

# CSE 842

## Natural Language Processing

### Lecture 23: Dialogue Systems

4/22/2009

CSE842, Spring 2009, MSU

1

## Talk Announcement

- Title: Using Millions of Web Blogs as a Knowledge Base
- Speaker: Dr. Andrew Gordon
- Time: 11:00am Friday, April 24
- Location: EB3540

4/22/2009

CSE842, Spring 2009, MSU

2

## Spoken Dialogue Systems

- Today's spoken dialogue systems make it possible to accomplish **real tasks**, over the phone, **without talking to a person**
  - Real-time speech technology enables real-time interaction
  - Speech recognition and understanding is '*good enough*' for limited, goal-directed interactions
- Many applications
  - Information seeking (restaurant, movies)
  - Reservation (airlines, trains, etc)
  - Call center applications
  - Automated tutoring systems

4/22/2009

CSE842, Spring 2009, MSU

3

## Properties of Human Conversation

- Turn-taking
- Joint activities
  - Speech acts/dialogue acts
  - Grounding
- Conversation implicatures

4/22/2009

CSE842, Spring 2009, MSU

4

## Turns and Utterances

C<sub>1</sub>: ... I need to travel in May.  
A<sub>1</sub>: And, what day in May did you want to travel?  
C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.  
A<sub>2</sub>: And you're flying into what city?  
C<sub>3</sub>: Seattle.  
A<sub>3</sub>: And what time would you like to leave Pittsburgh?  
C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.  
A<sub>4</sub>: Right. There's three non-stops today.  
C<sub>5</sub>: What are they?  
A<sub>5</sub>: The first one departs PGH at 10:00am arrives Seattle at 12:05 their time. The second flight departs PGH at 5:55pm, arrives Seattle at 8pm. And the last flight departs PGH at 8:15pm arrives Seattle at 10:28pm.  
C<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.  
A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.  
C<sub>7</sub>: OK.

4/22/2009

CSE842, Spring 2009, MSU

5

## Turns and Utterances

- Dialogue is characterized by **turn-taking**: *who* should talk next, and *when* they should talk
- How do we identify turns in recorded speech?
  - Little speaker overlap (around 5% in English -- although depends on domain)
  - But little silence between turns either
- How are turns allocated? How do we know when to take a turn?

4/22/2009

CSE842, Spring 2009, MSU

6

## Simplified Turn-Taking Rule

- Conversation Analysis provides a socio-linguistic approach to turn-taking (e.g., Sacks et al.)
- At each transition-relevance place of each turn:
  - If during this turn current speaker has selected A as the next speaker, then A must speak next
  - If 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
- Transition-relevance places are where the structure of the language allows speaker shifts to occur.

4/22/2009

CSE842, Spring 2009, MSU

7

## Conversation Analysis

- **Adjacency pairs** (set up next speaker expectations)
  - GREETING GREETING
  - QUESTION ANSWER
  - COMPLIMENT DOWNPLAYER
  - REQUEST GRANT
- **Significant silence** follows first element of adjacency pair
  - A: *Is there something bothering you or not? (1.0s)*
  - A: *Yes or no? (1.5s)*
  - A: *Eh?*
  - B: *No.*

4/22/2009

CSE842, Spring 2009, MSU

8

# Utterances

- **Transition-relevance places** typically occur at utterance boundaries but how to define “utterance”
  - Spoken utterances typically shorter, contain more pronouns, include repairs ..., compared to written
  - Cue: well, so.
  - A single utterance may span several turns  
*A: We've got you on USAir flight 99*  
*B: Yep*  
*A: leaving on December 1.*
  - Multiple utterances may occur in single turn  
*A: We've got you on USAir flight 99 leaving on December. Do you need a rental car?*

# Speech Acts

- Austin (1962) observed that dialogue utterances are a kind of speaker **action**, or speech act: an utterance in a dialogue is a kind of action being performed by the speaker.
  - E.g.: performative sentences  
*I name this ship the Titanic.*  
*I second this motion.*  
*I bet you five dollars it will be sunny tomorrow.*

# Types of Speech Acts

- **Locutionary acts:** the utterance of a sentence with a particular meaning
- **Illocutionary acts:** the act of asking, answering, promising, etc. in uttering a sentence
- **Perlocutionary acts:** the (often intentional) production of certain effects upon the feelings, thoughts, or actions of the addressee in uttering a sentence
- **You can't do that.**
  - locutionary: utterance
  - illocutionary force: protesting
  - perlocutionary effect: stopping or annoying the hearer
- The term **speech act** is generally used to describe illocutionary acts.

# Types of Illocutionary Acts

- Searle's term to classify **illocutionary acts** (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 (e.g., 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 (including many of the performative acts above: I resign, you're fired)

# Grounding

- Conversational participants must continually establish **common ground** (or, **mutual belief**)
- **Principle of Closure** (Clark 96): Agents performing an action require evidence, sufficient for current purpose, that they have succeeded in performing it.
- Hearer must **ground** a speaker's utterances (by making it clear that (believed) understanding has occurred), or else indicate that a grounding problem occurred

4/22/2009

CSE842, Spring 2009, MSU

13

# The Contribution Model (Clark&Schaefer'89)

- Each joint linguistic contribution has two phases:
  - **Presentation:** a speaker presents the hearer with an utterance, performing a speech act
  - **Acceptance:** the hearer grounds the utterance, indicating whether the understanding was achieved

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

4/22/2009

CSE842, Spring 2009, MSU

14

# Examples

- **Acknowledgement**
  - continuer / backchannel / acknowledgement token (also nods if vision available)
  - A: ... returning on U.S. flight one.
  - C: *Mm hmm*
  - grounds A's utterance, and also returns turn
- **Display** (stronger method)
  - display all or part of utterance to be grounded verbatim
  - C: *OK I'll take the 5ish flight on the 11th.*
  - A: *On the 11<sup>th</sup>?*
  - Request for repair indicates lack of grounding
  - C: *OK I'll take the 5ish flight on the 11th.*
  - A: *Huh?*
  - C: *I'll take the 5ish flight on the 11th.*

4/22/2009

CSE842, Spring 2009, MSU

15

# Conversational Implicature

- **Example:**
  - A: *What day in May did you want to travel?*
  - C: *I need to be there for a meeting that is from the 12<sup>th</sup> to 15<sup>th</sup>.*

4/22/2009

CSE842, Spring 2009, MSU

16

## Conversational Implicature

- Example:

*A: What day in May did you want to travel?*

*C: I need to be there for a meeting that is from the 12<sup>th</sup> to 15<sup>th</sup>.*

*A: OK. There are 3 non-stops on the 11<sup>th</sup>.*

## Conversational Implicature

- Example:

*A: What day in May did you want to travel?*

*C: I need to be there for a meeting that is from the 12<sup>th</sup> to 15<sup>th</sup>.*

*A: OK. There are 3 non-stops on the 11<sup>th</sup>.*

- Implicature-licensed inferences

- The meeting information answers the request for travel dates
- Not seven non-stops

## Conversational Implicature

- Pragmatics: the study of how language is used to accomplish goals; beyond literal meaning. The interpretation of utterance relies on more than just the literal meaning of the sentences
- Conversational Implicature means a particular class of inferences (that the speaker expects the hearer to draw)
- Grice's maxims for conversation explain what enables hearers draw such inferences.

## Grice's Maxims

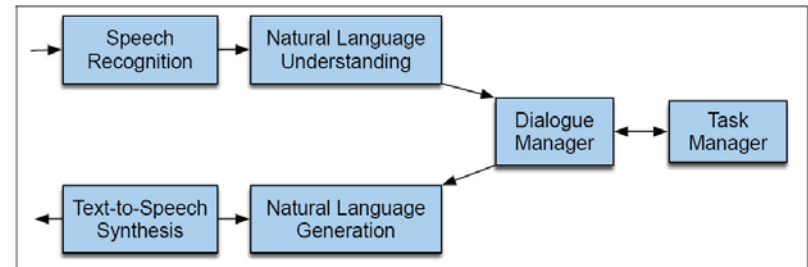
- Maxim of Quantity: Be exactly as informative as is required:
  - Make your contribution as informative as is required (for the current purposes of the exchange).
  - Do not make your contribution more informative than is required.

# Grice's Maxims

- Maximum of Quality: Try to make your contribution one that is true:
  - Do not say what you believe to be false.
  - Do not say that for which you lack adequate evidence
- Maxim of Relevance: Be relevant
- Maxim of Manner: Be perspicuous:
  - Avoid obscurity of expression
  - Avoid ambiguity
  - Be brief (avoid unnecessary prolixity)
  - Be orderly

# Basic Dialogue Systems

Simplified architecture



## Dialog Manager – Motivating Examples

- Dialog 1
  - S: Would you like movie showtime or theater playlist information
  - U: Movie showtime.
  - S: What movie do you want showtime information about?
  - U: Saving Private Ryan.
  - S: At what theater do you want to see Saving Private Ryan?
  - U: Paramount theater
  - S: Saving Private Ryan is not playing at the Paramount theater

## Dialog Manager – Motivating Examples (Cont'd)

- Dialog 2
  - S: How may I help you?
  - U: When is Saving Private Ryan playing?
  - S: For what theater?
  - U: The Paramount theater
  - S: Saving Private Ryan is not playing at the Paramount theater, but it's playing at the Madison theater at 3:00.

# Dialog Management

- Goal: determine what to accomplish in response to user utterances, e.g.:
  - Answer user question
  - Solicit further information
  - Confirm/Clarify user utterances
  - Notify invalid query
  - Suggest alternative
- Interface between user/language processing components and system knowledge base

# Dialog Management (Cont'd)

- Main design issues:
  - Functionality: how much should the system do?
  - Process: how should the system do them? Dialog strategies?
- Affected by:
  - Task complexity: how hard the task is
  - Affects:
    - Robustness
    - Naturalness
    - Perceived intelligence

# Dialogue Management: Process

- Determines how the system will go about selecting among the possible goals
  - Frame-based, plan-based, learning-based, etc.
- At the dialogue level, determined by system designer in terms of initiative strategies:
  - System-initiative: system always has control, user only responds to system questions
  - User-initiative: user always has control, system passively answers user questions
  - Mixed-initiative: control switches between system and user

# Types of Confirmation Strategies

U: I want to go to Baltimore.

- Explicit  
S: Did you say you want to go to Baltimore?
- Implicit  
S: What time do you want to leave Baltimore?

# Evaluating Dialogue Systems

- “**Performance**” of a dialogue system is affected both by *what* gets accomplished by the user and the dialogue agent and *how* it gets accomplished
- **Metrics:**
  - **Efficiency** of the Interaction: User Turns, System Turns, Elapsed Time
  - **Quality** of the Interaction: ASR rejections, Time Out Prompts, Help Requests, Barge-Ins, Mean Recognition Score (concept accuracy), Cancellation Requests
  - **User Satisfaction**
  - **Task Success:** perceived completion, information extracted

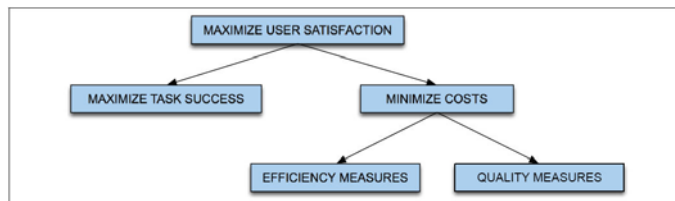
# User Satisfaction

## Survey questionnaires

<b>TTS Performance</b>	Was the system easy to understand ?
<b>ASR Performance</b>	Did the system understand what you said?
<b>Task Ease</b>	Was it easy to find the message/flight/train you wanted?
<b>Interaction Pace</b>	Was the pace of interaction with the system appropriate?
<b>User Expertise</b>	Did you know what you could say at each point?
<b>System Response</b>	How often was the system sluggish and slow to reply to you?
<b>Expected Behavior</b>	Did the system work the way you expected it to?
<b>Future Use</b>	Do you think you'd use the system in the future?

# Evaluating Dialogue Systems

- PARADISE framework (walker et al, 97)
- Multiple regression to learn the weights of factors related to evaluation metrics.



# New Course Announcement

- CSE891: Language and Interaction
  - An introduction to dialogue systems
- Will be offered in Spring 2010
  - 3 credits
  - Time: Mon, Wed. 12:40-2:00p
  - Location: EB2320