The Dialog State Tracking Challenge and Bayesian approach

Quan Nguyen

5qnguyen@informatik.uni-hamburg.de

22.06.2016

UNIVERSITÄT HAMBURG, DEPARTMENT OF INFORMATICS
NATURAL LANGUAGE SYSTEMS GROUP

Outline

- Motivation
- Dialog State Tracking Problem
- Bayesian Approach
- Conclusion
- References

Motivation

- From agent's perspective:

- **What do we have so far**?

(common knowledge)

- **What do the user need**?

(requested slots)

- **What do I need from the user**?

(informable slots)

- **What should I say**?

(conversation acts)

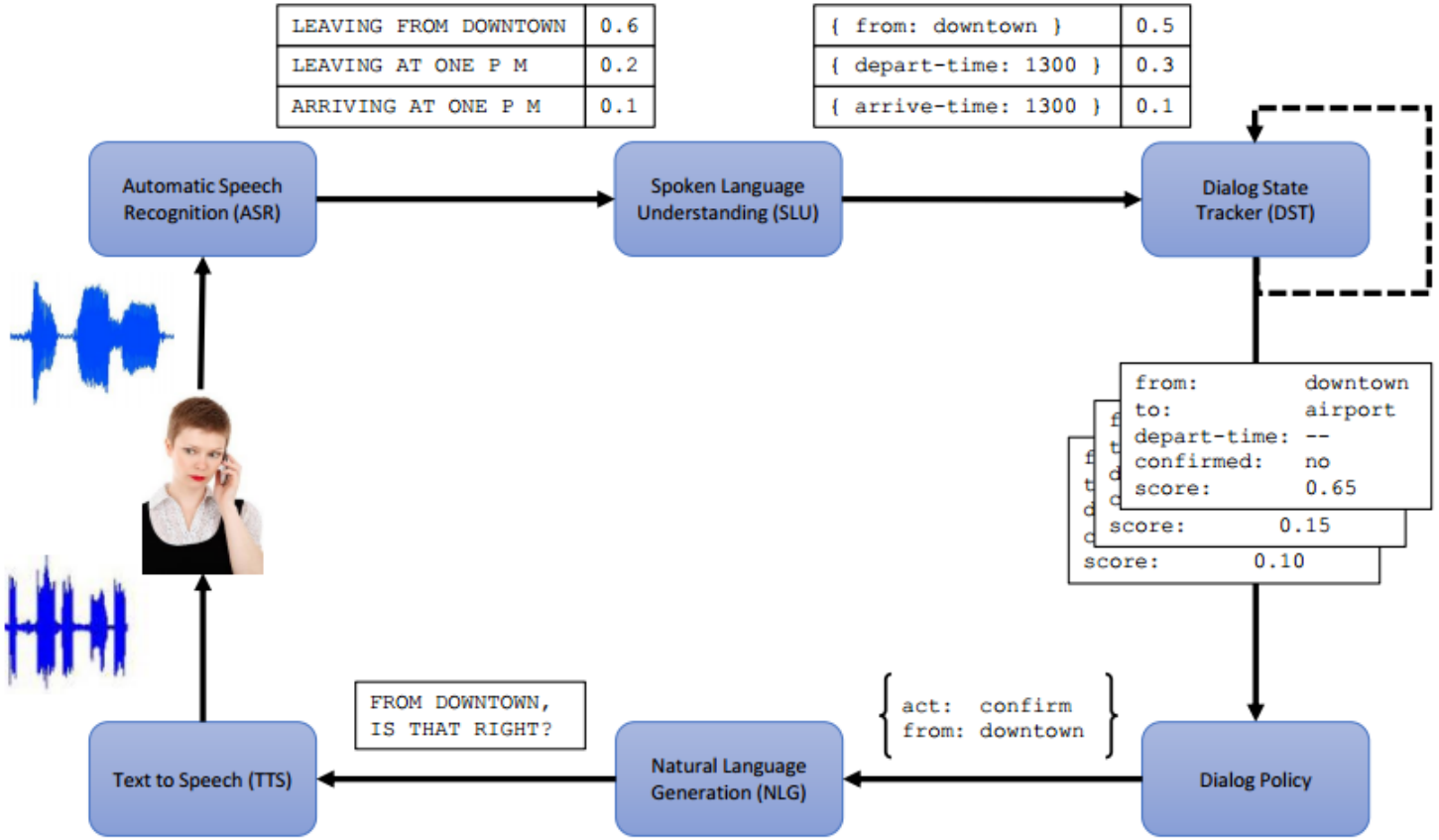
Motivation (2)

- Measuring the accuracy of agent's understanding
  - No common evaluation metrics
  - Different domain
  - Different techniques
  - Different (learning) data

- One state tracker for multiple strategies, multiple domain?

Dialog State Tracking Problem

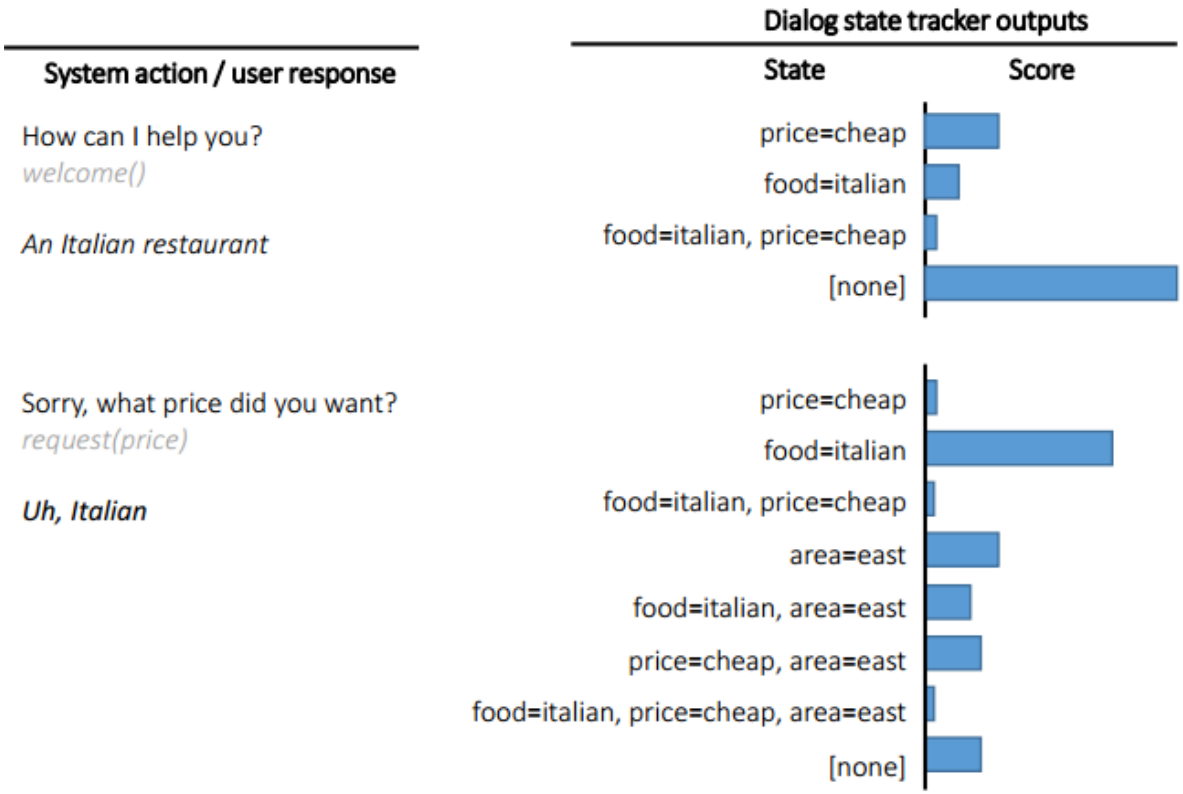
- State of Dialog



Wiliams et.al, 2007

Dialog State Tracking Problem (2)

- Tracker's output



Wiliams et.al, 2007

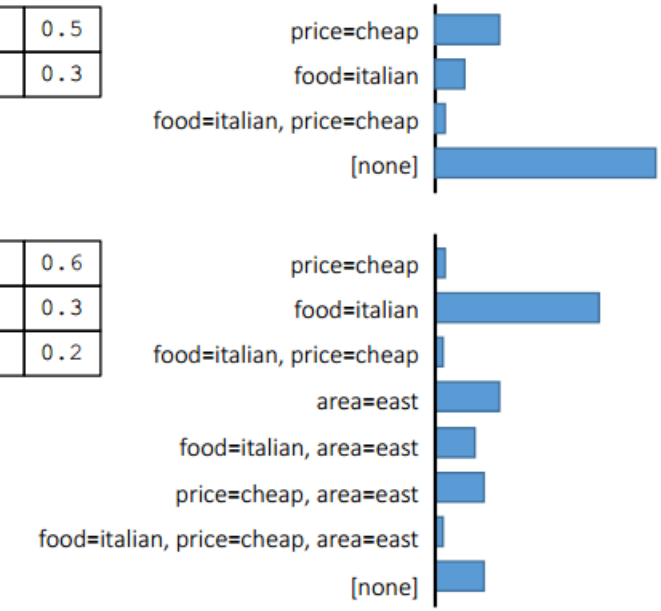
Dialog State Tracking Problem (3)

- Tracker's input (end-to-end system)

Dialog state tracker inputs

System action / user response	ASR output		SLU output	
How can I help you? *welcome()*	CHEAP RESTAURANT	0.6	inform(price=cheap)	0.5
	RESTAURANT	0.2	inform(food=italian)	0.3
An Italian restaurant	ITALIAN	0.1		
Sorry, what price did you want? *request(price)*	EAST AREA	0.5	inform(area=east)	0.6
	ITALIAN	0.3	inform(food=italian)	0.3
Uh, Italian	YEAH	0.1	affirm()	0.2

Dialog state tracker outputs



Wiliams et. al, 2007

The DSTC challenge

- Different trackers – similar train and test data
  - Mixed strategies
  - Shared ontology?
- Baseline
  - Admissible heuristic
  - Check for improvement
  - Incorporate past solutions
- Evaluation metrics
  - Schedule – when to measure.
  - Accuracy – number of correct turns
  - L2 – vector distance with one-hot encoding
  - AvgP – average score

Bayesian approach

- How to build statistical frameworks?
- How to estimate Bayesian networks' parameters (tracker's experience)?
- Tractability?

Bayesian approach (2)

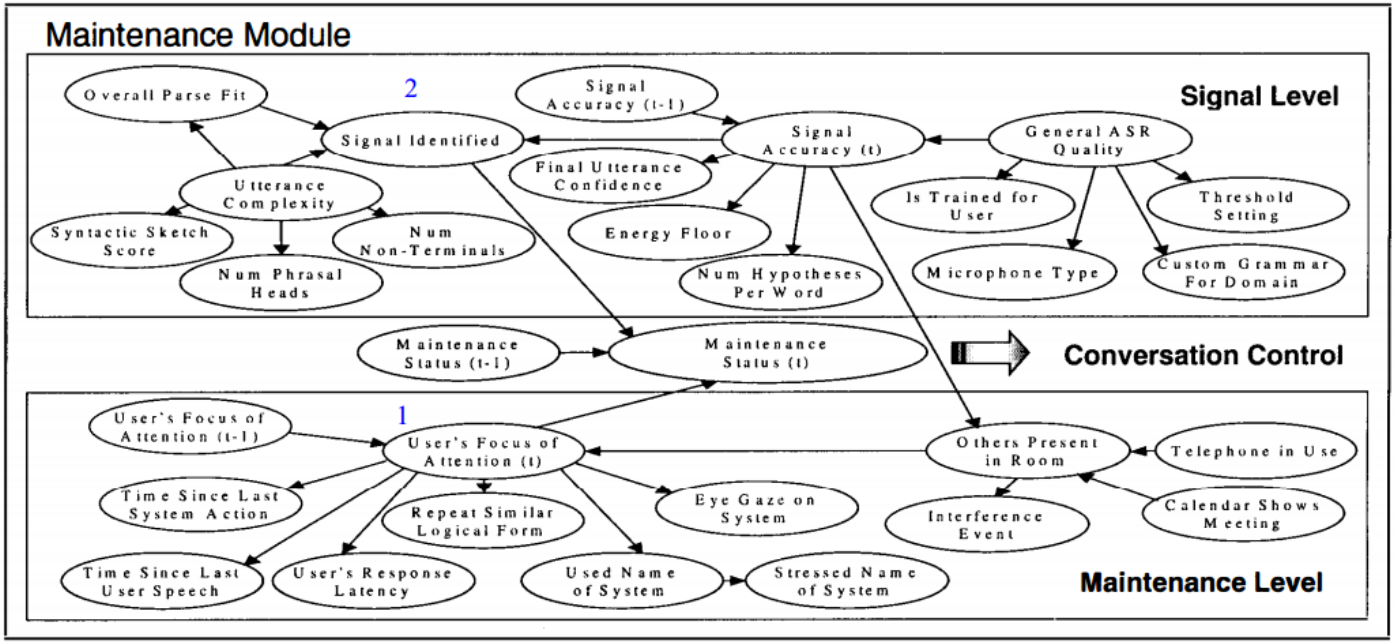
- Basic statistical framework (Williams et. al, 2007)

$$b'(s') = k \cdot p(o'|s', a) \sum_{s \in S} p(s'|a, s) b(s).$$

- Where
  - s and s' are dialog state
  - o' is agent's observations (ASR and SLU output)
  - a is agent's action (speech act)
  - k is normalization constant
  - b(s) and b'(s'): state distribution before and after action a

Bayesian approach (2)

- Tracker's experiences:
  - Hand-crafted nodes and edges (Paek et. al, 2000)



  - Reinforcement learning: Optimal Policy (Williams et. al, 2007)
    - Value Iteration

Bayesian approach (3)

- Tracker's experiences (cont):
  - States with continuous values and/or search methods
    - Particle filters: Continuous -> Discrete (Williams 2007)
    - Beam search: M-best most likely states
- Tractability
  - Training phase + Value iteration approach: quadratic in number of states and number of actions (per iteration)
  - End-to-end system: quadratic in number of states

Bayesian approach (4)

- Hand-crafted nodes and rewards
  - Infeasible in complex domain
- Only works under certain assumptions
  - Independence
  - Short-term dependence (dialog history)

Conclusion

- Trackers are crucial in dialog systems
  - What the agents know
  - What the agents should do
- The DSTC challenge
  - Common testing framework
- Bayesian approach
  - Statistical framework -> room for tracking methods
  - Works under certain assumptions (Markov, short-term dependence, small state space)

Adaptation to

- DSTC2: user goal changes
- DSTC3: unseen concepts in test data

The End

Thank you for your attention.

Any question?

Literature:

- Williams, Raux and Henderson. The DSTC Series: A review*. Dialog and Discourse*, 2007
- Williams and Young. Partially observable Markov decision processes for spoken dialog system. *Computer Speech and Language*, 21(2):393-422, 2007a
- Paek and Horvitz. Conversation as Action Under Uncertainty. *UNCERTAINTY IN ARTIFICIAL INTELLIGENCE PROCEEDINGS*, 2000
- Williams. Using Particle filters to track dialogue state. In *Proc IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), Kyoto, Japan,* 2007