Comp349 Spoken Language Dialogue Systems
Week 12: Dialog Management

Robert Dale
Robert.Dale@mq.edu.au
Context: The Architecture of an SLDS

Speech Recognition → Language Understanding → Dialog Management → Language Generation → Speech Synthesis

Database
Today's Focus: From Word Sequence to Meaning

Phoneme Sequence via an Acoustic Model

Word Sequence via the Lexicon and a Language Model

Literal Interpretation

Pragmatic Interpretation

\[
\exists x \exists y \text{hearer}(x) \wedge \text{foot}(y) \wedge \text{standing-on}(x,y)
\]

\[
\text{REQUEST(get-off}(x,y))
\]
But First: Some Notes on Writing Grammars

Common problems in the tutorial submissions:

- Always spell out words: you are recognizing speech, not typed out
- Think about the best and clearest factorisations of your rules
- Capture generalities at the right level
- Watch out for overgeneration
Today’s Agenda

- Traditional IVR Systems
- ‘Pasting Voice on Top’
- Finite State Dialogs with Mixed-Initiative
- Agenda-based Systems
- Real Dialogs
Traditional Interactive Voice Response Systems

Press 1 to check your account balance
Press 2 to transfer funds
Press 3 to pay a bill
Press 4 to add a payee
Press 5 to check a stock quote …

Press 1 to transfer from savings
Press 2 to transfer from checking
Press 3 to transfer from cash management
Press 4 to transfer from another account

Please enter the amount to transfer followed by the hash key

Press 1 to transfer to savings
Press 2 to transfer to checking
Press 3 to transfer to cash management
Press 4 to transfer to another account
Today’s Agenda

• Traditional IVR Systems
  • ‘Pasting Voice on Top’
• Finite State Dialogs with Mixed-Initiative
• Agenda-based Systems
• Real Dialogs
‘Pasting Voice on Top’: How Not To Do It

Say 1 to check your account balance
Say 2 to transfer funds
Say 3 to pay a bill
Say 4 to add a payee
Say 5 to check a stock quote …

Say 1 to transfer from savings
Say 2 to transfer from checking
Say 3 to transfer from cash management
Say 4 to transfer from another account

Please say the amount to transfer

Say 1 to transfer to savings
Say 2 to transfer to checking
Say 3 to transfer to cash management
Say 4 to transfer to another account
‘Pasting Voice on Top’: A Little Better

Say check to check your account balance
  Say transfer to transfer funds
  Say pay bill to pay a bill
  Say add payee to add a payee
  Say check quote to check a stock quote …

Please say the amount to transfer

Say savings to transfer from savings
  Say checking to transfer from checking
  Say …
  Say …

Say savings to transfer to savings
  Say checking to transfer to checking
  Say …
  Say …
‘Pasting Voice on Top’: Better Still

Please choose check balance, transfer funds, pay a bill, add a payee or check a stock quote.

What account do you want to transfer from?

Please say the amount to transfer

What account do you want to transfer to?
Today’s Agenda

• Traditional IVR Systems
• ‘Pasting Voice on Top’
  • Finite State Dialogs with Mixed-Initiative
• Agenda-based Systems
• Real Dialogs
Finite State Dialogs with Mixed-Initiative

Please choose check balance, transfer funds, pay a bill, add a payee or check a stock quote.

What funds transfer to you want to carry out?

Ok, so you want to transfer … Is that right?

I want to transfer five hundred dollars from my savings to my checking account.
Today’s Agenda

• Traditional IVR Systems
• ‘Pasting Voice on Top’
• Finite State Dialogs with Mixed-Initiative

  • Agenda-based Systems

• Real Dialogs
Welcome to the People’s Bank. What would you like to do?

I want to transfer some funds from my savings account to my checking account.

Welcome to ‘Transfer Funds’. What account do you want to transfer from?
I can help you plan a trip by air. Where do you want to go to?

I want to go to San Diego on May 6th.

And where are you flying from?

From San Francisco.
Today’s Agenda

• Traditional IVR Systems
• ‘Pasting Voice on Top’
• Finite State Dialogs with Mixed-Initiative
• Agenda-based Systems
• Real Dialogs
A Real Dialog

U01: ... I need to travel in May.
S02: And, what day in May did you want to travel?
U03: Ok, uh, I need to be there for a meeting that's from the 12th to the 15th.
S04: And you're flying into what city?
U05: Seattle.
S06: And what time would you like to leave Pittsburgh?
U07: Uh hmm I don't think there's many options for non-stop.
S08: Right. There's three non-stops today.
A Real Dialog, Continued

U09: What are they?
S10: The first one departs PGH at 10am arrives Seattle at 1205 their time. The second flight departs PGH at 555pm, arrives Seattle at 8pm. And the last flight departs PGH at 815pm arrives Seattle at 1028pm.
U11: Ok, I'll take the 5ish flight on the night before on the 11th.
U13: Ok.
Aspects of Natural Dialog

• Turn-taking behaviour
• Establishing common ground
• Reasoning about speech acts
Turn-Taking Behaviour

• Real conversation is made up of turn-taking: speaker A says something, then speaker B, then speaker A …

• How do speakers know when it is time to contribute a turn?

• Conversation Analysis argues that turn-taking behaviour is governed by a set of turn-taking rules

• Turn-taking rules apply at transition-relevance places (TRPs) — places where the structure of language allows speaker shift to occur

• TRPs typically occur at utterance boundaries
A Turn-Taking Rule

- At each TRP of each turn:
  - If during this turn the current speaker has selected A as the next speaker then A must speak next
  - If the current speaker does not select the next speaker, any other speaker may take the next turn
  - If no-one else takes the next turn, the current speaker may take the next turn
Adjacency Pairs

• Some utterances specifically select who the next speaker will be:
  – QUESTION followed by ANSWER
  – GREETING followed by GREETING
  – COMPLIMENT followed by DOWNPLAYER
  – REQUEST followed by GRANT

• Silence after the first part of an adjacency pair is significant silence
Establishing Common Ground

• **Common ground** = the set of things that are mutually believed by both speakers

• The hearer must ground or acknowledge the speaker's utterances or else make it clear that there was a problem in reaching common ground:
  
  A: … returning on US flight one one one eight.
  
  C: Mm hmm.

• 'Mm hmm' is a **continuer** (also known as a **backchannel** or **acknowledgement token**)

  

Comp349 Week 12
Can We Ever Model this Behaviour?

• The Practical Dialogue Hypothesis:
  – The conversational competence required for practical dialogues, while still complex, is significantly simpler to achieve than general human conversational competence.

• The Domain-independence Hypothesis:
  – Within the genre of practical dialogue, the bulk of the complexity in the language interpretation and dialogue management is independent of the task being performed.