COMP349
Spoken Language Dialogue Systems

Robert Dale
Robert.Dale@mq.edu.au
The Class So Far

- Peter Samaras
- JunXin Hu (Noel)
- MingYin Ng (Boris)
What This Unit is About: First Half

- The techniques required to build spoken language dialog systems
- The challenges that language poses for machine processing
- The limitations of spoken language dialogue systems
- Issues in dialogue design, prompt design and grammar writing
- The importance of error recovery techniques
- Practical implementation of spoken language dialogue systems using VoiceXML
What This Unit is About: Second Half

- The development lifecycle for spoken language systems
- The range of approaches available for dialog design
- The issues that arise in building embodied conversational agents ('talking heads')
- Practical development using existing talking head technology
- Future directions for talking head research
The Aims of Today's Class

• To give an idea of state of the art in spoken language dialog systems
• To provide a high-level overview of the different approaches to developing spoken language dialog systems
• To review the challenges in building conversational systems
Outline

• Logistics
• Spoken Language Dialog Systems Today
• Why Conversational Systems are Challenging
• Revisiting Current Applications
• Next Week's Tutorial and Practical
Logistics

- Class meets every Wednesday 6-9pm in C5A 310
- Tutorials and Practicals begin in Week 2
- Class taught by
  - Robert Dale (Week 1; Weeks 8-12)
  - Rolf Schwitter (Weeks 2-7)
- All materials via WebCT:
  http://online.mq.edu.au/pub/COMP349/index.html
Outline

• Logistics
  • Spoken Language Dialog Systems Today
• Why Conversational Systems are Challenging
• Revisiting Current Applications
• Next Week's Tutorial and Practical
Spoken Language Dialog Systems Today

- What is an SLDS?
- Why SLDSs are Important
- The Architecture of an SLDS
- Key Concepts
- The Bigger Picture
Examples of Spoken Language Dialog Systems

- Early interactions with machines
- Modern portrayals: The Jackal
- Real Systems Today:
  - Thrifty Car Rental
  - VeCommerce's VeBet
  - Nuance’s Airline Reservations
- In the labs: CMU's airline booking demo
What is a Spoken Language Dialog System?

• An SLDS is a computer system that you can talk to in order to carry out some task
• SLDSs are typically of two kinds:
  – Transaction-based systems allow you to undertake some transaction, such as buying or selling stocks, or reserving a seat on a plane
  – Information-provision systems provide information in response to a query, such as a request for timetable information or weather information
Today's Agenda

• What is an SLDS?
• Why SLDSs are Important
• The Architecture of an SLDS
• Key Concepts
• The Bigger Picture
The Evolving Interface

• 1960s through 1980s: command line interfaces
• 1980s and 1990s: graphical user interfaces
• The new millennium: the voice user interface
Two Uses of Speech Recognition Technology

• **On the desktop:**
  – Speaker-dependent
  – Large vocabulary (tens of thousands of words)
  – Dictation tasks

• **Telephony-based:**
  – Speaker-independent
  – Relatively small vocab (hundreds of words)
  – Interactive tasks
Uses of Telephony-Based SLDSs

- Remote banking
- Travel reservations
- Information enquiry
- Stock transactions
- Auto-attendants
- Directory assistance
- Taxi bookings
- Pizza ordering
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 ...

Please enter the amount to transfer followed by the hash key

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

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
Speech-Enabled Interaction

Transfer 500 dollars from savings to checking next Wednesday after 3pm

~25 seconds via speech — as compared with two minutes via touch tone
Today's Agenda

- What is an SLDS?
- Why SLDSs are Important
- The Architecture of an SLDS
- Key Concepts
- The Bigger Picture
The Architecture of an SLDS

Speech Recognition -> Language Understanding

Database

Dialog Management

Language Generation -> Speech Synthesis
What an SLDS contains

- **Speech Recognition**—analyses the audio speech input signal to extract linguistic units such as words or phonemes
- **Language Understanding**—determines the meaning of the input
- **Dialog Management**—manages the flow of the conversation, maintaining history and context, directing its course, accessing the database, and formulating responses
- **Database**—stores the information which provides the dialog content
- **Language Generation**—puts the responses into words
- **Speech Synthesis**—produces the audio speech output signal
Today's Agenda

- What is an SLDS?
- Why SLDSs are Important
- The Architecture of an SLDS
- Key Concepts
- The Bigger Picture
What's Involved in Building an SLDS Today

• Dialog Design
  – The process of working how the interaction between human and machine will move from stage to stage
  – Also referred to as script writing or call-flow layout

• Grammar Writing
  – Specifying what the user is permitted to say at any given state

• Prompt Design
  – What the system says to get the user to say something that is permitted to be said!
Key Things to Learn in Building SLDSs

• How to craft effective prompts
• How to predict and manage social user-behaviours
• How to exploit the capabilities of speech recognition technologies
• How to use time to the advantage of both user and application
• How to keep a consistent 'look and feel' to the interface
• How to balance tradeoffs when selecting recognition vocabulary
• How to approach the problem of application 'personality'
• How to prevent errors from looking like errors
Today's Agenda

- What is an SLDS?
- Why SLDSs are Important
- The Architecture of an SLDS
- Key Concepts
- The Bigger Picture
Where Do SLDSs Fit?

- Speech and Language Technology Applications
- Human Computer Interaction Applications
- Spoken Language Dialog Systems
Other Language Technology Applications

- Search and Information Retrieval
- Question Answering
- Information Extraction
- Text Summarisation
- Grammar Checking
- Natural Language Report Generation
- Text Personalisation
Outline

• Logistics
• Spoken Language Dialog Systems Today
  • Why Conversational Systems are Challenging
• Revisiting Current Applications
• Next Week's Tutorial and Practical
Word Boundaries

- recognise speech
- wreck a nice peach
Stages in Processing Language

- Speech Stream
- Word Sequence

Word Boundary Detection
Lexical Ambiguity

- The astronomer saw the star.
- The astronomer married the star.
- King Kong sat on the bank.
Stages in Processing Language

Speech Stream

Word Sequence

Tagged Words

Word Boundary Detection

Lexical Lookup
Structural Ambiguity

• The astronomer saw the star with a telescope.
• Visiting relatives can be nuisance.
Stages in Processing Language

Speech Stream
Word Sequence
Tagged Words
Syntactic Structure

Word Boundary Detection
Lexical Lookup
Syntactic Analysis
Semantic Interpretation

- The green box is on the table
- $\exists s \exists x \exists y \text{state}(s, \text{on-top-of}) \land \text{holds}(s, \text{NOW}) \land \text{type}(x, \text{box}) \land \text{colour}(x, \text{green}) \land \text{type}(y, \text{table})$
Semantic Interpretation

• What I’d really like to do is take the first flight that will get me in before nine in the morning but only if I can fly on a Qantas flight that doesn’t prevent me from using my Frequent Flyer points for an upgrade.
Anaphora Resolution

- Andy put the cake on the table and ate it.
- Sue went to Mary’s house and she cooked her dinner.
- Maisy swore at Sabine then she insulted her.
Stages in Processing Language
Non-literal Meaning

- Can you pass the salt?
- You’re standing on my foot.
- It’s cold in here.
Stages in Processing Language

- Speech Stream
- Word Sequence
- Tagged Words
- Syntactic Structure
- Literal Interpretation
- Speech Act Interpretation

- Word Boundary Detection
- Lexical Lookup
- Syntactic Analysis
- Semantic Analysis
- Pragmatic Analysis
Plan Recognition

Plan inference and co-operative response:

User: Which students got an F in Comp248 in 1999?
System: None.

User: Did anyone fail Comp248 in 1999?
System: No.

User: How many people passed Comp248 in 1999?
System: Zero.

User: Was Comp248 given in 1999?
System: No.
Outline

• Logistics
• Spoken Language Dialog Systems Today
• Why Conversational Systems are Challenging
  • Revisiting Current Applications
• Next Week's Tutorial and Practical
Examples of Spoken Language Dialog Systems

- Early interactions with machines
- Modern portrayals: The Jackal
- Real Systems Today:
  - Thrifty Car Rental
  - VeCommerce's VeBet
  - Nuance’s Airline Reservations
- In the labs: CMU's airline booking demo
Outline

• Logistics
• Spoken Language Dialog Systems Today
• Why Conversational Systems are Challenging
• Revisiting Current Applications
• Next Week's Tutorial and Practical
Week 2 Tutorial

- You need to find three speech system vendors and investigate their sites
- You need to provide summary results in a PowerPoint file via WebCT by Monday, August 6, 2007, 9am
- You will present these results and discuss your findings in the tutorial
- See the WebCT unit website for more information
Week 2 Practical

• You will investigate SLDS demo systems online
• You will answer questions about these systems and rate them on their usability
• See the WebCT unit website for more information