

# Wrapping Up

Ling575  
Spoken Dialog Systems  
June 5, 2013

# Roadmap

- Overview
  - Distinctive factors in dialog:
    - Human-human
    - Human-computer
  - Dialog components & dialog management
  - Specialized topics:
    - Detailed analysis of:
      - Distinctive factors
      - Techniques and applications
  - Discussion:
    - Trends, techniques, interrelations

# Characteristics of Dialog

- Human-human:
  - Multi-party interaction:
    - Flexible turn-taking, mixed initiative
  - Speech acts:
    - Actions via speech, levels of interpretation
  - Implicature:
    - Grice's maxims
  - Cooperativity & closure:
    - Grounding and levels of display
      - Corrections, repairs, and confirmations

# Characteristics of Dialog

- Human-computer – most deployed systems
  - Multi-party interaction:

# Characteristics of Dialog

- Human-computer – most deployed systems
  - Multi-party interaction:
    - Rigid silence-based turn-taking, system or “mixed” initiative
  - Speech acts:

# Characteristics of Dialog

- Human-computer – most deployed systems
  - Multi-party interaction:
    - Rigid silence-based turn-taking, system or “mixed” initiative
  - Speech acts:
    - Actions via speech: dialog acts, NLU
  - Implicature:

# Characteristics of Dialog

- Human-computer – most deployed systems
  - Multi-party interaction:
    - Rigid silence-based turn-taking, system or “mixed” initiative
  - Speech acts:
    - Actions via speech: dialog acts, NLU
  - Implicature:
    - Um... depends on dialog management, NLU
  - Grounding:

# Characteristics of Dialog

- Human-computer – most deployed systems
  - Multi-party interaction:
    - Rigid silence-based turn-taking, system or “mixed” initiative
  - Speech acts:
    - Actions via speech: dialog acts, NLU
  - Implicature:
    - Um... depends on dialog management, NLU
  - Grounding:
    - Confirmation: implicit/explicit: learned?
    - Corrections, repairs: problematic
  - Why?



# Characteristics of Dialog

- Human-computer – most deployed systems
  - Multi-party interaction:
    - Rigid silence-based turn-taking, system or “mixed” initiative
  - Speech acts:
    - Actions via speech: dialog acts, NLU
  - Implicature:
    - Um... depends on dialog management, NLU
  - Grounding:
    - Confirmation: implicit/explicit: learned?
    - Corrections, repairs: problematic
  - Constrained by complexity, processing, speed, etc

# Dialog System Components

- HMM-based ASR models
- NLU: call-routing, semantic grammars
- Dialog acts and recognition
- Dialog management:
  - Finite-state
  - Frame-based
    - VoiceXML
  - Information state
  - Statistical dialog management
- Lots of examples!

# Topics

- In-depth discussions:
  - Computational approaches to make human-computer interaction more like human-human interaction
    - Many issues raised in characterizing dialog:
      - Multi-party

# Topics

- In-depth discussions:
  - Computational approaches to make human-computer interaction more like human-human interaction
    - Many issues raised in characterizing dialog:
      - Multi-party: multi-party interaction, turn-taking, initiative
      - Grounding

# Topics

- In-depth discussions:
  - Computational approaches to make human-computer interaction more like human-human interaction
    - Many issues raised in characterizing dialog:
      - Multi-party: multi-party interaction, turn-taking, initiative
      - Grounding: Miscommunication & repair, incremental processing
      - Interpretation:

# Topics

- In-depth discussions:
  - Computational approaches to make human-computer interaction more like human-human interaction
    - Many issues raised in characterizing dialog:
      - Multi-party: multi-party interaction, turn-taking, initiative
      - Grounding: Miscommunication & repair, incremental processing
      - Interpretation: Reference, affect, subjectivity, personification, information structure, prosody
      - Multi-modality
  - Applications and issues:
    - Tutoring, machine translation, information-seeking
    - Non-native speech

# Interconnections

Apps: MT

Tutoring

Non-  
native

Turn-  
taking

Affect

Sentiment

Info.  
Struct

Reference

Increment

Prosody

Initiative

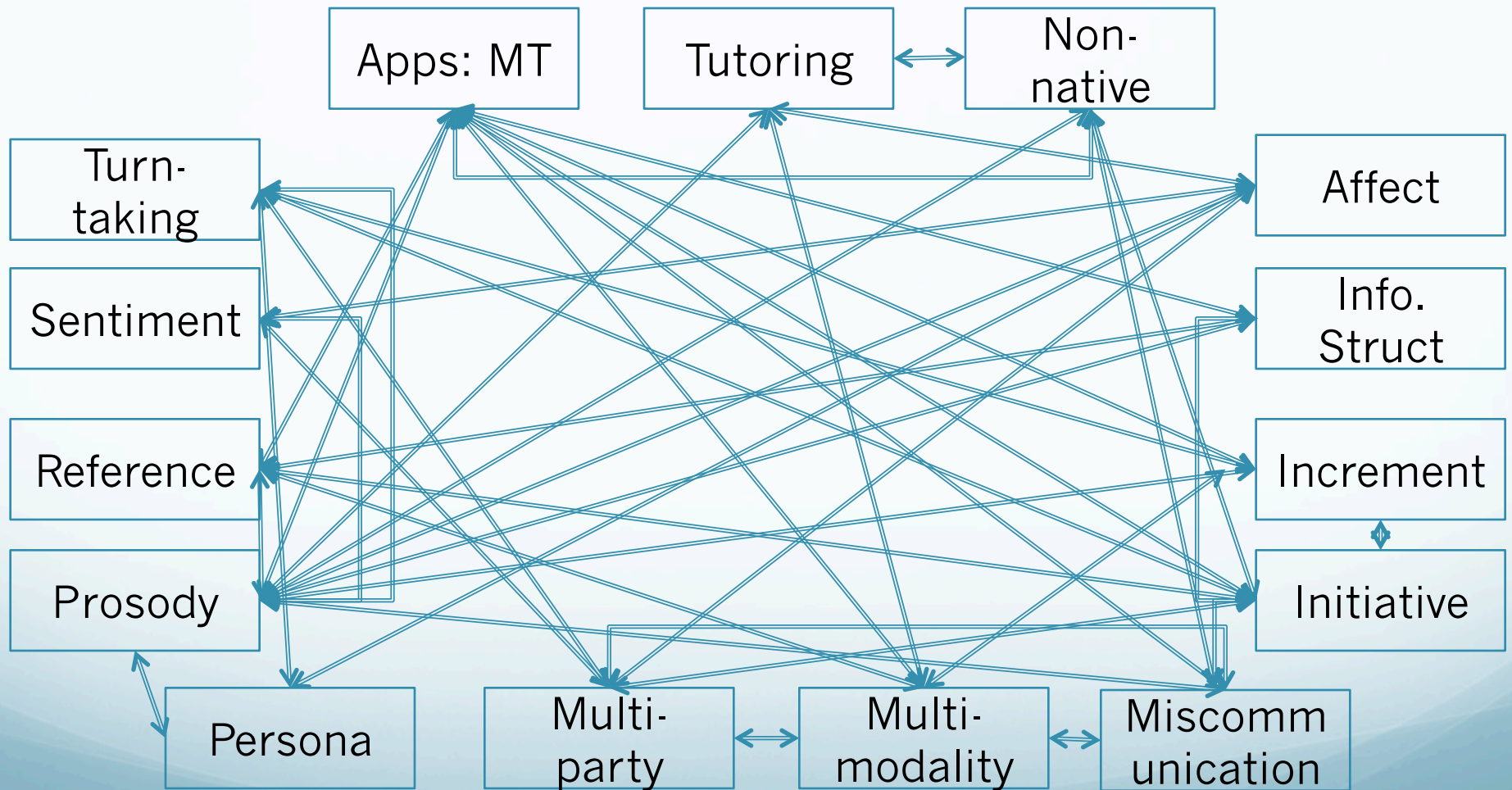
Persona

Multi-  
party

Multi-  
modality

Miscomm  
unication

# Interconnections





# Techniques & Sources of Information

- Range of techniques:

# Techniques & Sources of Information

- Range of techniques:
  - Deep processing, shallow processing, manual rules
- Machine learning:

# Techniques & Sources of Information

- Range of techniques:
  - Deep processing, shallow processing, manual rules
- Machine learning:
  - Anything from decision trees to POMDPs
- Information sources:

# Techniques & Sources of Information

- Range of techniques:
  - Deep processing, shallow processing, manual rules
- Machine learning:
  - Anything from decision trees to POMDPs
- Information sources:
  - Acoustic, lexical, prosodic, timing, syntactic, semantic, pragmatic, etc
    - Multimodal: gaze, gesture, etc
  - Integration

# Techniques & Sources of Information

- Range of techniques:
  - Deep processing, shallow processing, manual rules
- Machine learning:
  - Anything from decision trees to POMDPs
- Information sources:
  - Acoustic, lexical, prosodic, timing, syntactic, semantic, pragmatic, etc
    - Multimodal: gaze, gesture, etc
  - Integration: Complex and varied
    - Huge feature vectors, tandem models, blackboards, learned
- Substantial strides, but huge remaining challenges

# Questions?

- Favorite topic?
- Most surprising result?
- Most obvious result?
- Most surprising gap?