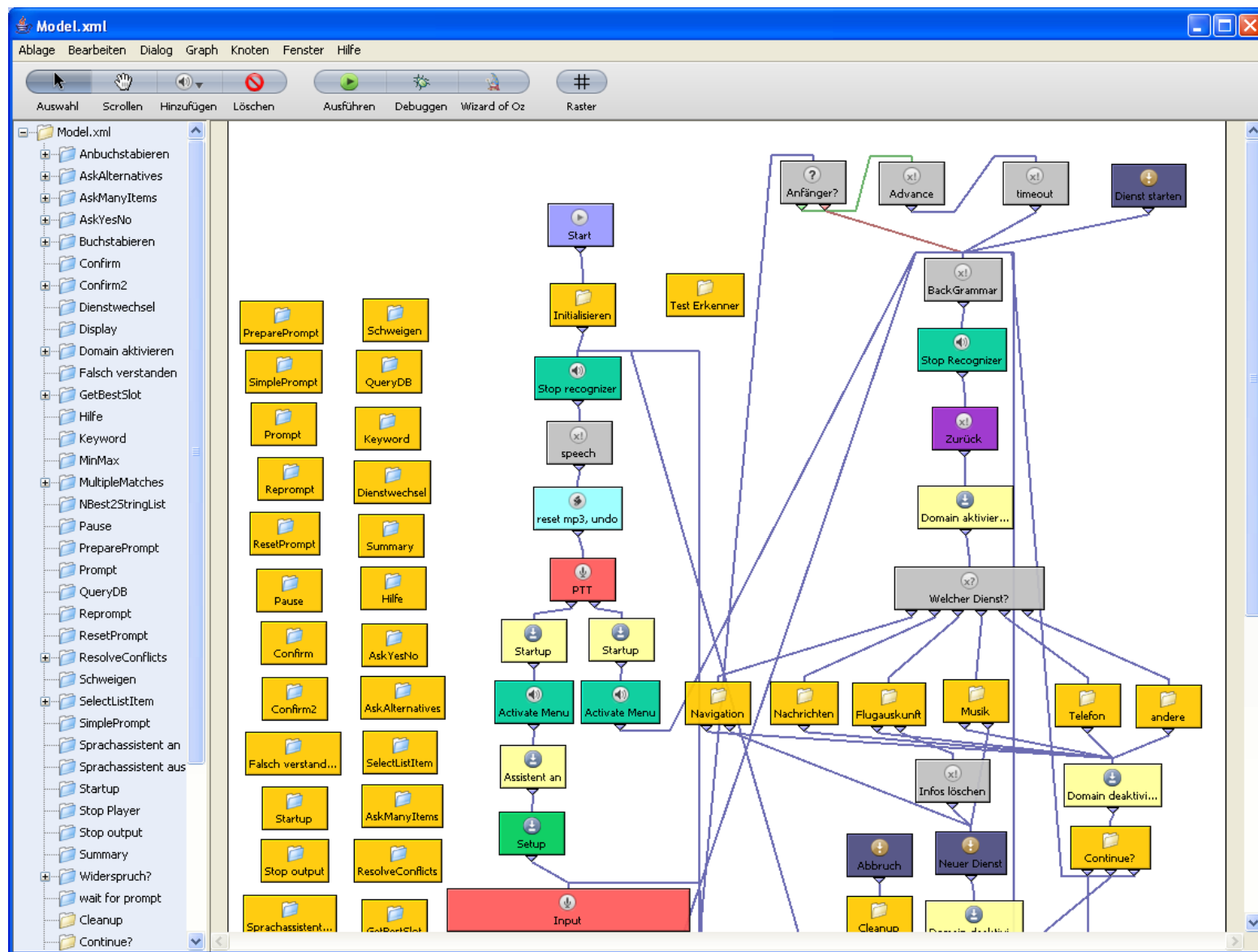
Dialogmodell mit Grounding



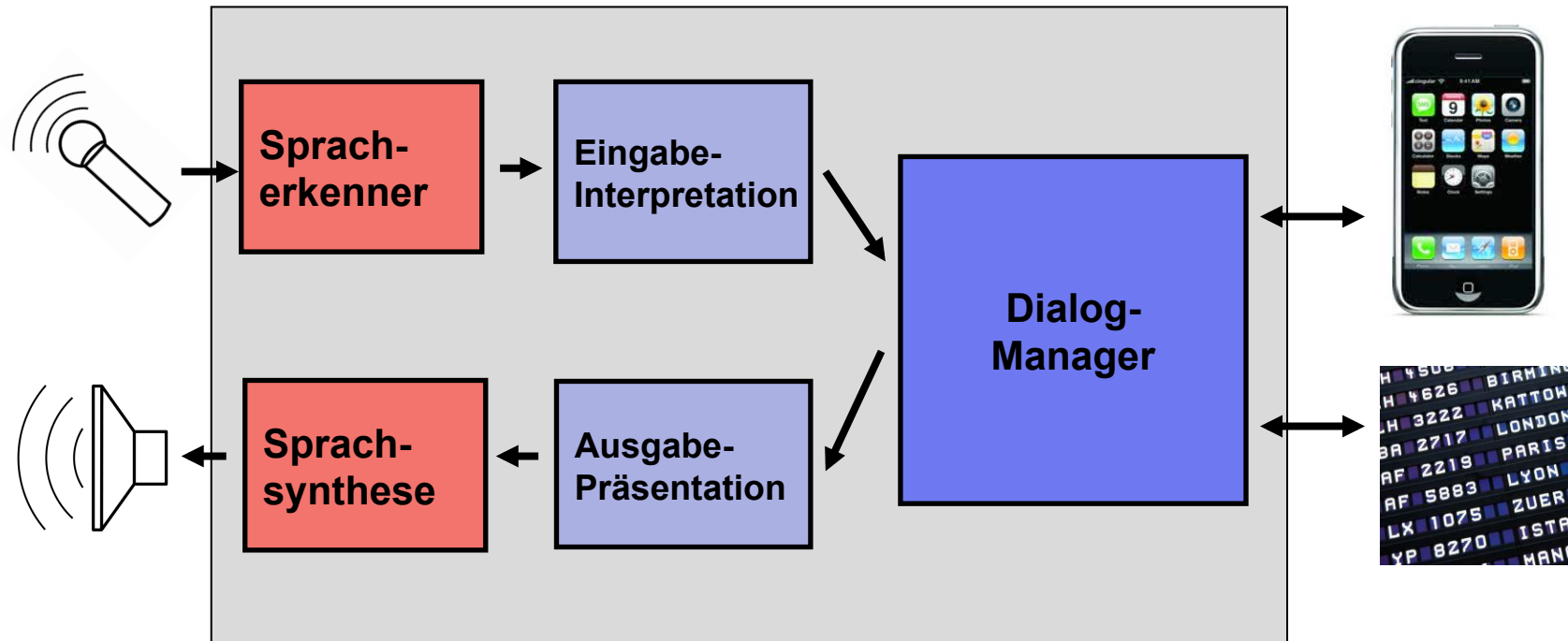
Ein Werkzeug für die Dialogmodellierung

- DialogOS-Screenshot:





Dialog-Management



Ausgabe-Präsentation und Sprachsynthese

- Template-basierte Generierung der Sprachausgabe:
 - Der nächste Flug nach **\$AIRPORT** geht um **\$DAYTIME**.
- Sprachsynthese: Standard ist Unit Selection.

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

NaDia I – Demonstrator für BMW





Vorlesung "Einführung in die CL" 2013/2014 © M. Pinkal UdS Computerlinguistik



Vorlesung "Einführung in die CL" 2013/2014 © M. Pinkal UdS Computerlinguistik