

# CS 224S / LINGUIST 236 Speech Recognition and Synthesis

Dan Jurafsky

## Lecture 14: More sophisticated Conversational Agents

IP Notice:

2/17/05

CS 224S Winter 2005

1

## Outline

- **Grounding**,
  - Explicit vs. implicit confirmation
  - Rejection
  - Progressive reprompting
  - Recognizing and handling user corrections
- **Dialogue Acts**
  - Cues for them
  - HMM model for dialogue act interpretation
- **Planning and Intention**
  - Gricean implicature
  - BDI model of dialogue acts
  - Planning for understanding and planning for production

2/17/05

CS 224S Winter 2005

2

## Review

- **Finite-state dialogue management**
- **Frame-based dialogue management**
  - Semantic grammars
  - ASR
  - System, User, and Mixed-initiative
  - Voice XML

2/17/05

CS 224S Winter 2005

3

## We want more complex dialogue

- **Finite-state and form-based agents**
  - Practical but limited
  - No error handling
  - No handling of unexpected questions
  - In fact, can't even tell when they are asked a question!
- **Today's Goal:**
  - Introduce more sophisticated ideas about dialogue
  - Draw from human-human conversation
  - Look at current research paradigms in dialogue systems

2/17/05

CS 224S Winter 2005

4

## On the way to more powerful dialogue systems

- **Grounding**
  - Performing grounding
  - Recognizing user's grounding
- **Dialogue Acts**
  - Using correct dialogue acts
  - Recognizing user's dialogue acts
- **Planning and Intention**
  - Planning what to say
  - Recognizing user's intentions and plans

2/17/05

CS 224S Winter 2005

5

## Grounding and Confirmation

- Dialogue is a collective act performed by speaker and hearer
- **Common ground:** set of things mutually believed by both speaker and hearer
- Need to achieve common ground, so hearer must **ground** or **acknowledge** speakers utterance.
- Clark (1996):
  - *Principle of closure.* Agents performing an action require evidence, sufficient for current purposes, that they have succeeded in performing it
- (Interestingly, Clark points out that this idea draws from Norman (1988) work on non-linguistic acts)
- Need to know whether an action succeeded or failed

2/17/05

CS 224S Winter 2005

6

## Clark and Schaefer: Grounding

- **Continued attention:** B continues attending to A
- **Relevant next contribution:** B starts in on next relevant contribution
- **Acknowledgement:** B nods or says continuer like *uh-huh, yeah, assessment (great!)*
- **Demonstration:** B demonstrates understanding A by paraphrasing or reformulating A's contribution, or by collaboratively completing A's utterance
- **Display:** B displays verbatim all or part of A's presentation

2/17/05

CS 224S Winter 2005

7

## A human-human conversation

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.

2/17/05

CS 224S Winter 2005

8

## Grounding examples

- **Display:**
  - C: I need to travel in May
  - A: And, what day in May did you want to travel?
- **Acknowledgement**
  - C: He wants to fly from Boston
  - A: mm-hmm
  - C: to Baltimore Washington International
  - [Mm-hmm (usually transcribed "uh-huh") is a backchannel, continuer, or acknowledgement token]

2/17/05

CS 224S Winter 2005

9

## Grounding Examples (2)

- **Acknowledgement + next relevant contribution**
  - And, what day in May did you want to travel?
  - And you're flying into what city?
  - And what time would you like to leave?
- The **and** indicates to the client that agent has successfully understood answer to the last question.

2/17/05

CS 224S Winter 2005

10

## Grounding and Dialogue Systems

- Grounding is not just a tidbit about humans
- Is key to design of conversational agent
- Why?

2/17/05

CS 224S Winter 2005

11

## Grounding and Dialogue Systems

- Grounding is not just a tidbit about humans
- Is key to design of conversational agent
- Why?
  - HCI researchers find users of speech-based interfaces are confused when system doesn't give them an explicit acknowledgement signal
  - Stifelman et al. (1993), Yankelovich et al. (1995)

2/17/05

CS 224S Winter 2005

12

## Grounding negative responses

From Cohen et al. (2004)

- System: Did you want to review some more of your personal profile?
- Caller: No. **Good!**
- System: *Okay*, what's next?
  
- System: Did you want to review some more of your personal profile?
- Caller: No. **Bad!**
- System: What's next?

2/17/05

CS 224S Winter 2005

13

## Confirmation

- Another reason for grounding
- Errors: Speech is a pretty errorful channel
  - Even for humans; so they use grounding to *confirm* that they heard correctly
- ASR is way worse than humans!
- So dialogue systems need to do even more grounding and confirmation than humans

2/17/05

CS 224S Winter 2005

14

## Explicit confirmation

- S: Which city do you want to leave from?
- U: Baltimore
- S: Do you want to leave from Baltimore?
- U: Yes

2/17/05

CS 224S Winter 2005

15

## Explicit confirmation

- U: *I'd like to fly from Denver Colorado to New York City on September 21st in the morning on United Airlines*
- S: Let's see then. I have you going from Denver Colorado to New York on September 21st. Is that correct?
- U: *Yes*

2/17/05

CS 224S Winter 2005

16

## Implicit confirmation: display

- U: *I'd like to travel to Berlin*
- S: When do you want to travel to Berlin?
  
- U: *Hi I'd like to fly to Seattle Tuesday morning*
- S: Traveling to Seattle on Tuesday, August eleventh in the morning. Your name?

2/17/05

CS 224S Winter 2005

17

## Implicit vs. Explicit

- Complementary strengths
- Explicit: easier for users to correct systems's mistakes (can just say "no")
- But explicit is cumbersome and long
- Implicit: much more natural, quicker, simpler (if system guesses right).

2/17/05

CS 224S Winter 2005

18

## Implicit and Explicit

- Early systems: all-implicit or all-explicit
- Modern systems: adaptive
- How to decide?
  - ASR system can give **confidence metric**.
  - This expresses how convinced system is of its transcription of the speech
  - If high confidence, use implicit confirmation
  - If low confidence, use explicit confirmation

2/17/05

CS 224S Winter 2005

19

## Computing confidence

- Simplest: use acoustic log-likelihood of user's utterance
- More features
  - Prosodic: utterances with longer pauses, FO excursions, longer durations
  - Backoff: did we have to backoff in the LM?
  - Cost of an error: Explicit confirmation before moving money or booking flights

2/17/05

CS 224S Winter 2005

20

## Rejection

- e.g., VoiceXML "nomatch"
- "I'm sorry, I didn't understand that."
- Reject when:
  - ASR confidence is low
  - Best interpretation is semantically ill-formed
- Might have four-tiered level of confidence:
  - Below confidence threshold, reject
  - Above threshold, explicit confirmation
  - If even higher, implicit confirmation
  - Even higher, no confirmation

2/17/05

CS 224S Winter 2005

21

## Progressive reprompting or "escalating detail"

- System: When would you like to leave?
- Caller: Well, um, I need to be in New York in time for the first World Series game.
- System: <reject>. Sorry, I didn't get that. Please say the month and day you'd like to leave.
- Caller: I wanna go on October fifteenth

2/17/05

CS 224S Winter 2005

22

## Rapid reprompting

- Reject an utterance just by saying
  - "I'm sorry"
  - "What was that?"
- Idea:
  - Use rapid reprompting the first time
  - Then use progressive prompting afterwards
  - Cohen et al. (2004) summarizes experiments showing that users greatly prefer this

2/17/05

CS 224S Winter 2005

23

## Clarification subdialogues

- If the system knew exactly which word the recognition had failed on/had low confidence
- We could ask the user questions about just that word instead of making them repeat the whole sentence

2/17/05

CS 224S Winter 2005

24

## Error Handling

- Despite all these clever confirmation/rejection strategies, dialogue systems still make mistakes (Surprise!)
- If system misrecognizes an utterance, and either
  - Rejects
  - Via confirmation, displays its misunderstanding
- Then user has a chance to make a **correction**
  - Repeat themselves
  - Rephrasing
  - Saying "no" to the confirmation question.

2/17/05

CS 224S Winter 2005

25

## Corrections

- Unfortunately, corrections are **harder** to recognize than normal sentences!
  - Swerts et al (2000): corrections misrecognized twice as often (in terms of WER) as non-corrections!!!
  - Why?
    - Prosody seems to be largest factor: [hyperarticulation](#)
    - English Example from Liz Shriberg
      - "NO, I am DE-PAR-TING from Jacksonville"
    - A German example from Bettina Braun from a talking elevator

2/17/05

CS 224S Winter 2005

26

## Swerts: 5 Strategies users use

- Repetition (REP)
- Paraphrase (PAR)
- Add task relevant content (ADD)
- Omit content (OMIT)
- About two-thirds of all are repetitions and omissions.

2/17/05

CS 224S Winter 2005

27

## A Labeled dialogue (Swerts et al)

Turn	Corr Type
S On which day of the week do you want to leave?	
S Sorry, I didn't hear anything from you. Please tell me your departure date	
U <i>Um tomorrow</i>	NA
S You can say the departure date. For example, you can say 'today' or 'tomorrow'. Or you can say 'Monday', 'Tuesday', 'Wednesday'...	
U <i>Tomorrow</i>	REP
S You can say the departure date. For example, you can say 'today' or 'tomorrow'. Or you can say 'Monday', 'Tuesday', 'Wednesday'...	
U <i>Saturday</i>	PAR
S Do you want to leave on Saturday?	
U Yes	
S What time do you want to leave?	
U Eight o'clock A M	
S Do you want to leave around ten o'clock p-m	
U <i>Eight o'clock</i>	OMIT
S Do you want to leave around ten o'clock p-m	
U No	OMIT
S What time do you want to leave?	
U <i>Eight o'clock A M</i>	ADD

## Machine learning to detect user corrections

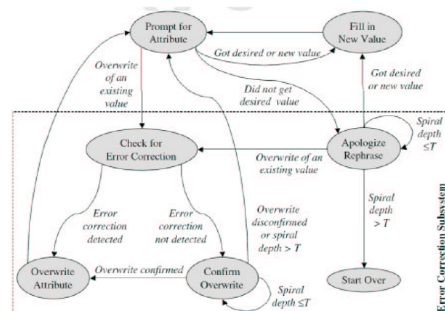
- Build classifiers using features like
  - Lexical information (words "no", "correction", "I don't", swear words)
  - Prosodic features (various increases in FO range, pause duration, and word duration that correlation with hyperarticulation)
  - Length
  - ASR confidence
  - LM probability
  - Various dialogue features (repetition)

2/17/05

CS 224S Winter 2005

29

## Fitting error corrector into a dialogue manager: (Bulyko et al 2004)



## On the way to more powerful dialogue systems

- **Grounding**
  - Performing grounding
  - Recognizing user's grounding
- **Dialogue Acts**
  - Using correct dialogue acts
  - Recognizing user's dialogue acts
- **Planning and Intention**
  - Planning what to say
  - Recognizing user's intentions and plans

2/17/05

CS 224S Winter 2005

31

## 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**

2/17/05

CS 224S Winter 2005

32

## 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.

2/17/05

CS 224S Winter 2005

33

## Locutionary and illocutionary

- "You can't do that!"
- **Illocutionary force:**
  - Protesting
- **Perlocutionary force:**
  - Intent to annoy addressee
  - Intent to stop addressee from doing something

2/17/05

CS 224S Winter 2005

34

## The 3 levels of act revisited

	Locutionary Force	Illocutionary Force	Perlocutionary Force
Can I have the rest of your sandwich?	Question	Request	Intent: You give me sandwich
I want the rest of your sandwich	Declarative	Request	Intent: You give me sandwich
Give me your sandwich!	Imperative	Request	Intent: You give me sandwich

2/17/05

CS 224S Winter 2005

35

## Illocutionary Acts

- **What are they?**

2/17/05

CS 224S Winter 2005

36

## 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*)

2/17/05

CS 224S Winter 2005

37

## Dialogue acts

- Also called "conversational moves"
- An act with (internal) structure related specifically to its dialogue function
- Incorporates ideas of grounding
- Incorporates other dialogue and conversational functions that Austin and Searle didn't seem interested in

2/17/05

CS 224S Winter 2005

38

## Verbmobil task

- Two-party scheduling dialogues
- Speakers were asked to plan a meeting at some future date
- Data used to design conversational agents which would help with this task
- (cross-language, translating, scheduling assistant)

2/17/05

CS 224S Winter 2005

39

## Verbmobil Dialogue Acts

THANK	Thanks
GREET	Hello Dan
INTRODUCE	It's me again
BYE	Allright, 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?
INIT	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?

2/17/05

CS 224S Winter 2005

40

A01: Tag // Herr Scheytt.  
(GREET, INTRODUCE\_NAME)  
(Hello, Mr Scheytt)

B02: Guten Tag // Frau Klein // Wir müssen noch einen Termin ausmachen // für die Mitarbeiterbesprechung.  
(GREET, INTRODUCE\_NAME, INIT\_DATE, SUGGEST\_SUPPORT\_DATE)  
(Hello, Mrs. Klein, we should arrange an appointment, for the team meeting)

A03: Ja, // ich würde Ihnen vorschlagen in Januar, // zwischen dem fünfzehnten und neunzehnten.  
(SUGGEST\_SUPPORT\_DATE, SUGGEST\_SUPPORT\_DATE, SUGGEST\_SUPPORT\_DATE)  
(Well, I would suggest in January, between the fifteenth and the nineteenth)

B04: Oh // das ist ganz schlecht. // zwischen dem elften und achtzehnten Januar bin ich in Hamburg.  
(COUNTER, REJECT\_DATE, SUGGEST\_SUPPORT\_DATE)  
(Oh, that is really inconvenient, I'm in Hamburg between the eighteenth of January and the eleventh, )

...

A09: Doch ich habe Zeit von sechsten Februar bis neunten Februar.  
(SUGGEST\_SUPPORT\_DATE)  
(I have time afterall from the 6th of February to the 9th of February)

B10: Sehr gut // das paßt bei mir auch // Dann machen wir's gleich aus // für Donnerstag // den achten // Wie wäre es denn um acht Uhr dreißig //  
(FEEDBACK\_ACKNOWLEDGEMENT, ACCEPT\_DATE, INIT\_DATE, SUGGEST\_SUPPORT\_DATE, SUGGEST\_SUPPORT\_DATE, SUGGEST\_SUPPORT\_DATE)  
(Very good, that too suits me, we will arrange for it, for thursday, the eighth, how about half past eighth)

A11: Am achten // ginge es bei mir leider nur bis zehn Uhr // Bei mir geht es besser nachmittags.  
(SUGGEST\_SUPPORT\_DATE, SUGGEST\_SUPPORT\_DATE, ACCEPT\_DATE)  
(on the eighth, is it only unfortunately possible for me until 10 o'clock, it suits me better in the afternoon)

B12: gut // um wieviel Uhr sollen wir uns dann treffen ?  
(FEEDBACK\_ACKNOWLEDGEMENT, SUGGEST\_SUPPORT\_DATE)  
(good, when should we meet)

A13: ich v"urde "ahm vierzehn Uhr vorzuschlagen // geht es bei Ihnen.  
(SUGGEST\_SUPPORT\_DATE, REQUEST\_COMMENT\_DATE)  
( I would suggest 2 o'clock, is that possible for you?)

B14: sehr gut // das pa"st bei mir auch // das k"onnen wir festhalten  
(ACCEPT\_DATE, ACCEPT\_DATE, ACCEPT\_DATE)  
(very good, that suits me too, we can make a note of that)

## DAMSL: forward looking func.

STATEMENT	a claim made by the speaker
INFO-REQUEST	a question by the speaker
CHECK	a question for confirming information
INFLUENCE-ON-ADDRESSEE	(=Searle's directives)
OPEN-OPTION	a weak suggestion or listing of options
ACTION-DIRECTIVE	an actual command
INFLUENCE-ON-SPEAKER	(=Austin's commissives)
OFFER	speaker offers to do something
COMMIT	speaker is committed to doing something
CONVENTIONAL	other
OPENING	greetings
CLOSING	farewells
THANKING	thanking and responding to thanks

2/17/05

CS 224S Winter 2005

42

## DAMSL: backward looking func.

<b>AGREEMENT</b>	speaker's response to previous proposal
<b>ACCEPT</b>	accepting the proposal
<b>ACCEPT-PART</b>	accepting some part of the proposal
<b>MAYBE</b>	neither accepting nor rejecting the proposal
<b>REJECT-PART</b>	rejecting some part of the proposal
<b>REJECT</b>	rejecting the proposal
<b>HOLD</b>	putting off response, usually via subdialogue
<b>ANSWER</b>	answering a question
<b>UNDERSTANDING</b>	whether speaker understood previous
<b>SIGNAL-NON-UNDER.</b>	speaker didn't understand
<b>SIGNAL-UNDER.</b>	speaker did understand
<b>ACK</b>	demonstrated via continuer or assessment
<b>REPEAT-REPHRASE</b>	demonstrated via repetition or reformulation
<b>COMPLETION</b>	demonstrated via collaborative completion

2/17/05

CS 224S Winter 2005

43

[assert]	C <sub>1</sub> : ... I need to travel in May.
[info-req,ack]	A <sub>1</sub> : And, what day in May did you want to travel?
[assert,answer]	C <sub>2</sub> : OK uh I need to be there for a meeting that's from the 12th to the 15th.
[info-req,ack]	A <sub>2</sub> : And you're flying into what city?
[assert,answer]	C <sub>3</sub> : Seattle.
[info-req,ack]	A <sub>3</sub> : And what time would you like to leave Pittsburgh?
[check,hold]	C <sub>4</sub> : Uh humm I don't think there's many options for non-stop.
[accept,ack]	A <sub>4</sub> : Right.
[assert]	There's three non-stops today.
[info-req]	C <sub>5</sub> : What are they?
[assert, open-option]	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.
[accept,ack]	C <sub>6</sub> : OK I'll take the 5ish flight on the night before on the 11th.
[check,ack]	A <sub>6</sub> : On the 11th?
[assert,ack]	OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
[ack]	C <sub>7</sub> : OK.

2/17/05

CS 224S Winter 2005

44

## Automatic Interpretation of Dialogue Acts

- How do we automatically identify dialogue acts?
- Given an utterance:
  - Decide whether it is a QUESTION, STATEMENT, SUGGEST, or ACK
- Recognizing illocutionary force will be crucial to building a dialogue agent
- Perhaps we can just look at the form of the utterance to decide?

2/17/05

CS 224S Winter 2005

45

## Can we just use the surface syntactic form?

- YES-NO-Q's have auxiliary-before-subject syntax:
  - Will breakfast be served on USAir 1557?
- STATEMENTS have declarative syntax:
  - I don't care about lunch
- COMMAND's have imperative syntax:
  - Show me flights from Milwaukee to Orlando on Thursday night

2/17/05

CS 224S Winter 2005

46

## Surface form != speech act type

	Locutionary Force	Illocutionary Force
Can I have the rest of your sandwich?	Question	Request
I want the rest of your sandwich	Declarative	Request
Give me your sandwich!	Imperative	Request

2/17/05

CS 224S Winter 2005

47

## Dialogue act disambiguation is hard!

- Who's on First - Abbott and Costello routine



2/17/05

CS 224S Winter 2005

48

## Dialogue act ambiguity

---

- Who's on first?
  - INFO-REQUEST
  - or
  - STATEMENT

2/17/05 CS 224S Winter 2005 49

## Dialogue Act ambiguity

---

- Can you give me a list of the flights from Atlanta to Boston?
  - This looks like an INFO-REQUEST.
  - If so, the answer is:
    - YES.
  - But really it's a DIRECTIVE or REQUEST, a polite form of:
    - Please give me a list of the flights...
- What looks like a QUESTION can be a REQUEST

2/17/05 CS 224S Winter 2005 50

## Dialogue Act ambiguity

---

- Similarly, what looks like a STATEMENT can be a QUESTION:

Us	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
Ag	HOLD	OK uh let me pull up your profile and I'll be right with you here. [pause]
Ag	CHECK	And you said you wanted to travel next week?
Us	ACCEPT	Uh yes.

2/17/05 CS 224S Winter 2005 51

## Indirect speech acts

---

- Utterances which use a surface statement to ask a question
- Utterances which use a surface question to issue a request

2/17/05 CS 224S Winter 2005 52

## DA interpretation as statistical classification

---

- Lots of clues in each sentence that can tell us which DA it is:
- Words and Collocations:
  - *Please* or *would you*: good cue for REQUEST
  - *Are you*: good cue for INFO-REQUEST
- Prosody:
  - Rising pitch is a good cue for INFO-REQUEST
  - Loudness/stress can help distinguish *yeah*/AGREEMENT from *yeah*/BACKCHANNEL
- Conversational Structure
  - *Yeah* following a proposal is probably AGREEMENT; *yeah* following an INFORM probably a BACKCHANNEL

2/17/05 CS 224S Winter 2005 53

## HMM model of dialogue act interpretation

---

- A dialogue is an HMM
- The hidden states are the dialogue acts
- The observation sequences are sentences
  - Each observation is one sentence
- The observation likelihood model is a word N-gram

2/17/05 CS 224S Winter 2005 54

## HMMs for dialogue act interpretation

- **Goal of HMM model:**

- to compute labeling of dialogue acts  $D = d_1, d_2, \dots, d_n$
- that is most probable given evidence  $E$

$$D^* = \underset{D}{\operatorname{argmax}} P(D|E) = \underset{D}{\operatorname{argmax}} \frac{P(W|D)P(E)}{P(E)}$$

$$= \underset{D}{\operatorname{argmax}} P(E|D)P(D)$$

2/17/05

CS 224S Winter 2005

55

## HMMs for dialogue act interpretation

$$D^* = \underset{D}{\operatorname{argmax}} P(E|D)P(D)$$

- Let  $W$  be word sequence in sentence and  $F$  be prosodic feature sequence
- Simplifying (wrong) independence assumption
- (What are implications of this?)

$$P(E|D) = P(F|D)P(W|D)$$

$$D^* = \underset{D}{\operatorname{argmax}} P(D)P(F|D)P(W|D)$$

2/17/05

CS 224S Winter 2005

56

## HMM model for dialogue

$$D^* = \underset{D}{\operatorname{argmax}} P(D)P(F|D)P(W|D)$$

- **Three components**

- $P(D)$ : probability of sequence of dialogue acts
- $P(F|D)$ : probability of prosodic sequence given one dialogue act
- $P(W|D)$ : probability of word string in a sentence given dialogue act

2/17/05

CS 224S Winter 2005

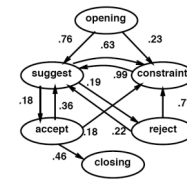
57

## $P(D)$

- **Markov assumption**
- Each dialogue act depends only on previous  $N$ . (In practice,  $N$  of 3 is enough).

$$P(D) = \prod_{i=2}^M P(d_i | d_{i-1}, \dots, d_{i-M+1})$$

- Woszczyna and Waibel (1994):



2/17/05

CS 224S Winter 2005

## $P(W|D)$

- Each dialogue act has different words
- Questions have "are you...", "do you...", etc

$$P(W|D) = \prod_{i=2}^N P(w_i | w_{i-1}, \dots, w_{i-N+1}, d_i)$$

2/17/05

CS 224S Winter 2005

59

## $P(F|D)$

- Shriberg et al. (1998)
- Decision tree trained on simple acoustically-based prosodic features
  - Slope of F0 at the end of the utterance
  - Average energy at different places in utterance
  - Various duration measures
  - All normalized in various ways
- These helped distinguish
  - Statement (S)
  - Yes-No-Question (QY)
  - Declarative-Question (QD)
  - Wh-Question (QW)

2/17/05

CS 224S Winter 2005

60

## Prosodic Decision Tree for making S/QY/QW/QD decision



## Getting likelihoods from decision tree

- Decision trees give posterior  $p(d|F)$  [discriminative, good]
- But we need  $p(F|d)$  to fit into HMM

$$p(d|F) = \frac{P(F|d)p(d)}{p(F)}$$

- Rearranging terms to get a likelihood:

$$p(F|d) = \frac{P(d|F)p(F)}{p(d)}$$

- scaled likelihood is ok since  $p(F)$  is constant

2/17/05

CS 224S Winter 2005

62

## Final HMM equation for dialogue act tagging

$$D^* = \operatorname{argmax}_D P(D)P(F|D)P(W|D)$$

$$= \prod_{i=2}^M P(d_i | d_{i-1} \dots d_{i-M+1}) \prod_{i=1}^N \frac{P(d_i | F)}{P(d_i)} \prod_{i=2}^N P(w_i | w_{i-1} \dots w_{i-N+1}, d_i)$$

- Then can use Viterbi decoding to find  $D^*$
- In real dialogue systems, obviously can't use FUTURE dialogue acts, so predict up to current act
- In rescoring passes (for example for labeling human-human dialogues for meeting summarization), can use future info.
- Most other supervised ML classifiers have been applied to DA tagging task

2/17/05

CS 224S Winter 2005

63

## On the way to more powerful dialogue systems

- Grounding**
  - Performing grounding
  - Recognizing user's grounding
- Dialogue Acts**
  - Using correct dialogue acts
  - Recognizing user's dialogue acts
- Planning and Intention**
  - Planning what to say
  - Recognizing user's intentions and plans

2/17/05

CS 224S Winter 2005

64

## Conversational Implicature

- 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.
- Note that client did not answer question.
- Meaning of client's sentence:
  - Meeting
    - Start-of-meeting: 12th
    - End-of-meeting: 15th
  - Doesn't say anything about flying!!!!
- What is it that licenses agent to infer that client is mentioning this meeting so as to inform the agent of the travel dates?

2/17/05

CS 224S Winter 2005

65

## Conversational Implicature (2)

- A:** ... there's 3 non-stops today.
- This would still be true if 7 non-stops today.
- But no, the agent means: 3 and only 3.
- How can client infer that agent means:
  - only 3

2/17/05

CS 224S Winter 2005

66

## Grice: conversational implicature

- Implicature means a particular class of licensed inferences.
- Grice (1975) proposed that what enables hearers to draw correct inferences is:
- Cooperative Principle
  - This is a tacit agreement by speakers and listeners to cooperate in communication

2/17/05

CS 224S Winter 2005

67

## 4 Gricean Maxims

- **Relevance:** Be relevant
- **Quantity:** Do not make your contribution more or less informative than required
- **Quality:** try to make your contribution one that is true (don't say things that are false or for which you lack adequate evidence)
- **Manner:** Avoid ambiguity and obscurity; be brief and orderly

2/17/05

CS 224S Winter 2005

68

## Relevance

- **A:** Is Regina here?
- **B:** Her car is outside.
- **Implication: yes**
  - **Hearer thinks:** why would he mention the car? It must be relevant. How could it be relevant? It could since if her car is here she is probably here.
- **Client:** I need to be there for a meeting that's from the 12th to the 15th
  - **Hearer thinks:** Speaker is following maxims, would only have mentioned meeting if it was relevant. How could meeting be relevant? If client meant me to understand that he had to depart in time for the mtg.

2/17/05

CS 224S Winter 2005

69

## Quantity

- **A:** How much money do you have on you?
- **B:** I have 5 dollars
  - **Implication:** not 6 dollars
- Similarly, 3 non-stops can't mean 7 non-stops (**hearer thinks:**
  - if speaker meant 7 non-stops she would have said 7 non-stops
- **A:** Did you do the reading for today's class?
- **B:** I intended to
  - **Implication:** No
  - B's answer would be true if B intended to do the reading AND did the reading, but would then violate maxim

2/17/05

CS 224S Winter 2005

70

## Planning-based Conversational Agents

- How to do the kind of Gricean inference that could solve the problems we've discussed?
- Researchers who work on this use sophisticated AI models of planning and reasoning.
- Involves planning, plus various extensions to logic to create logic for Belief, Desire, Intention.
- These are called BDI models (belief, desire, intention)

2/17/05

CS 224S Winter 2005

71

## Plan-based agent interpreting user utterance

- **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
- **Use the Principle of Relevance**
  - System knows that one precondition for having a meeting is being at the place where the meeting is
  - One way of being at a place is flying there
  - Booking a flight is a precondition for flying there.
  - System follows this chain of inference, abducting that user wants to fly on a date before the 12th

2/17/05

CS 224S Winter 2005

72

## Plan-based agent producing system utterance

- **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
- **Inference chain:**
  - In order to help a client book a flight, must know enough information about the flight to book it
  - Knowing the month (May) is insufficient information to specify a departure or return date
  - The simplest way to find out the needed date information is simply to ask the client

2/17/05

CS 224S Winter 2005

73

## BDI Logic

- $B(S,P)$  = "speaker S believes proposition P"
- $KNOW(S,P) = P$  and  $B(S,P)$
- $KNOWIF(S,P)$  = "S knows whether P" =  $KNOW(S,P)$  or  $KNOW(S,notP)$
- $W(S,P)$  "S wants P to be true", where P is a state or the execution of some action
- $W(S,ACT(H)) = S$  wants H to do ACT

2/17/05

CS 224S Winter 2005

74

## How to represent actions

- **Preconditions:**
  - Conditions that must already be true in order to successfully perform the action
- **Effects:**
  - conditions that become true as a result of successfully performing the action
- **Body:**
  - A set of partially ordered goal states that must be achieved in performing the action

2/17/05

CS 224S Winter 2005

75

## How to represent the action of going to the beach

- $GOTOBEACH(P,B)$
- **Constraints:** Person(P) & Beach(B) & Car(C)
- **Precondition:** Know(P,location(B)) & Have(A,C) & working(C) & Want(P,AtBeach(P,B)) &...
- **Effect:** AtBeach(P,B)
- **Body:** Drive(P,C)

2/17/05

CS 224S Winter 2005

76

## How to represent the action of booking a flight

- $BOOK-FLIGHT(A,C,F)$
- **Constraints:** Agent(A) & Flight(F) & Client(C)
- **Precondition:** Know(A,dep-date(F)) & Know(A,dep-time(F)) & Know(A,origin(F)) & Has-Seats(F) & W(C,BOOK,A,C,F) & ...
- **Effect:** Flight-Booked(A,C,F)
- **Body:** Make-Reservation(A,F,C)

2/17/05

CS 224S Winter 2005

77

## Speech acts

- $INFORM(S,H,P)$
- **Constraints:** Speaker(S) & Hearer(H) & Proposition(P)
- **Precondition:** Know(S,P) & W(S,INFORM(S,H,P))
- **Effect:** Know(H,P)
- **Body:** B(H(W(S,Know(H,P))))

2/17/05

CS 224S Winter 2005

78

## Speech acts

- **REQUEST-INFORM(A,C,I)**
- **Constraints:** Agent(A) & Client(C)
- **Precondition:** Know(C,I)
- **Effect:** Know(A,I)
- **Body:** B(C(W(A,Know(A,I))))

2/17/05

CS 224S Winter 2005

79

## How a plan-based conversational agent works

- While conversation is not finished
  - If user has completed a turn
    - Then interpret user's utterance
  - If system has obligations
    - Then address obligations
  - Else if system has turn
    - Then if system has intended conversation acts
      - Then call generator to produce utterances
    - Else if some material is ungrounded
      - Then address grounding situation
    - Else if high-level goals are unsatisfied
      - Then address goals
    - Else release turn or attempt to end conversation
  - Else if no one has turn or long pause
    - Then take turn

2/17/05

CS 224S Winter 2005

80

## Plan-based agent data

- Queue of conversation acts it needs to generate, based on:
- **Grounding:** need to ground previous utterance
- **Dialogue obligations:** answer questions, perform commands
- **Goals:** agent must reason about its own goals

2/17/05

CS 224S Winter 2005

81

## A made-up example

- **C:** I want to go to Pittsburgh in May
- **System current state:**
  - Discourse obligations: **NONE**
  - Turn holder: **system**
  - Intended speech acts: **NONE**
  - Unacknowledged speech acts: **INFORM-1**
  - Discourse goals: **get-travel-goal, create-travel-plan**

2/17/05

CS 224S Winter 2005

82

## A made-up example

- **System decides to add 2 conversation acts to queue:**
  - Acknowledge user's inform act
  - Ask next travel-goal question of user
- **How?**
  - Given goal "get-travel-goal"
  - Request-info action scheme tells system that asking the user something is one way of finding out.

2/17/05

CS 224S Winter 2005

83

## A made-up example

- **System current state:**
  - Discourse obligations: **NONE**
  - Turn holder: **system**
  - Intended speech acts: **REQUEST-INFORM-1, ACKNOWLEDGE-1**
  - Unacknowledged speech acts: **INFORM-1**
  - Discourse goals: **get-travel-goal, create-travel-plan**
- This would be combined by clever generator:
  - And, what day in May did you want to travel

2/17/05

CS 224S Winter 2005

84

## A made-up example

- C .. I don't think there's many options for non-stop.
- Assume DA interpreter correctly interprets this as REQUEST-INFORM3
  - Discourse obligations: address(REQUEST-INFORM3)
  - Turn holder: system
  - Intended speech acts: NONE
  - Unacknowledged speech acts: REQUEST-INFORM-3
  - Discourse goals: get-travel-goal, create-travel-plan
- Manager would address discourse goal by calling planner to find out how many non-stop flights there are. Also needs to ground.

2/17/05

CS 224S Winter 2005

85

## A made-up example

- C .. I don't think there's many options for non-stop.
- Since this was in the form of an indirect request, we can do an ACKNOWLEDGEMENT (if a direct request, we would do ANSWER-YES). Also need to answer the question:
- Right. There's three non-stops today.

2/17/05

CS 224S Winter 2005

86

## Summary

- Dialogue manager architectures
  - Finite-state, Frame-based, Plan-based
    - (in 2 weeks: Markov Decision Process)
- Grounding,
  - Explicit vs. implicit confirmation
  - Rejection, reprompting
  - Recognizing and handling user corrections
- Dialogue Acts
  - Cues for them
  - HMM model for dialogue act interpretation
- Planning and Intention
  - Gricean implicature
  - BDI model of dialogue acts
  - Planning for understanding and planning for production

2/17/05

CS 224S Winter 2005

87