(Hidden) Information State Models

Ling575
Discourse and Dialogue
May 25, 2011

Roadmap

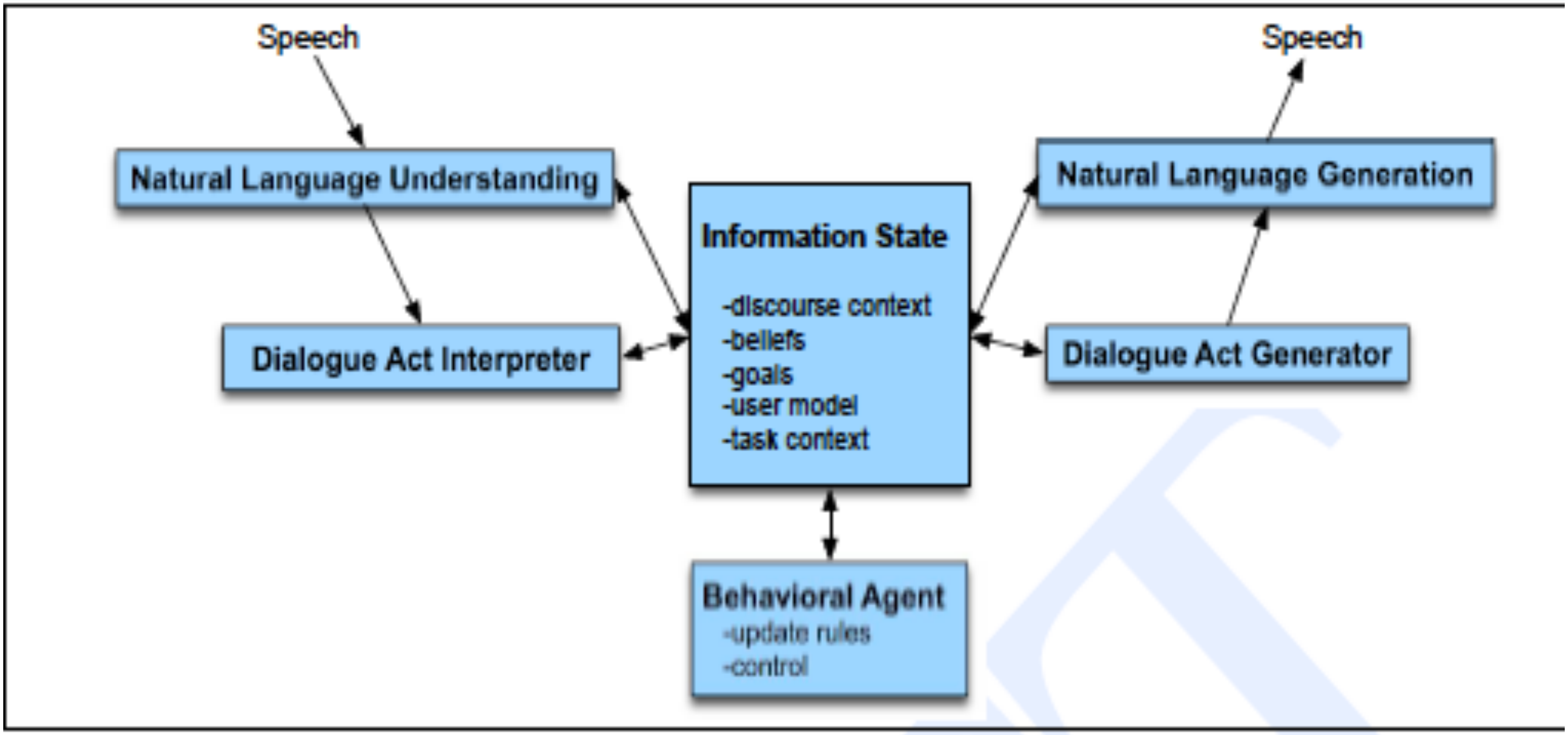
- Information State Models
  - Dialogue Acts
  - Dialogue Act Recognition
- Hidden Information State Models
  - Learning dialogue behavior
- Politeness and Speaking Style
  - Generating styles

Information State Systems

- Information state :
  - Discourse context, grounding state, intentions, plans.
- Dialogue acts:
  - Extension of speech acts, to include grounding acts
    - Request-inform; Confirmation
- Update rules
  - Modify information state based on DAs
    - When a question is asked, answer it
    - When an assertion is made,
      - Add information to context, grounding state

Information State Architecture

- Simple ideas, complex execution



Dialogue Acts

- Extension of speech acts
  - Adds structure related to conversational phenomena
    - Grounding, adjacency pairs, etc

Dialogue Acts

- Extension of speech acts
  - Adds structure related to conversational phenomena
    - Grounding, adjacency pairs, etc
- Many proposed tagsets
  - Verbmobil: acts specific to meeting sched domain

Dialogue Acts

- Extension of speech acts
  - Adds structure related to conversational phenomena
    - Grounding, adjacency pairs, etc
- Many proposed tagsets
  - Verbmobil: acts specific to meeting sched domain
  - DAMSL: Dialogue Act Markup in Several Layers
    - Forward looking functions: speech acts
    - Backward looking function: grounding, answering

Dialogue Acts

- Extension of speech acts
  - Adds structure related to conversational phenomena
    - Grounding, adjacency pairs, etc
- Many proposed tagsets
  - Verbmobil: acts specific to meeting sched domain
  - DAMSL: Dialogue Act Markup in Several Layers
    - Forward looking functions: speech acts
    - Backward looking function: grounding, answering
  - Conversation acts:
    - Add turn-taking and argumentation relations

Verbmobil DA

- 18 high level tags

Tag	Example
THANK	*Thanks*
GREET	*Hello Dan*
INTRODUCE	*It's me again*
BYE	*Allright bye*
REQUEST-COMMENT	*How does that look?*
SUGGEST	*from thirteenth through seventeenth June*
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?*
DIGRESS	*[we could meet for lunch] and eat lots of ice cream*
MOTIVATE	*We should go to visit our subsidiary in Munich*
GARBAGE	*Oops, I-*

Figure 24.17 The 18 high-level dialogue acts used in Verbmobil-1, abstracted over a total of 43 more specific dialogue acts. Examples are from Jekat et al. (1995).

Dialogue Act Interpretation

- Automatically tag utterances in dialogue
- Some simple cases:
  - Will breakfast be served on USAir 1557?

Dialogue Act Interpretation

- Automatically tag utterances in dialogue
- Some simple cases:
  - **YES-NO-Q**: Will breakfast be served on USAir 1557?
  - I don't care about lunch.

Dialogue Act Interpretation

- Automatically tag utterances in dialogue
- Some simple cases:
  - **YES-NO-Q:** Will breakfast be served on USAir 1557?
  - **Statement:** I don't care about lunch.
  - Show be flights from L.A. to Orlando

Dialogue Act Interpretation

- Automatically tag utterances in dialogue
- Some simple cases:
  - **YES-NO-Q:** Will breakfast be served on USAir 1557?
  - **Statement:** I don't care about lunch.
  - **Command:** Show be flights from L.A. to Orlando
- Is it always that easy?
  - Can you give me the flights from Atlanta to Boston?

Dialogue Act Interpretation

- Automatically tag utterances in dialogue
- Some simple cases:
  - **YES-NO-Q:** Will breakfast be served on USAir 1557?
  - **Statement:** I don't care about lunch.
  - **Command:** Show be flights from L.A. to Orlando
- Is it always that easy?
  - Can you give me the flights from Atlanta to Boston?
    - Syntactic form: question; Act: request/command
  - Yeah.

Dialogue Act Interpretation

- Automatically tag utterances in dialogue
- Some simple cases:
  - **YES-NO-Q:** Will breakfast be served on USAir 1557?
  - **Statement:** I don't care about lunch.
  - **Command:** Show be flights from L.A. to Orlando
- Is it always that easy?
  - Can you give me the flights from Atlanta to Boston?
  - Yeah.
    - Depends on context: Y/N answer; agreement; back-channel

Dialogue Act Ambiguity

- Indirect speech acts

A	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.
B	OK uh let me pull up your profile and I'll be right with you here. [pause]
B	**And you said you wanted to travel next week?**
A	Uh yes.

Dialogue Act Ambiguity

- Indirect speech acts

A	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.
B		OK uh let me pull up your profile and I'll be right with you here. [pause]
B		**And you said you wanted to travel next week?**
A		Uh yes.

Dialogue Act Ambiguity

- Indirect speech acts

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

Dialogue Act Ambiguity

- Indirect speech acts

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

Dialogue Act Ambiguity

- Indirect speech acts

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

Dialogue Act Recognition

- How can we classify dialogue acts?
- Sources of information:

Dialogue Act Recognition

- How can we classify dialogue acts?
- Sources of information:
  - Word information:
    - *Please, would you*: request; *are you*: yes-no question

Dialogue Act Recognition

- How can we classify dialogue acts?
- Sources of information:
  - Word information:
    - *Please, would you*: request; *are you*: yes-no question
    - N-gram grammars
  - Prosody:

Dialogue Act Recognition

- How can we classify dialogue acts?
- Sources of information:
  - Word information:
    - *Please, would you*: request; *are you*: yes-no question
    - N-gram grammars
  - Prosody:
    - Final rising pitch: question; final lowering: statement
    - Reduced intensity: *Yeah:* agreement vs backchannel

Dialogue Act Recognition

- How can we classify dialogue acts?
- Sources of information:
  - Word information:
    - *Please, would you*: request; *are you*: yes-no question
    - N-gram grammars
  - Prosody:
    - Final rising pitch: question; final lowering: statement
    - Reduced intensity: *Yeah:* agreement vs backchannel
  - Adjacency pairs:

Dialogue Act Recognition

- How can we classify dialogue acts?
- Sources of information:
  - Word information:
    - *Please, would you*: request; *are you*: yes-no question
    - N-gram grammars
  - Prosody:
    - Final rising pitch: question; final lowering: statement
    - Reduced intensity: *Yeah:* agreement vs backchannel
  - Adjacency pairs:
    - Y/N question, agreement vs Y/N question, backchannel
    - DA bi-grams

Task & Corpus

- Goal:
  - Identify dialogue acts in conversational speech

Task & Corpus

- Goal:
  - Identify dialogue acts in conversational speech
- Spoken corpus: Switchboard
  - Telephone conversations between strangers
  - Not task oriented; topics suggested
  - 1000s of conversations
    - recorded, transcribed, segmented

Dialogue Act Tagset

- Cover general conversational dialogue acts
  - No particular task/domain constraints

Dialogue Act Tagset

- Cover general conversational dialogue acts
  - No particular task/domain constraints
- Original set: ~50 tags
  - Augmented with flags for task, conv mgmt
    - 220 tags in labeling: some rare

Dialogue Act Tagset

- Cover general conversational dialogue acts
  - No particular task/domain constraints
- Original set: ~50 tags
  - Augmented with flags for task, conv mgmt
    - 220 tags in labeling: some rare
- Final set: 42 tags, mutually exclusive
  - SWBD-DAMSL
  - Agreement: K=0.80 (high)

Dialogue Act Tagset

- Cover general conversational dialogue acts
  - No particular task/domain constraints
- Original set: ~50 tags
  - Augmented with flags for task, conv mgmt
    - 220 tags in labeling: some rare
- Final set: 42 tags, mutually exclusive
  - SWBD-DAMSL
  - Agreement: K=0.80 (high)
- 1,155 conv labeled: split into train/test

Common Tags

- **Statement** & **Opinion**: declarative +/- op
- **Question**: Yes/No&Declarative: form, force
- **Backchannel:** Continuers like uh-huh, yeah
- **Turn Exit/Adandon:** break off, +/- pass
- **Answer :** Yes/No, follow questions
- **Agreement**: Accept/Reject/Maybe

Probabilistic Dialogue Models

- HMM dialogue models

Probabilistic Dialogue Models

- HMM dialogue models
  - States = Dialogue acts; Observations: Utterances
    - Assume decomposable by utterance
    - Evidence from true words, ASR words, prosody

$$d^* = \operatorname*{argmax}_d P(d \mid o) = \operatorname*{argmax}_d \frac{P(o \mid d)P(d)}{P(o)} = \operatorname*{argmax}_d P(o \mid d)P(d)$$

Probabilistic Dialogue Models

- HMM dialogue models
  - States = Dialogue acts; Observations: Utterances
    - Assume decomposable by utterance
    - Evidence from true words, ASR words, prosody

$$d^* = \operatorname*{argmax}_d P(d \mid o) = \operatorname*{argmax}_d \frac{P(o \mid d)P(d)}{P(o)} = \operatorname*{argmax}_d P(o \mid d)P(d)$$

$$P(o \mid d) = P(f \mid d)P(W \mid d)$$

Probabilistic Dialogue Models

- HMM dialogue models
  - States = Dialogue acts; Observations: Utterances
    - Assume decomposable by utterance
    - Evidence from true words, ASR words, prosody

$$d^* = \underset{d}{\operatorname{argmax}} P(d \mid o) = \underset{d}{\operatorname{argmax}} \frac{P(o \mid d)P(d)}{P(o)} = \underset{d}{\operatorname{argmax}} P(o \mid d)P(d)$$

$$P(o \mid d) = P(f \mid d)P(W \mid d)$$

$$P(W \mid d) = \prod_{i=2}^{N} P(w_i \mid w_{i-1}, w_{i-2} \ldots w_{i-N+1}, d)$$

Probabilistic Dialogue Models

- HMM dialogue models
  - States = Dialogue acts; Observations: Utterances
    - Assume decomposable by utterance
    - Evidence from true words, ASR words, prosody

$$d^* = \underset{d}{\operatorname{argmax}} P(d \mid o) = \underset{d}{\operatorname{argmax}} \frac{P(o \mid d)P(d)}{P(o)} = \underset{d}{\operatorname{argmax}} P(o \mid d)P(d)$$

$$P(o \mid d) = P(f \mid d)P(W \mid d)$$

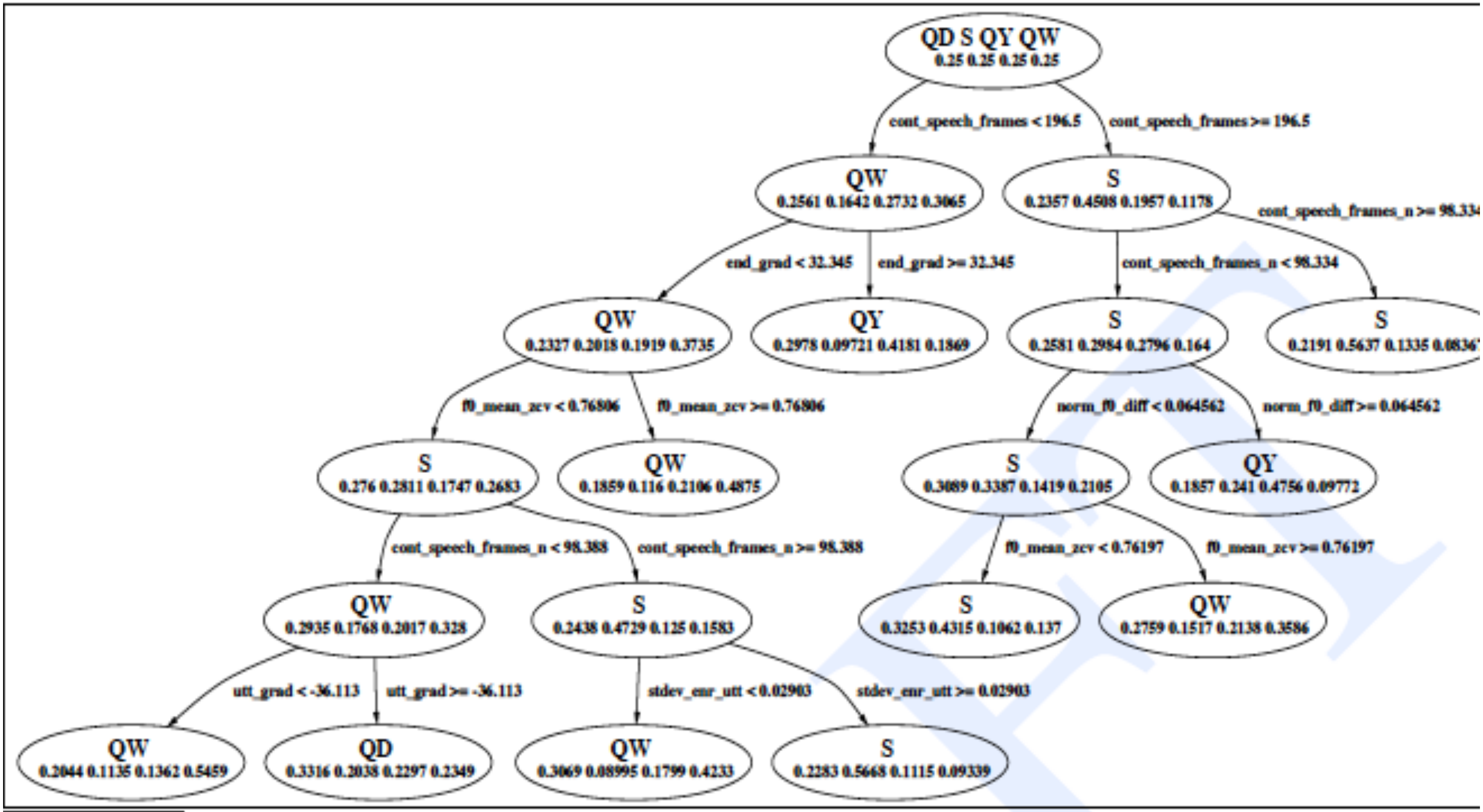
$$P(W \mid d) = \prod_{i=2}^{N} P(w_i \mid w_{i-1}, w_{i-2} \ldots w_{i-N+1}, d)$$

$$d^* = \underset{d}{\operatorname{argmax}} P(d \mid d_{t-1})P(f \mid d)P(W \mid d)$$

DA Classification - Prosody

- Features:
  - Duration, pause, pitch, energy, rate, gender
    - Pitch accent, tone
- Results:
  - Decision trees: 5 common classes
    - 45.4% - baseline=16.6%

Prosodic Decision Tree



DA Classification -Words

- Words
  - Combines notion of discourse markers and collocations:
    - e.g. uh-huh=Backchannel
  - Contrast: true words, ASR 1-best, ASR n-best
- Results:
  - Best: 71%- true words, 65% ASR 1-best

DA Classification - All

- Combine word and prosodic information
  - Consider case with ASR words and acoustics

DA Classification - All

- Combine word and prosodic information
  - Consider case with ASR words and acoustics
  - Prosody classified by decision trees
    - Incorporate decision tree posteriors in model for $P(f|d)$

DA Classification - All

- Combine word and prosodic information
  - Consider case with ASR words and acoustics
  - Prosody classified by decision trees
    - Incorporate decision tree posteriors in model for P(f|d)

$$d^* = P(d \mid d_{t-1}) \frac{P(d \mid f)}{P(d)} \prod_{i=2}^{N} P(w_i \mid w_{i-1} \ldots w_{i-N+1}, d)$$

- Slightly better than raw ASR

Integrated Classification

- Focused analysis
  - Prosodically disambiguated classes
    - Statement/Question-Y/N and Agreement/Backchannel
    - Prosodic decision trees for agreement vs backchannel
      - Disambiguated by duration and loudness

Integrated Classification

- Focused analysis
  - Prosodically disambiguated classes
    - Statement/Question-Y/N and Agreement/Backchannel
    - Prosodic decision trees for agreement vs backchannel
      - Disambiguated by duration and loudness
  - Substantial improvement for prosody+words
    - True words: S/Q: 85.9%-> 87.6; A/B: 81.0%->84.7

Integrated Classification

- Focused analysis
  - Prosodically disambiguated classes
    - Statement/Question-Y/N and Agreement/Backchannel
    - Prosodic decision trees for agreement vs backchannel
      - Disambiguated by duration and loudness
  - Substantial improvement for prosody+words
    - True words: S/Q: 85.9%-> 87.6; A/B: 81.0%->84.7
    - ASR words: S/Q: 75.4%->79.8; A/B: 78.2%->81.7
  - More useful when recognition is iffy

Many Variants

- Maptask: (13 classes)
  - Serafin & DiEugenio 2004
    - Latent Semantic analysis on utterance vectors
    - Text only
    - Game information; No improvement for DA history

Many Variants

- Maptask: (13 classes)
  - Serafin & DiEugenio 2004
    - Latent Semantic analysis on utterance vectors
    - Text only
    - Game information; No improvement for DA history
  - Surendran & Levow 2006
    - SVMs on term n-grams, prosody
    - Posteriors incorporated in HMMs
      - Prosody, sequence modeling improves

Many Variants

- Maptask: (13 classes)
  - Serafin & DiEugenio 2004
    - Latent Semantic analysis on utterance vectors
    - Text only
    - Game information; No improvement for DA history
  - Surendran & Levow 2006
    - SVMs on term n-grams, prosody
    - Posteriors incorporated in HMMs
      - Prosody, sequence modeling improves
- MRDA: Meeting tagging: 5 broad classes

Observations

- DA classification can work on open domain
  - Exploits word model, DA context, prosody
  - Best results for prosody+words
  - Words are quite effective alone – even ASR
- Questions:

Observations

- DA classification can work on open domain
  - Exploits word model, DA context, prosody
  - Best results for prosody+words
  - Words are quite effective alone – even ASR
- Questions:
  - Whole utterance models? – more fine-grained
  - Longer structure, long term features

Detecting Correction Acts

- Miscommunication is common in SDS
  - Utterances after errors misrecognized >2x as often
    - Frequently repetition or paraphrase of original input

Detecting Correction Acts

- Miscommunication is common in SDS
  - Utterances after errors misrecognized >2x as often
    - Frequently repetition or paraphrase of original input
- Systems need to detect, correct

Detecting Correction Acts

- Miscommunication is common in SDS
  - Utterances after errors misrecognized >2x as often
    - Frequently repetition or paraphrase of original input
- Systems need to detect, correct
- Corrections are spoken differently:
  - Hyperarticulated (slower, clearer) -> lower ASR conf.

Detecting Correction Acts

- Miscommunication is common in SDS
  - Utterances after errors misrecognized >2x as often
    - Frequently repetition or paraphrase of original input
- Systems need to detect, correct
- Corrections are spoken differently:
  - Hyperarticulated (slower, clearer) -> lower ASR conf.
  - Some word cues: ‘No’,’ I meant’, swearing..

Detecting Correction Acts

- Miscommunication is common in SDS
  - Utterances after errors misrecognized >2x as often
    - Frequently repetition or paraphrase of original input
- Systems need to detect, correct
- Corrections are spoken differently:
  - Hyperarticulated (slower, clearer) -> lower ASR conf.
  - Some word cues: 'No',' I meant', swearing..
- Can train classifiers to recognize with good acc.

Generating Dialogue Acts

- Generation neglected relative to generation

Generating Dialogue Acts

- Generation neglected relative to generation
- Stent (2002) model: Conversation acts, Belief model
  - Develops update rules for content planning, e.g.
    - If user releases turn, system can do 'TAKE-TURN' act
    - If system needs to summarize, use ASSERT act

Generating Dialogue Acts

- Generation neglected relative to generation
- Stent (2002) model: Conversation acts, Belief model
  - Develops update rules for content planning, i.e.
    - If user releases turn, system can do ‘TAKE-TURN’ act
    - If system needs to summarize, use ASSERT act
  - Identifies turn-taking as key aspect of dialogue gen.

Cue	Turn-taking acts signaled
um	KEEP-TURN, TAKE-TURN, RELEASE-TURN
<lipsmack>, <click>, so, uh	KEEP-TURN, TAKE-TURN
you know, isn’t that so	ASSIGN-TURN

Figure 24.21 Language used to perform turn-taking acts, from Stent (2002).

Generating Confirmation

- Simple systems use fixed confirmation strategy
  - Implicit or explicit

Generating Confirmation

- Simple systems use fixed confirmation strategy
  - Implicit or explicit
- More complex systems can select dynamically
  - Use information state and features to decide

Generating Confirmation

- Simple systems use fixed confirmation strategy
  - Implicit or explicit
- More complex systems can select dynamically
  - Use information state and features to decide
    - Likelihood of error:
      - Low ASR confidence score
        - If very low, can reject

Generating Confirmation

- Simple systems use fixed confirmation strategy
  - Implicit or explicit
- More complex systems can select dynamically
  - Use information state and features to decide
    - Likelihood of error:
      - Low ASR confidence score
        - If very low, can reject
      - Sentence/prosodic features: longer, initial pause, pitch range

Generating Confirmation

- Simple systems use fixed confirmation strategy
  - Implicit or explicit
- More complex systems can select dynamically
  - Use information state and features to decide
    - Likelihood of error:
      - Low ASR confidence score
        - If very low, can reject
      - Sentence/prosodic features: longer, initial pause, pitch range
    - Cost of error:

Generating Confirmation

- Simple systems use fixed confirmation strategy
  - Implicit or explicit
- More complex systems can select dynamically
  - Use information state and features to decide
    - Likelihood of error:
      - Low ASR confidence score
        - If very low, can reject
      - Sentence/prosodic features: longer, initial pause, pitch range
    - Cost of error:
      - Book a flight vs looking up information
- Markov Decision Process models more detailed

Statistical Dialogue Management

- Pioneered by Steve Young's group at Cambridge
- Model dialogue as probabilistic agent
  - Markov Decision Process (MDP)

Statistical Dialogue Management

- Pioneered by Steve Young's group at Cambridge
- Model dialogue as probabilistic agent
  - Markov Decision Process (MDP)
- Characterized by:
  - S: set of **states** agent can be in
  - A: set of **actions** the agent can take
  - R(a,s): **reward** agent gets for action a in state s

Statistical Dialogue Management

- Pioneered by Steve Young's group at Cambridge
- Model dialogue as probabilistic agent
  - Markov Decision Process (MDP)
- Characterized by:
  - S: set of **states** agent can be in
  - A: set of **actions** the agent can take
  - R(a,s): **reward** agent gets for action a in state s
- Learn:
  - π : **Policy**: Which action a should agent in state s take to achieve highest reward?

Dialogue States

- Encapsulate information about current dialogue
  - History:
    - Everything (all states) so far?

Dialogue States

- Encapsulate information about current dialogue
  - History:
    - Everything (all states) so far?
      - Explosive

Dialogue States

- Encapsulate information about current dialogue
  - History:
    - Everything (all states) so far?
      - Explosive
    - Markov assumptions
  - Typically:
    - Value of current frame slots, Most recent system question
    - Most recent user answer, ASR confidence, etc

Dialogue States

- Encapsulate information about current dialogue
  - History:
    - Everything (all states) so far?
      - Explosive
    - Markov assumptions
  - Typically:
    - Value of current frame slots, Most recent system question
    - Most recent user answer, ASR confidence, etc
  - For day, month frame:
    - 411 states!

Actions & Rewards

- For day, month input:

Actions & Rewards

- For day, month input:
  - A_1: question asking for day

Actions & Rewards

- For day, month input:
  - A_1: question asking for day
  - A_2: question asking for month
  - A_3: question asking for day and month
  - A_4: submitting the form

Actions & Rewards

- For day, month input:
  - A_1: question asking for day
  - A_2: question asking for month
  - A_3: question asking for day and month
  - A_4: submitting the form
- Reward:
  - Correct answer with shortest interaction
  - $R = (w_i n_i + w_c n_c + w_f n_{f)}$
    - N_i:# interactions; n_c:# errors; n_f: # filled slots

Policies

- 1) Asking for Day, Month together
- 2) Asking for Day, Month separately
- Compute reward for each policy, given some P(error)

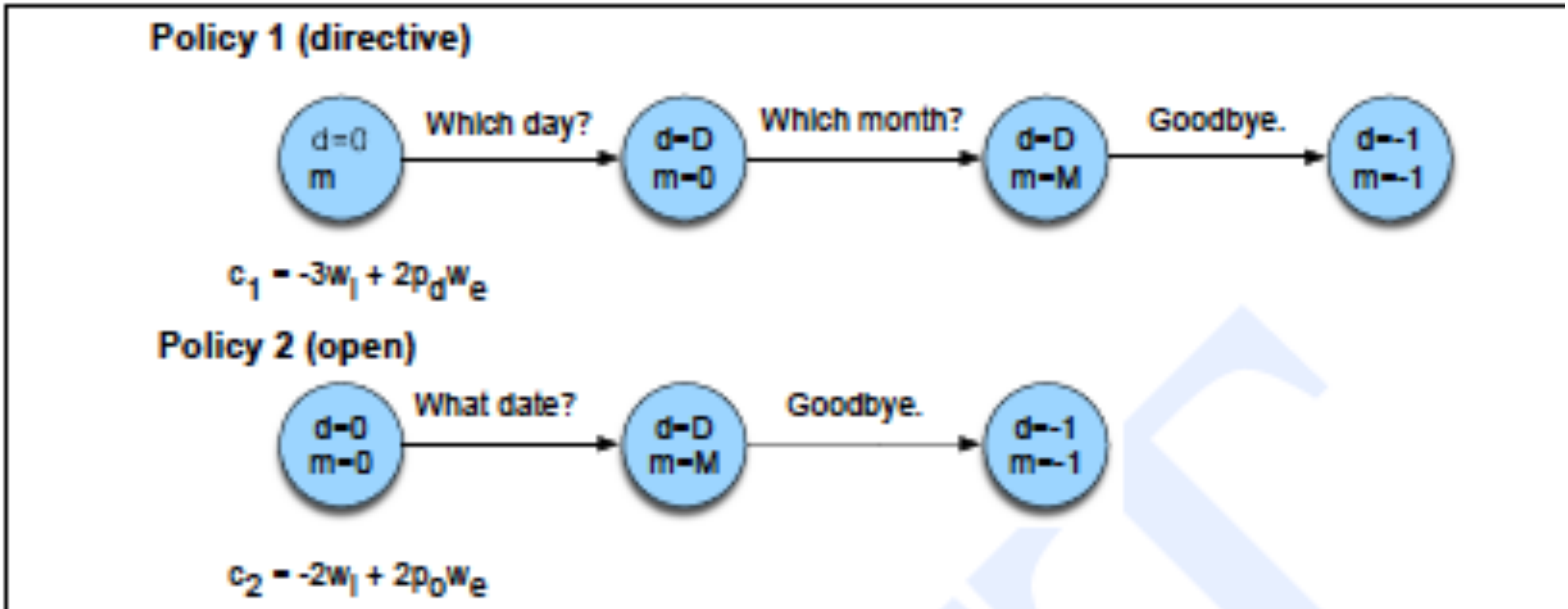


Figure 24.22 Two policies for getting a month and a day. After Levin et al. (2000).

Utility

- A utility function
  - maps a state or state sequence
  - onto a real number
  - describing the goodness of that state
  - I.e. the resulting "happiness" of the agent

Utility

- A utility function
  - maps a state or state sequence
  - onto a real number
  - describing the goodness of that state
  - I.e. the resulting "happiness" of the agent
- Principle of Maximum Expected Utility:
  - A rational agent should choose an action that maximizes the agent's expected utility

Learning Policies

- Simple system:
  - Can enumerate policies and select

Learning Policies

- Simple system:
  - Can enumerate policies and select
- Complex system:

Learning Policies

- Simple system:
  - Can enumerate policies and select
- Complex system:
  - Huge number of actions, states, policies
  - Selection is complex optimization problem
  - Can describe expected cumulative reward w/Bellman eqn
    - Standard approach in reinforcement learning
    - Solvable with value iteration algorithm

$$Q(a,s) = R(s,a) + \gamma \sum_{s'} P(s' \mid s,a) \max_{a'} Q(s',a')$$

Training the Model

- State transition probabilities must be estimated
  - For small corpus

Training the Model

- State transition probabilities must be estimated
  - For small corpus
    - Get real users for system
    - Compute results for different choices (i.e. initiative)
    - Directly collect empirical estimate
  - For larger system,

Training the Model

- State transition probabilities must be estimated
  - For small corpus
    - Get real users for system
    - Compute results for different choices (i.e. initiative)
    - Directly collect empirical estimate
  - For larger system, too many alternatives
    - Need arbitrary number of users

Training the Model

- State transition probabilities must be estimated
  - For small corpus
    - Get real users for system
    - Compute results for different choices (i.e. initiative)
    - Directly collect empirical estimate
  - For larger system, too many alternatives
    - Need arbitrary number of users
    - Simulation!!
    - Stochastic state selection

Training the Model

- State transition probabilities must be estimated
  - For small corpus
    - Get real users for system
    - Compute results for different choices (i.e. initiative)
    - Directly collect empirical estimate
  - For larger system, too many alternatives
    - Need arbitrary number of users
    - Simulation!!
    - Stochastic state selection
- Learned policies can outperform hand-crafted

Politeness & Speaking Style

Agenda

- Motivation
- Explaining politeness & indirectness
  - Face & rational reasoning
  - Defusing Face Threatening Acts
- Selecting & implementing speaking styles
  - Plan-based speech act modeling
  - Socially appropriate speaking styles

Why be Polite to Computers?

- Computers don't have feelings, status, etc

- Would people be polite to a machine?

Why be Polite to Computers?

- Computers don't have feelings, status, etc
- Would people be polite to a machine?
  - Range of politeness levels:
  - Direct < Hinting < Conventional Indirectness
- Why?

Varying Politeness

- Direct Requests:

Varying Politeness

- Direct Requests:
  - Read it to me
  - Go to the next group
  - Next message
- Polite Requests: Conventional Indirectness

Varying Politeness

- Direct Requests:
  - Read it to me
  - Go to the next group
  - Next message
- Polite Requests: Conventional Indirectness
  - I'd like to check Nicole's calendar
  - Could I have the short term forecast for Boston?
  - Weather please
- Goodbye spirals

Why are People Polite to Each Other?

Why are People Polite to Each Other?

- "Convention"
  - Begs the question - why become convention?
- Indirectness

Why are People Polite to Each Other?

- “Convention”
  - Begs the question - why become convention?
- Indirectness
  - Not just adding as many hedges as possible
    - “Could someone maybe please possibly be able to..”

Why are People Polite to Each Other?

- "Convention"
  - Begs the question - why become convention?
- Indirectness
  - Not just adding as many hedges as possible
    - "Could someone maybe please possibly be able to.."
- Social relation and rational agency
  - Maintaining face, rational reasoning
    - Pragmatic clarity

Face

- Kernel of politeness
  - Cross-cultural
- Public self-image

Face

- Kernel of politeness
  - Cross-cultural
- Public self-image
  - Negative: Claim of freedom to action, from imposition
    - “Want” to be unimpeded by others: “Autonomy”

Face

- Kernel of politeness
  - Cross-cultural
- Public self-image
  - Negative: Claim of freedom to action, from imposition
    - “Want” to be unimpeded by others: “Autonomy”
  - Positive: Desire to be approved of
    - “Want” to be liked - usually by specific people for specific attr
- Generally cooperate to preserve face
  - Mutually vulnerable

Rational Reasoning

- Guarantee inferences from ends to means that satisfy those ends
  - Means to end is satisfactory only if
    - means is true implies end is true
  - Ability to weigh different means
    - Preference operator to select
  - Notion of least-cost satisfiability
    - No wasted effort

Threatening & Saving Face

- Communicative acts may threaten face
  - Negative:

Threatening & Saving Face

- Communicative acts may threaten face
  - Negative: Put pressure on H to do, accept
    - E.g. request, suggest, remind, offer, compliment,..
  - Positive

Threatening & Saving Face

- Communicative acts may threaten face
  - Negative: Put pressure on H to do, accept
    - E.g. request, suggest, remind, offer, compliment,..
  - Positive: Indicate dislike or indifference to face
    - E.g. criticism, disapproval, contradiction, boasting
  - Threats to H’s or S’s face; positive/negative

Threatening & Saving Face

- Communicative acts may threaten face
  - Negative: Put pressure on H to do, accept
    - E.g. request, suggest, remind, offer, compliment,..
  - Positive: Indicate dislike or indifference to face
    - E.g. criticism, disapproval, contradiction, boasting
  - Threats to H's or S's face; positive/negative
- Given threats, rational agents will minimize
  - Constraints: communicate content, be efficient, maintain H's face

How to be Polite

- On-record: with clear intent

How to be Polite

- On-record: with clear intent
  - Without redress, baldly:
    - Direct: clear and concise as possible
      - Very casual or very urgent

How to be Polite

- On-record: with clear intent
  - Without redress, baldly:
    - Direct: clear and concise as possible
      - Very casual or very urgent
  - With redress, positive:
    - Indicate S want H's wants

How to be Polite

- On-record: with clear intent
  - Without redress, baldly:
    - Direct: clear and concise as possible
      - Very casual or very urgent
  - With redress, positive:
    - Indicate S want H's wants
  - With redress, negative: avoidance-based
    - Conventional indirectness

How to be Polite

- On-record: with clear intent
  - Without redress, baldly:
    - Direct: clear and concise as possible
      - Very casual or very urgent
  - With redress, positive:
    - Indicate S want H's wants
  - With redress, negative: avoidance-based
    - Conventional indirectness
- Off-record: ambiguous intent - hint
- Don't ask....

Indirectness vs Politeness

- Politeness not just maximal indirectness
  - Not just maintain face
  - Balance minimizing inferential effort
  - If too indirect, inferential effort high
    - E.g. hinting viewed as impolite
- Conventionalized indirectness eases interp
  - Maintain face and pragmatic clarity

Generating Speaking Styles

- Stylistic choices
  - Semantic content, syntactic form, acoustic realiz' n
  - Lead listeners to make inferences about character and personality
- Base on:
  - Speech Acts
  - Social Interaction & Linguistic Style

Dialogue Act Modeling

- Small set of basic communicative intents
  - Initiating: Inform, offer, request-info, request-act
  - Response: Accept or reject: offer, request, act

Dialogue Act Modeling

- Small set of basic communicative intents
  - Initiating: Inform, offer, request-info, request-act
  - Response: Accept or reject: offer, request, act
- Distinguish: intention of act from realization
- Abstract representation for utterances
  - Each utterance instantiates plan operator

Dialogue Act Model

- Plan-based speech act decomposition
- Speech Act defined as plan
  - Header: request-act(s,h,a)
  - Precondition: want(s,a), cando(h,a)
  - Effects: want(h,a), know(h,want(s,a))
  - Decompositions
    - Different alternatives specify surface realization
    - Select based on social information

Decomposition & Realization

- Surface-request(s,h,a)
  - “Do a”.

Decomposition & Realization

- Surface-request(s,h,a)
  - "Do a".
- Surface-request(s,h,informif(h,s,cando(h,a)))
  - "Can you do a?"

Decomposition & Realization

- Surface-request(s,h,a)
  - “Do a”.
- Surface-request(s,h,informif(h,s,cando(h,a)))
  - “Can you do a?”
- Surface-request(s,h,~cando(s,a))
  - “I can’t do a”
- Surface-request(s,h,want(s,a))
  - “I want you to do a.”

Representing the Script

- (Manually) Model sequence in story/task
  - Sequence of dialogue acts and physical acts
    - Model world, domain, domain plans
  - Preconditions, effects, decompositions
    - => semantic content
  - Represent as input to linguistic realizer

Modeling Social Interaction

- Based on B&L model of speakers
  - Face: Autonomy and Approval; Rational meaning

Modeling Social Interaction

- Based on B&L model of speakers
  - Face: Autonomy and Approval; Rational meaning
- Based strategy on socially determined vars
  - Social distance, Power, Ranking of Imposition: 1-50
    - Requests, offer, inform: threat to auto; rejects: threat to approval

Modeling Social Interaction

- Based on B&L model of speakers
  - Face: Autonomy and Approval; Rational meaning
- Based strategy on socially determined vars
  - Social distance, Power, Ranking of Imposition: 1-50
    - Requests, offer, inform: threat to auto; rejects: threat to approval
- Try to avoid threats to face
  - Theta= social distance + power + imposition
  - Select strategies based on theta:
    - Direct < Approval-oriented < Autonomy-oriented<off-rec

Modeling Social Interaction

- Based on B&L model of speakers
  - Face: Autonomy and Approval; Rational meaning
- Based strategy on socially determined vars
  - Social distance, Power, Ranking of Imposition: 1-50
    - Requests, offer, inform: threat to auto; rejects: threat to approval
- Try to avoid threats to face
  - Theta= social distance + power + imposition
  - Select strategies based on theta:
    - Direct < Approval-oriented < Autonomy-oriented<off-rec
- Semantic content: plan rep; syntactic form: library

Modeling Social Interaction

- Based on B&L model of speakers
  - Face: Autonomy and Approval; Rational meaning
- Based strategy on socially determined vars
  - Social distance, Power, Ranking of Imposition: 1-50
    - Requests, offer, inform: threat to auto; rejects: threat to approval
- Try to avoid threats to face
  - Theta= social distance + power + imposition
  - Select strategies based on theta:
    - Direct < Approval-oriented < Autonomy-oriented<off-rec
- Semantic content: plan rep; syntactic form: library
- Affect: set acoustic realization
  - Angry, pleasant, disgusted, annoyed, distraught, sad, gruff

Generating Appropriate Style

- Input:
  - Sequence of speech acts
  - Social status: social distance, power, ranking
  - Emotional stance (view as orthogonal)

Generating Appropriate Style

- Input:
  - Sequence of speech acts
  - Social status: social distance, power, ranking
  - Emotional stance (view as orthogonal)
- Example: Speech act= request;
  - Status: D+P+R < 50
    - Direct: Imperative form: “Bring us two drinks”
  -

Generating Appropriate Style

- Input:
  - Sequence of speech acts
  - Social status: social distance, power, ranking
  - Emotional stance (view as orthogonal)
- Example: Speech act= request;
  - Status: D+P+R < 50
    - Direct: Imperative form: "Bring us two drinks"
  - Status: 91<D+P+R<120
    - Autonomy-oriented: query-capability-autonomy
      - "Can you bring us two drinks?" - Conventional indirect

Controlling Affect

- Affect editor (Cahn 1990)
- Input: POS, phrase boundaries, focus
- Acoustic parameters: Vary from neutral
  - 17: pitch, timing, voice and phoneme quality
- Prior evaluation:
  - Naïve listeners reliably assign to affect class

Summary

- Politeness and speaking style
  - Rational agent maintaining face, clarity
  - Indirect requests allow hearer to save face
    - Must be clear enough to interpret
  - Sensitive to power and social relationships
- Generate appropriate style based on
  - Dialogue acts (domain-specific plans)
  - Define social distance and power
  - Emotional state