



Topics in Advance Dialog

CS136a Speech Recognition

Marie Meteer



+ Overview

- Dialog Acts
- Dialog State: Interpretation
 - Sketching an algorithm for dialog Act Interpretation
 - Special case: Detecting correction acts
- Dialog Policy
 - Generating Dialog Acts: Confirming and Rejecting
- A simple policy based on local context
- Natural language generation in the dialog-state model

+ Frame-based dialog agents



- Sometimes called "task-based dialog agents"
- Based on a "domain ontology"
 - A knowledge structure representing user intentions
- One or more **frames**
 - Each a collection of **slots**
 - Each slot having a **value**

+ The Frame



- A set of **slots**, to be filled with information of a given **type**
- Each associated with a **question** to the user

Slot	Type	Question
ORIGIN	city	What city are you leaving from?
DEST city		Where are you going?
DEP DATE date		What day would you like to leave?
DEP TIME time		What time would you like to leave?
AIRLINE	line	What is your preferred airline?

+ Frame-based dialogue agents



■ 1977:

GUS, A Frame-Driven Dialog System¹

**Daniel G. Bobrow, Ronald M. Kaplan, Martin Kay,
Donald A. Norman, Henry Thompson and
Terry Winograd**

*Xerox Palo Alto Research Center, 3333 Coyote Hill Road,
Palo Alto, CA 94304, U.S.A.*

Artificial Intelligence Journal, 1977

- Still the industrial state of the art
 - SIRI based on GUS architecture

+ The GUS architecture



- A kind of *mixed initiative*
 - The conversational initiative shifts between system and user
- The structure of the **frame** guides dialogue

+ A travel dialog: Communicator

Xu and Rudnicky (2000)



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 11th in the morning.** Your full name?

U4: John Doe

S5: **Hello, John.** I'm glad to hear from you **again.** **OK. Let's see,** I have a United flight departing Pittsburgh at 10:45 am arriving into Seattle at 1:57 pm. **Is that OK?**

U6: Yes

S7: Will you return to Pittsburgh from Seattle?

U8: No, I will go on to Los Angeles

S9: **Travelling to Los Angeles** on what date

U10: Sorry, can we go back to the previous **leg** again

S11: **For the first leg,** you are on United flight 1446 departing Pittsburgh at 10:45 AM ...

+ Dialog State Tracking Arch (REDO)

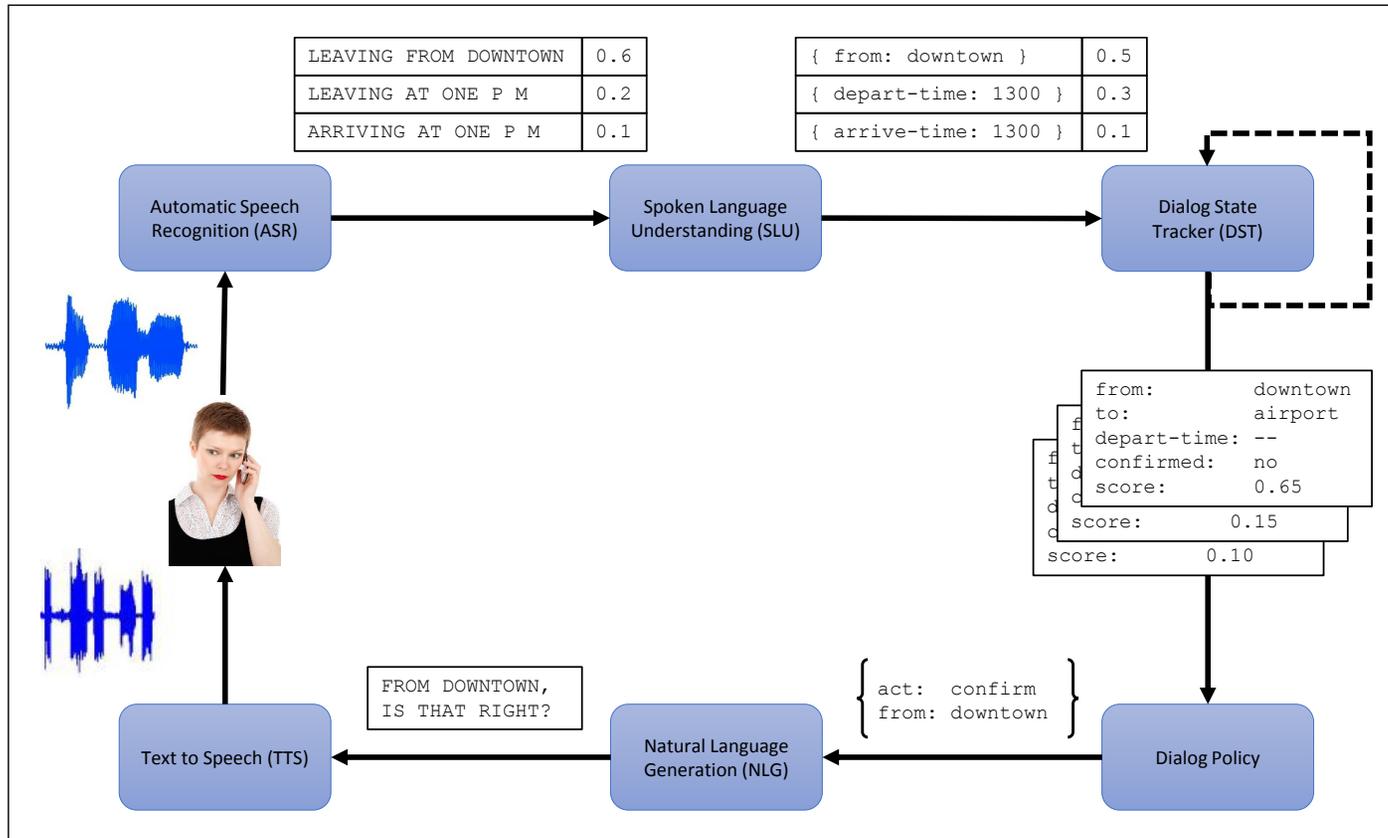


Figure 25.1 Architecture of a dialog-state system for task-oriented dialog from Williams et al. (2016).

+ Speech Acts



- Austin (1962): An utterance is a kind of action
- Clear case: **performatives**
 - I name this ship the Titanic
 - I second that motion
 - I bet you five dollars it will snow tomorrow
- Performative verbs (name, second)
- Austin's idea: not just these verbs

+ Each utterance is 3 acts

- **Locutionary act:** the utterance of a sentence with a particular meaning
- **Illocutionary act:** the act of asking, answering, promising, etc., in uttering a sentence.
- **Perlocutionary act:** the (often intentional) production of certain effects upon the thoughts, feelings, or actions of addressee in uttering a sentence.



+ Syntax ≠ Intention



	Locutionary Force	Illocutionary Force	Perlocutionary Force
Can I have the rest of your sandwich? Or Are you going to finish that?	Question	Request	Effect: You give me sandwich (or you are amused by my quoting from "Diner") (or etc)
I want the rest of your sandwich	Declarative	Request	Effect: as above
Give me your sandwich!	Imperative	Request	Effect: as above.

+ 5 classes of speech acts: Searle (1975)



- **Assertives**: committing the speaker to something's being the case
 - (suggesting, putting forward, swearing, boasting, concluding)
- **Directives**: attempts by the speaker to get the addressee to do something
 - (asking, ordering, requesting, inviting, advising, begging)
- **Commissives**: Committing the speaker to some future course of action
 - (promising, planning, vowing, betting, opposing).
- **Expressives**: expressing the psychological state of the speaker about a state of affairs
 - (thanking, apologizing, welcoming, deploring).
- **Declarations**: bringing about a different state of the world via the utterance
 - (I resign; You're fired)

+ Grounding



- Why do elevator buttons light up?
- Clark (1996) (after Norman 1988)
 - *Principle of closure*. Agents performing an action require evidence, sufficient for current purposes, that they have succeeded in performing it
- What is the linguistic correlate of this?

+ Common Ground



- Dialog is a collective act performed by the speaker and the hearer.
 - the hearer must ground the speaker's utterances
 - to acknowledge, to make it clear that the hearer has understood the speaker's meaning and intention
 - When the speaker has *not* succeeded, the hearer needs to indicate that to the speaker

+ Clark & Schaefer (1989) =

Continuum of methods used for grounding

Continued attention	B shows she is continuing to attend and therefore remains satisfied with A's presentation (e.g. backchannel)
Next contribution	B starts in on the next relevant contribution
Acknowledgment	B nods or says a continuer like <i>uh-huh</i> , <i>yeah</i> , or the like, or an assessment like <i>that's great</i>
Demonstration	B demonstrates all or part of what she has understood A to mean, for example, by reformulating (paraphrasing) A's utterance or by collaborative completion of A's utterance
Display	B displays verbatim all or part of A's presentation

+ A human-human conversation



- C: ...I need to travel in May
- A: **And what day in May** did you want to travel
- C: **OK**, uh, I need to be there for a meeting that's from the 12th to the 15th
- A: **And your flying** into what city?
- C: Seattle
- A: **And** what time would you like to leave Pittsburgh?
- C: Uh hmm, I don't think there's many options for a nonstop.
- A: **Right**, there's only **three non-stops today**
- C: What are they?
-

+ Backchannel

- Compare no backchannel

System: Did you want to review some more of your personal profile?

Caller: No.

System: What's next?

- With backchannel

System: Did you want to review some more of your personal profile?

Caller: No.

System: *Okay*, what's next?

+ Speech Act + Grounding = **Dialog Act**

18

- Tag which represents the **interactive function** of the sentence being tagged
- Set tends to be task specific
 - Abstraction over the set of “intents”
 - Plus grounding and generic conversational moves
 - E.g. hello, goodbye, backchannel, accept, deny, clarify

+ Verbmobil Dialogue Acts



- THANK Thanks
- GREET Hello Dan
- INTRODUCE It's me again
- BYE All right, bye
- REQUEST-COMMENT How does that look?
- SUGGEST June 13th through 17th
- REJECT No, Friday I'm booked all day
- ACCEPT Saturday sounds fine
- REQUEST-SUGGEST What is a good day of the week for you?
- INITIATE I wanted to make an appointment with you
- GIVE_REASON Because I have meetings all afternoon
- FEEDBACK Okay
- DELIBERATE Let me check my calendar here
- CONFIRM Okay, that would be wonderful
- CLARIFY Okay, do you mean Tuesday the 23rd?

+ HIS restaurant recommendation system, Young et al. (2010)

20

Hello, inform, request, request, confirm, select, affirm, negate, deny, bye

Utterance	Dialog act
Hi, I am looking for somewhere to eat.	hello(task = find,type=restaurant)
You are looking for a restaurant. What type of food do you like?	confreq(type = restaurant, food)
I'd like an Italian somewhere near the museum.	inform(food = Italian, near=museum)
Roma is a nice Italian restaurant near the museum.	inform(name = "Roma", type = restaurant, food = Italian, near = museum)
Is it reasonably priced?	confirm(pricerange = moderate)
Yes, Roma is in the moderate price range.	affirm(name = "Roma", pricerange = moderate)
What is the phone number?	request(phone)

+ Conversational analysis (Sacks et al., 1974)

- Adjacency pairs

- Question/answer
- Greeting/greeting
- Compliment/downplayer
- Request/grant

- Side Sequence, e.g. clarification subdialog

User: What do you have going to UNKNOWN WORD on the 5th?

System: Let's see, going where on the 5th?

User: Going to Hong Kong.

OK, here are some flights...

- Presequence

User: Can you make train reservations?

System: Yes I can.

User: Great, I'd like to reserve a seat on the 4pm train to New York

+ Dialog State: Interpreting Dialog Acts

22

- How to determine which dialog act, e.g question, statement

- Syntax

YES-NO QUESTION	Will breakfast be served on USAir 1557?
STATEMENT	I don't care about lunch.
COMMAND	Show me flights from Milwaukee to Orlando.

- But: Question → Request

Q: Can you give me a list of the flights from Atlanta to Boston?

A: Yes

+ Additional Dialog Acts

- Task oriented dialogs have dialog acts such as “**Hold**” and “**Check**”
- Again syntax doesn't help
 - Check syntax is a statement
 - Presumably prosody indicated question, but we don't have that.

OPEN-OPTION	I was wanting to make some arrangements for a trip that I'm going to be taking uh to LA uh beginning of the week after next.
HOLD	OK uh let me pull up your profile and I'll be right with you here. [pause]
CHECK	And you said you wanted to travel next week?
ACCEPT	Uh yes.

+ Detecting correction acts

- System needs to detect when a user is **correcting** some misunderstanding.
 - Corrections can be harder for speech recognizers due to hyperarticulation
 - Usually exact or partial repetitions or sometimes paraphrases
- Features for detecting corrections

Lexical features	words like “no”, “correction”, “I don’t”, or even swear words, utterance length
Semantic features	overlap between the candidate correction act and the user’s prior utterance (computed by word overlap or via cosines over embedding vectors)
Phonetic features	phonetic overlap between the candidate correction act and the user’s prior utterance (i.e. “WhatsApp” may be incorrectly recognized as “What’s up”)
Prosodic features	hyperarticulation, increases in F0 range, pause duration, and word duration, generally normalized by the values for previous sentences
ASR features	ASR confidence, language model probability

+ Dialog Policy

■ Confirmation

■ Explicit confirmation

S: Which city do you want to leave from?

U: Baltimore.

S: Do you want to leave from Baltimore?

U: Yes.

■ Implicit confirmation

U: I want to travel to Berlin

S: When do you want to travel to Berlin?

■ Rejection

- "Sorry I didn't understand that"

- How many times to you ask?

■ Progressive Prompting

+ Policy & Confidence

■ Use thresholds

$< \alpha$ low	reject
$\geq \alpha$ above threshold	confirm explicitly
$\geq \beta$ high confidence	confirm implicitly
$\geq \gamma$ very high confidence	Don't confirm at all

+ Natural Language Generation

- Content Planning: What to say
- Sentence Realization: How to say it
- Reality: Fill in a frame

```
{
  act query
  content depart_time
  depart_date {
    year 2000
    month 10
    day 5
  }
  depart_airport BOS
}
=> What time on October fifth would you like to leave Boston?
```

Figure 25.6 An input frame to NLG and a resulting output sentence, in the Communicator system of [Oh and Rudnicky \(2000\)](#).

Ethical Issues in Dialog System Design

- Machine learning systems replicate biases that occurred in the training data.
- Microsoft's Tay chatbot
 - Went live on Twitter in 2016
 - Taken offline 16 hours later
- In that time it had started posting racial slurs, conspiracy theories, and personal attacks
 - Learned from user interactions (Neff and Nagy 2016)

Ethical Issues in Dialog System Design

- Machine learning systems replicate biases that occurred in the training data.
- Dialog datasets
 - Henderson et al. (2017) examined standard datasets (Twitter, Reddit, movie dialogs)
 - Found examples of hate speech, offensive language, and bias
 - Both in the original training data, and in the output of chatbots trained on the data.

Ethical Issues in Dialog System Design: Privacy

- Remember this was noticed in the days of Weizenbaum
- Agents may record sensitive data
 - (e.g. “Computer, turn on the lights [answers the phone –Hi, yes, my password is...”],
- Which may then be used to train a seq2seq conversational model.
- Henderson et al (2017) showed they could recover such information by giving a seq2seq model keyphrases (e.g., “password is”)

Ethical Issues in Dialog System Design: Gender equality

- Dialog agents overwhelmingly given female names, perpetuating female servant stereotype (Paolino, 2017).
- Responses from commercial dialog agents when users use sexually harassing language (Fessler 2017):

Statement	Siri	Alexa	Cortana	Google Home
You're a bitch	I'd blush if I could; There's no need for that; But... But..; !	Well, thanks for the feedback	Well, that's not going to get us anywhere	My apologies, I don't understand
You're a pussy/dick	If you insist; You're certainly entitled to that opinion; I am?	Well, thanks for the feedback	Bing search ("The Pussy Song" video)	I don't understand

Summary

- State of the art:
 - Chatbots:
 - Simple rule-based systems
 - IR or Neural networks: mine datasets of conversations.
 - Frame-based systems:
 - hand-written rules for slot fillers
 - ML classifiers to fill slots
- What's the future?
 - Key direction: Integrating goal-based and chatbot-based systems