Language Technology 2:
Natural Language(-based) Interaction

www.coli.uni-saarland.de/courses/late2-09

Manfred Pinkal, Magdalena Wolska

dialog systems

today

dialog systems
- example tasks, modes of interaction
- sub-systems
- dialog management and dialog models

dialog systems

goal-oriented conversational systems

challenges:

- need to understand
- interpretation context-dependent
- intention recognition
- anaphora resolution
- people don’t talk in sentences...
- user’s self-revisions

dialog systems

goal-oriented conversational systems

how:

- interactions in a limited domain
- prime users to adopt vocabulary the system knows
- partition interaction into manageable stages
- let the system take the initiative (predictability)

dialog systems

example tasks:

- retrieve information → information-seeking dialogue
- seek to satisfy constraints → negotiation dialogue
- perform action → command-control dialog
- collaborate on solving a problem → problem-solving dialog
- instruct → tutorial/instructional dialogue

applications:

- travel arrangements, telephone directory
- customer service, call routing
- tutoring
- communicating with robots
- voice-operated devices

dialog systems: travel arrangements (Communicator)

S1:	Hello. You've reached the [Communicator]. Tell me your full name
U2:	*hi i'd like to fly to seattle tuesday morning*
S3:	Traveling to Seattle on Tuesday, August eleventh in the morning. Your full name?
U4:	*[John Doe]*
S5:	Hello, [John], I'm glad to hear from you again. You can now ask me for help.
S6:	OK. Let's see I have a United flight departing Pittsburgh at ten forty five a.m. arriving into Seattle at one fifty seven p.m. Is that OK?
U7:	*Yes*
S8:	Will you return to Pittsburgh from Seattle?
U9:	*no i will go to Los Angeles*
S10:	Traveling to Los Angeles. on what date
U11:	*Sorry can we go back to talk about the previous leg again*
S12:	For the first leg, you are on a United flight 1446 departing Pittsburgh at ten forty five A.M. tomorrow, arriving in Seattle at one fifty seven P.M. Do you want to take that?
U13:	*actually i want to leave on wednesday*
S14:	a flight on wednesday.
S15:	I have a flight departing Pittsburgh at ten forty five a.m., arrives Seattle at one fifty seven p.m. Is that OK?
U16:	*Yes*

Figure 19.1 The travel domain: a fragment from a successful conversation between a user (U) and the Communicator system (S) of Xu and Rudnicky (2000).

dialog systems: call routing (ATT HMIHY)

S:	How may I help you?
U:	*can you tell me how much it is to tokyo?*
S:	You want to know the cost of a call?
U:	*yes that's right*
S:	Please hold on for rate information.

S:	How may I help you?
U:	*what's the area code for chicago?*
S:	You want an area code.
U:	*(silence)*
S:	The area code for Chicago is 312.

Figure 19.2 The call routing/telephone information domain: two actual conversations between a user (U) and the AT&T How May I Help You (HMIHY) system (S) (Gorin et al., 1997).

dialog systems: tutorial dialog (ITSPOKE)

S: I didn't understand. What force acts on all objects within close proximity to earth?
U: *uh the force of gravity*
S: Fine. Besides the Earth's gravitational force, are there any other forces acting on the person, keys, and elevator after the release?
U: *no*
S: Yep. Thus, keep in mind that after the release, the only force on the person, keys, and elevator is the force of gravity. If gravity is the only force acting on an object, will it be moving or staying still?
U: *it will be moving*
S: Fine. Not only are the person, keys, and elevator moving, they have only gravitational forces on them. When an object is falling and has only gravitational force on it, it is said to be in what?
U: *freefall*

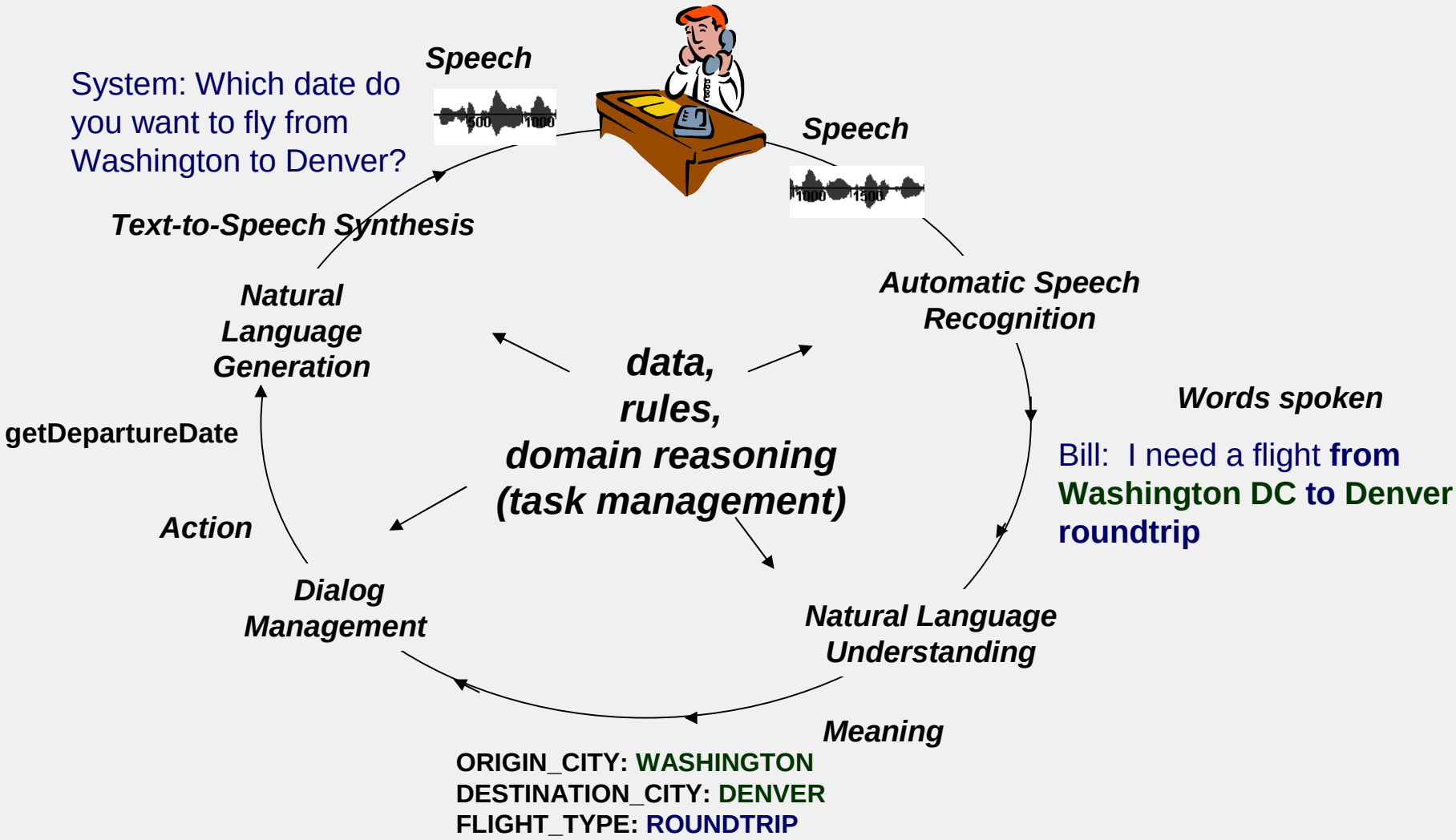
Figure 19.3 The tutoring domain: part of an actual conversation between a student user (U) and the ITSPOKE system (S) of (Litman and Silliman, 2004), based on the Why2-Atlas text-based tutoring system (?).

dialog systems

modality: type of communication channel used to convey or acquire information

natural-language: spoken or textual keyboard-based or both
pointing devices
graphics, drawing
gesture
combination of one of more of above (multi-modal systems)

dialog systems



dialog systems

typical components:

ASR, NLU: tell system what was said
Dialog Manager: when to say, what to say
Task Manager: perform domain-relevant action
NLG: how to say
TTS: say

dialog systems

additional components:

- speaker identification, verification; e.g. banking
  - system knows the speaker...
    - definitely: say „hi, Cindy”, go directly to appropriate account
    - probably: say “is that Cindy?”
    - possibly: say “have you used this service before?”
  - otherwise: say “hi, what’s your name”
- user model
- modality handlers (input fission, output fusion)
- ...

dialog systems: speech recognition

ASR: speech to words/meanings

language model + recognition grammar („semantic grammar”)

understanding user crucial → grammars typically hand-written context-free rather than statistical

```
REQUEST : tell me | I want | I’d like | …
DEPARTURE_TIME : (after|around|before) HOUR | morning | evening
HOUR : one|two|three| ... |twelve (am|pm)
FLIGHTS : (a) DEPARTURE_TIME flight | DEPARTURE_TIME flights
ORIGIN : from CITY
DESTINATION : to CITY
CITY : London | Warsaw | New York | ...
```

dialog systems: speech recognition

(some) problems

- grammar-writing time-consuming, expensive
- limited coverage (grammar writer will probably miss many possible formulations because he/she just doesn't think about them)

(some) things to consider:

- „restricted" language models: dependent on dialogue state
  - e.g. if asking for city name, model only with city names
- could make use of the fact that the system know who the speaker is
  - adapt to speaker: acoustic, language model, pronounciation
- will user be allowed to speak while the system is speaking?
  - need to correctly detect speech (esp. in noisy environments)
- using recognition confidence values
  - overall utterance, individual words, combined

dialog systems: generation and speech synthesis

NLG: based on content (meaning) to be expressed:

plans sentences

chooses how to express concepts with words; syntactic structures and lexemes → surface realization

simplest method: „canned" utterances (with variable slots)
→ „template-based" generation

if possible, assigns prosody (according to context)

Text-to-Speech component

takes NLG output

synthesizes a waveform

dialog systems: generation and speech synthesis

(some) NLG considerations:

system prompts influence dialog coherence and „naturalness” → variation:

S1: Please say the departure time
S2: Please say the departure city.
S3: Please say the destination city.

S1: First, tell me when you would like to travel
S2: And from which city would you like to leave?
S3: Thanks. Now, what is your destination?

dialog systems: generation and speech synthesis

(some) NLG considerations:

system prompts influence dialog coherence and „naturalness” → tapered prompts (gradually shorter if same sub-dialog):

S: Now, what’s the first company to add to your watch list?
U: Cisco
S: What’s the next company name? Or, you can say, „Done”
U: IBM
S: Tell me the next company name, or say, „Done.”
U: Intel
S: Next one?
U: America Online.
S: Next?
U: …

dialog systems: dialog management

dialog engine's tasks:

when to say? → control the flow of dialog
what to say? → dialog modeling

takes input from ASR/NLU
maintains some sort of „dialog state"
communicates with Task Manager
passes output to NLG/TTS

dialog systems: dialog management

control the flow of dialog

- when to say something and when to listen (turn-taking), when to stop
- update dialog context with current user's input and output the next action in the dialog
- deal with barge-in, hang-ups

dialog modeling

- what is the context
- what to say next

goal: achieve an ***application goal*** in an ***efficient*** way through a series of interaction with the user

dialog systems: dialog models

why need dialog models?

system and user work on a task
dialog structure reflects the task structure

BUT:

dialog need not follow the task-steps
need for grounding

dialog systems: dialog models

examples of dialog models

- FSA
- frame-based
- Information State (aka ISU)

the choice depends on the complexity and nature of the task