Towards relational POMDPs for adaptive dialogue management

Pierre Lison
German Research Centre for Artificial Intelligence (DFKI GmbH)
http://talkingrobots.dfki.de

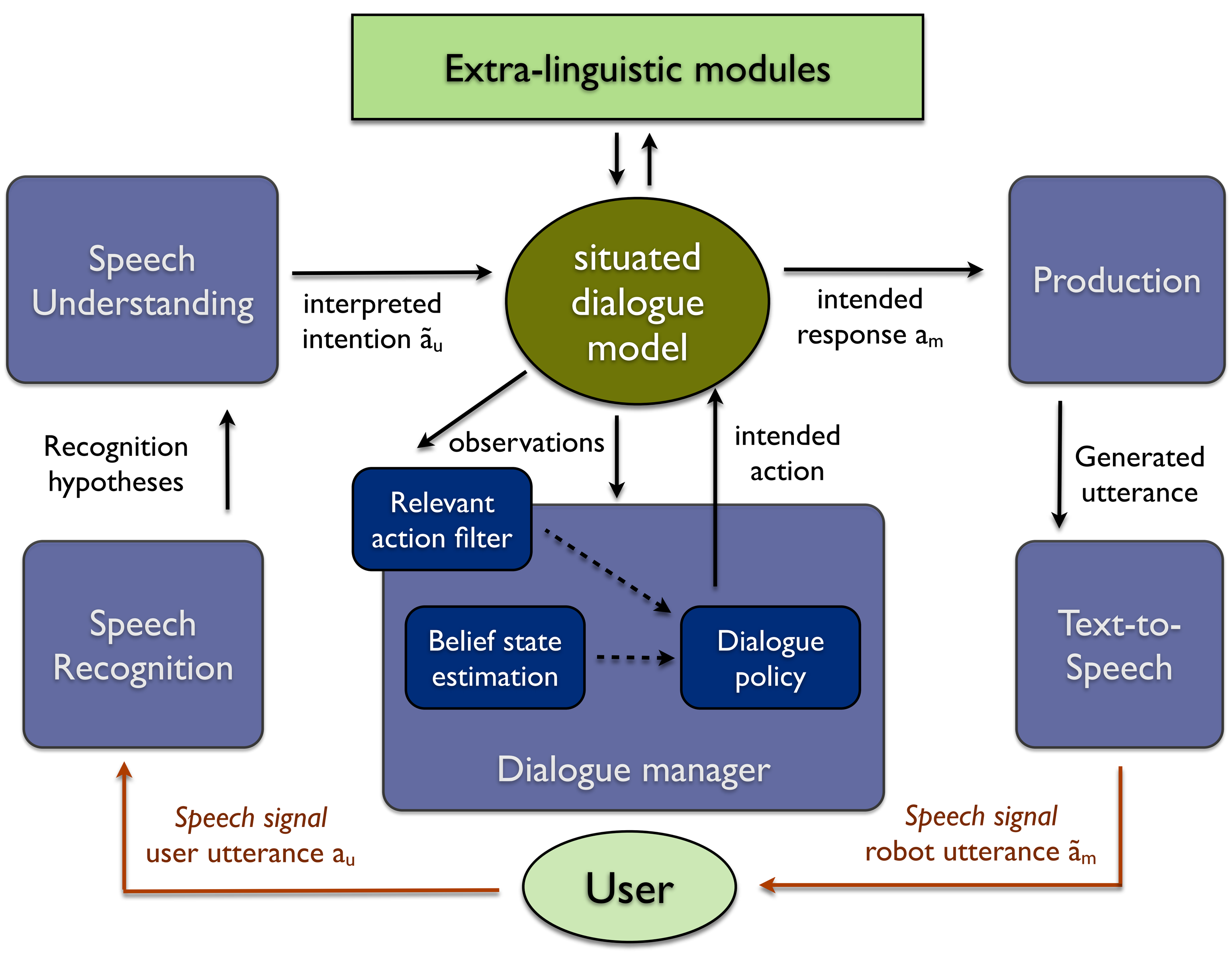
Introduction

- We are interested in developing *spoken dialogue systems* for *rich, open-ended interactions*
- For instance: human-robot interaction (HRI) in indoor or outdoor environments using spoken dialogue
- Two issues to solve:
  - **high levels of uncertainty** (speech recognition errors, limited grammar coverage, linguistic/pragmatic ambiguities)
  - **structural complexity** (dialogue history, task model, external context viewed as rich relational structures)
- In addition, the dialogue systems should also be **adaptive** to a variety of internal and external factors

→ How to develop a dialogue manager which takes these requirements into account?

- The issues of uncertainty and structural complexity have traditionally been addressed separately in the literature:
  - *Uncertainties* usually dealt with probabilistic models combined with decision-theoretic planning
  - *Structural complexity* addressed with logic-based approaches to pragmatic reasoning
- Our long-term goal is to develop a **hybrid approach** to dialogue management which combines these insights
- We present here preliminary work on this topic
- The core idea is to combine a *POMDP* framework with *pragmatic knowledge* over relevant dialogue moves
- This pragmatic knowledge is encoded in a first-order probabilistic language, **Markov Logic Networks (MLNs)**

Approach

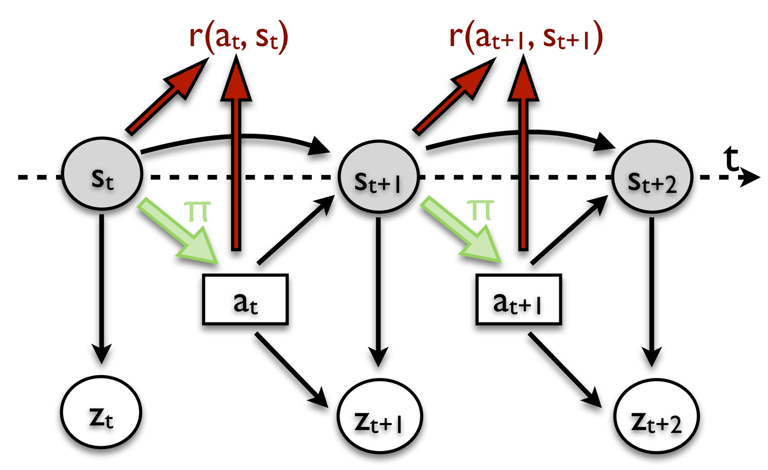


Architecture schema of the spoken dialogue system

- Dialogue manager based on POMDPs
- Powerful framework for control problems with partial observability, uncertain action effects, and multiple conflicting objectives.

\+ *optimisation* of dialogue policies

- Dialogue state is not directly accessible but must be inferred from observation



Simplified schema of the POMDP model

- Classical POMDP approaches operate directly on the full action space
- Significant planning time is spend on actions which should be discarded as irrelevant. Ex: "Robot, take the left mug!" replied with "Is the box green?"
- Instead, we formalise action selection as a *two-step* process:
  1. Using pragmatic knowledge encoded in MLNs, we extract locally relevant dialogue moves from the situated dialogue model
  2. The POMDP planner or precompiled dialogue policy then selects the optimal (highest-reward) action on the basis of this reduced action space
- Example of MLN rule (with infinite weight):

$$LastUserMove(s, a_u) \wedge PolarQuestion(a_u) \wedge YesNoAnswer(a_m) \rightarrow RelevantMove(a_m, s).$$

Discussion

- The current dialogue state is represented by both a general, open-world dialogue model (including situated, multi-agent beliefs, events and intentions) and a smaller, closed-world POMDP belief state
- The first-order expressivity of Markov Logic Networks allows us to leverage the *rich relational structure* of the problem
- Distinction between relevance and utility: pragmatic knowledge is used to determine the *relevance* of a given action, POMDP planning its *utility*

- We are currently implementing these ideas as part of a spoken dialogue system for human-robot (social) interaction
- Future work will focus on:
  1. The integration of these action relevance filters with offline/online POMDP planners and reinforcement learning techniques
  2. Using statistical relational learning to set the MLN weights
  3. And the evaluation of our approach on empirical data