International Workshop on Spoken Dialogue Systems - 2012, Paris (France)

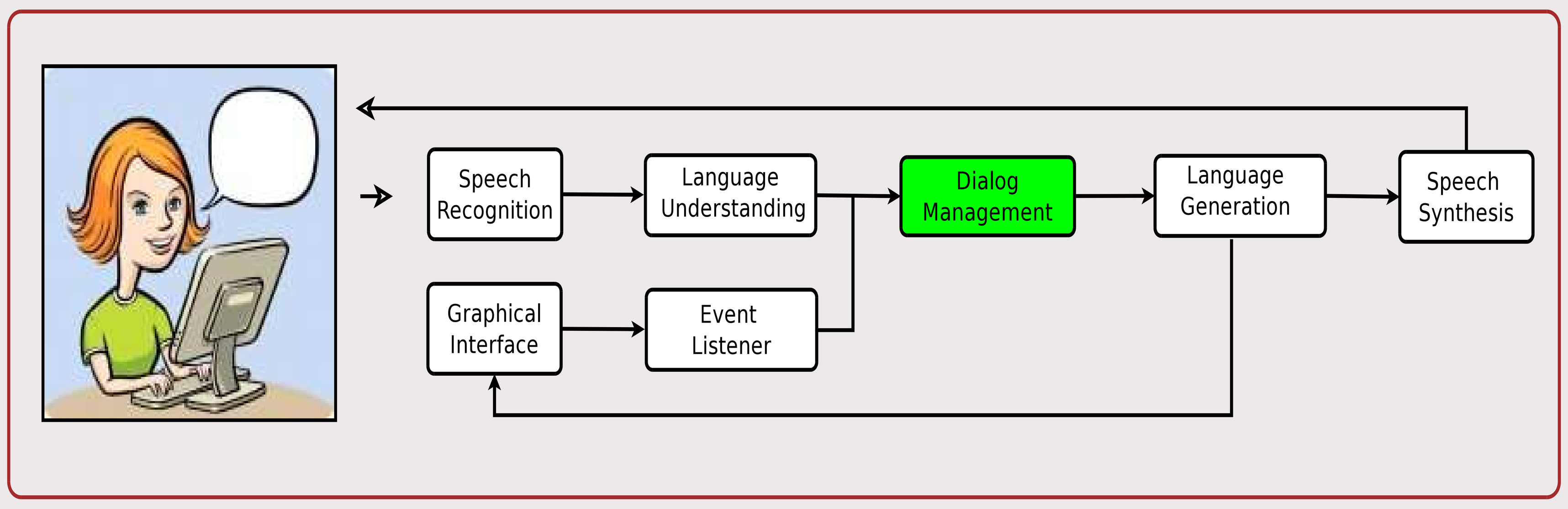
Co-adaptation in Spoken Dialogue Systems

Senthilkumar Chandramohan [1,3], Matthieu Geist [1], Fabrice Lefèvre [3], Olivier Pietquin [1,2]
[1] Supelec, MaLIS - IMS Research Group, Metz, France
[2] UMI 2958 (CNRS - GeorgiaTech), France
[3] Université d'Avignon et des Pays de Vaucluse, LIA-CERI, France
[1]firstname.lastname@supelec.fr [3]firstname.lastname@univ-avignon.fr

Spoken Dialogue Systems (SDS)

- Natural language interfaces for human-computer interaction
- Speech is a frequently used mode of interaction
- Involves: speech recognition/synthesis, language understanding/generation
- Often designed to perform specific tasks
- Examples for goal-oriented dialogue systems:
  - Flight ticket booking
  - Town-information
  - Language tutoring

Architecture of spoken dialogue systems

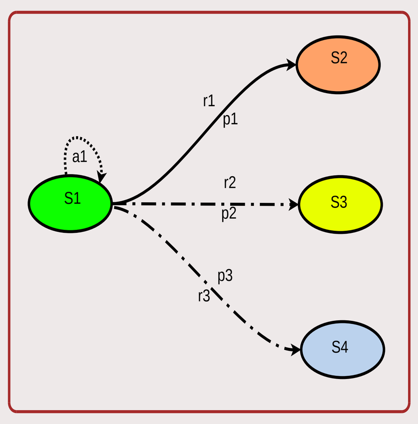


Spoken dialogue management

- Navigates the dialogue system
- Essentially a decision making problem
- What to ask some information?
- When to say some thing to the user?
- More than one possible action
- Choose best action given dialogue context
- Sequential decision making problem

Markov Decision Processes (MDP)

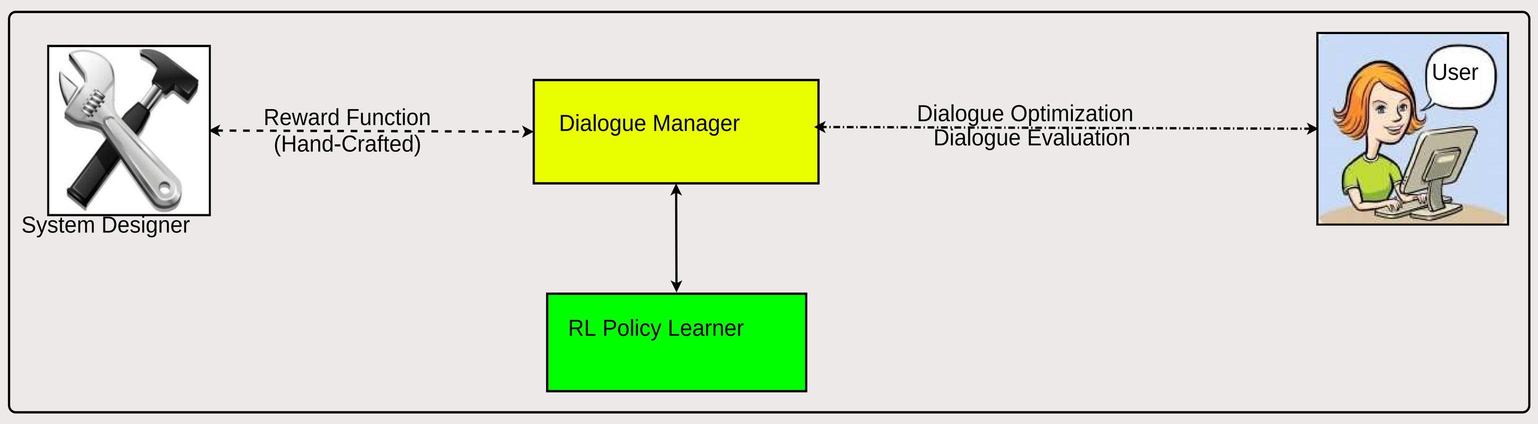
- Solve sequential decision making problems
- Markov Decision Process: $\{S,A,R,P,\gamma\}$ [2]
- Rewarded state transitions {s,a,r,s'}
- Solution of an MDP is an optimal policy
- Schemes to solve MDPs:
  - Dynamic programming [1] (model based)
  - Reinforcement learning [7] (model free)



Spoken dialogue optimization (MDP-SDS)

- Dialogue management is a sequential decision making problem
- Dialogue management problem is cast as a MDP [2]
- Transition probabilities i.e., model of the system is not available
- Reinforcement learning is used for dialogue optimization [5]
- System designer specifies the reward function to be maximized

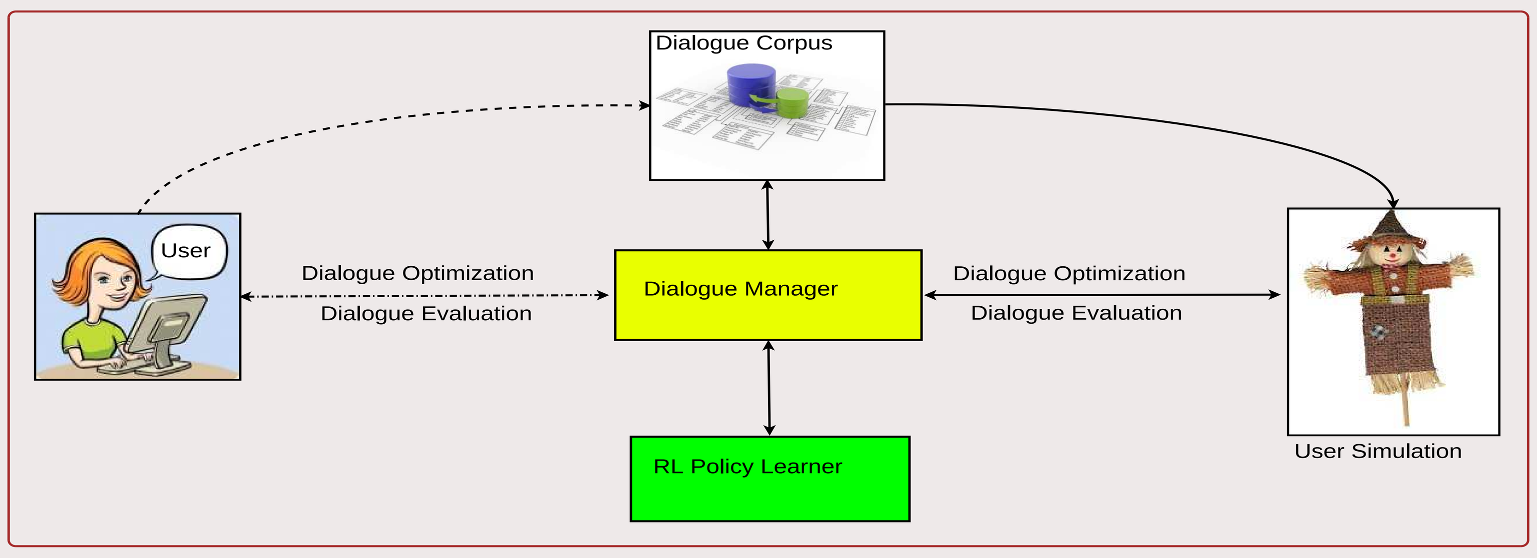
Dialogue optimization using reinforcement learning



User simulation in spoken dialogue systems

- However, RL needs large amount of dialogue corpora for policy optimization
- Dialogue corpora generation is expensive and time taking process
- User simulators [6] are built from corpora
- Simulators aims at generating synthetic dialogue corpus
- Simulators are used to estimate/evaluate dialogue policies
- Is it optimal to adapt dialogue manager to a fixed user or corpus?
- In real world users not only adapt but also change goal/behavior
- It is important to build dynamic and adaptive user simulators

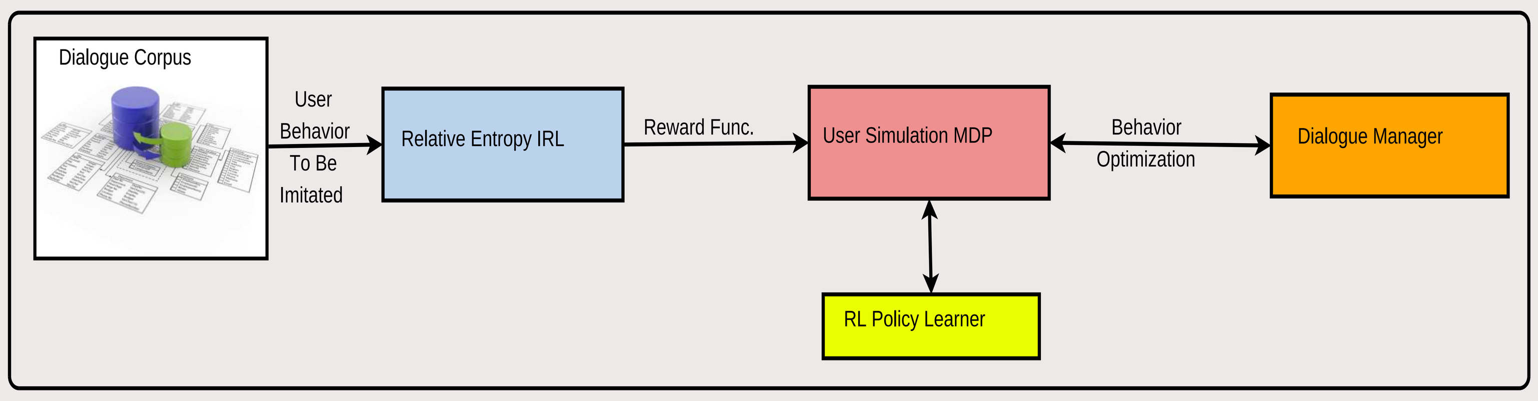
Role of user simulation in dialogue optimization



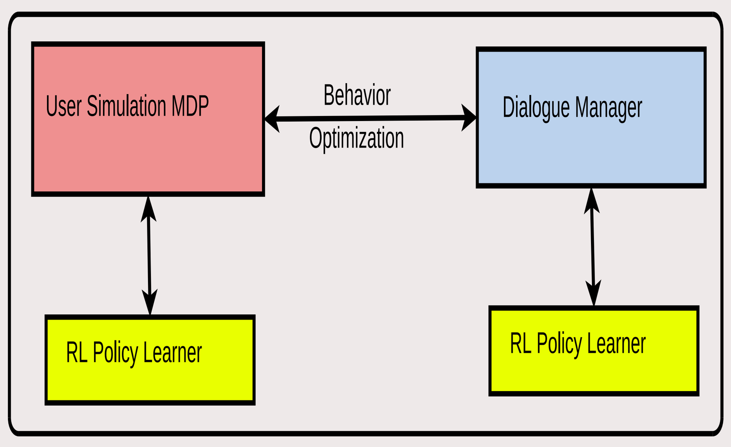
Modelling user simulation as a Markov Decision Process

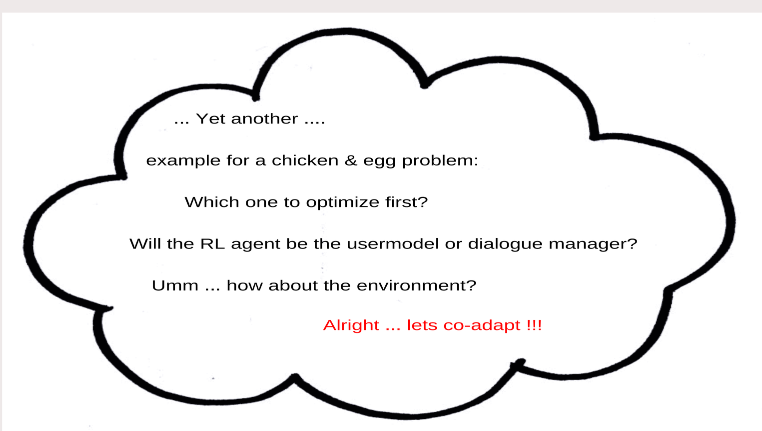
- Dialogue system end users tend to behave in goal oriented manner
- User behavior can be perceived as sequence of decisions
- Sequential decision making users can be modelled as an MDP
- Reward function required for RL optimization must be learned from corpus
- Relative Entropy IRL was employed to learn the reward function [3]
- User model is casted as an MDP and behavior is imitated using IRL and RL

Revisiting spoken dialogue optimization problem)



Consequence of casting both dialogue manager and user as MDPs

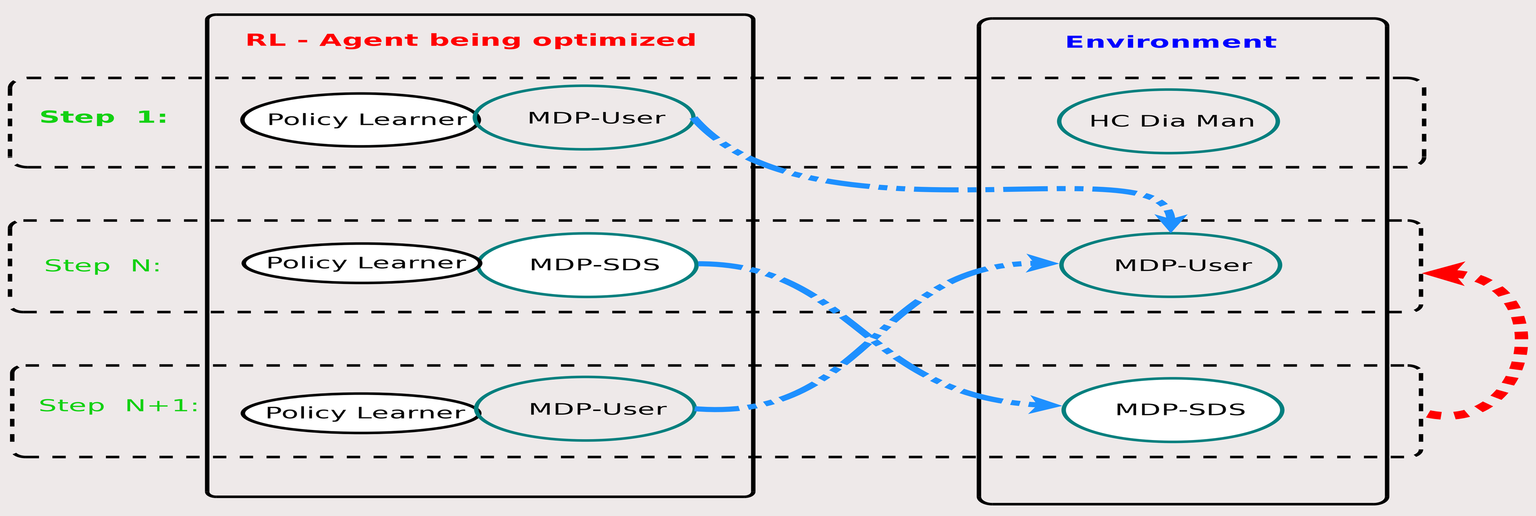




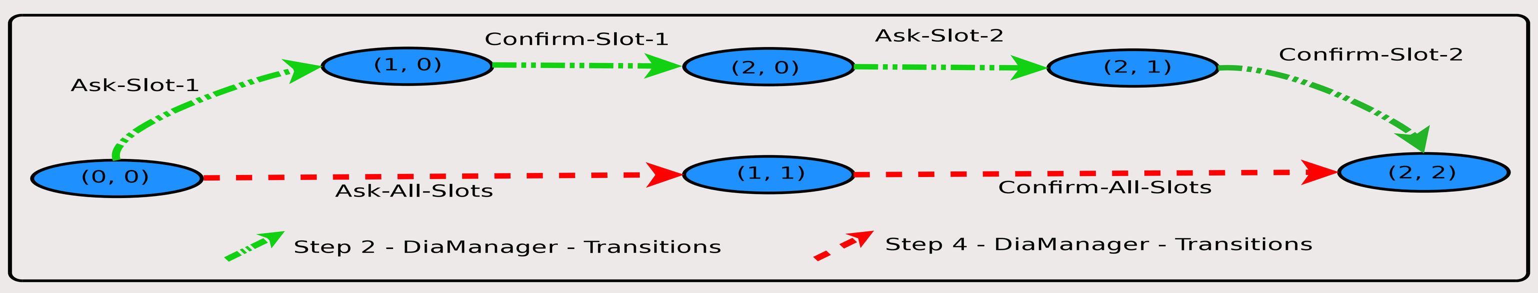
Revisiting spoken dialogue optimization problem)

- Naturalness observed in human-human communication is result of
  - (1) Dialogue initiation stage (focal point of current dialogue research)
  - (2) Dialogue evolution or co-alignment stage
- Online dialogue optimization may seem to facilitate co-alignment, however
  - (1) Users can better adapt to SDS than vice-versa
  - (2) Introduces some degree of over-confidence during adaptation [4]

Co-adaptation framework in spoken dialogue systems



Experiment: Co-adaptation in restaurant information SDS



- Dialogue and user policies evolve and improve over time
- Similar to human-human dialogue, co-adaptation is subjective to ASR errors
- Co-adaptation is a step towards building self evolving dialogue systems

References

- R. Bellman.
  Dynamic Programming.
  Dover Publications, sixth edition, 1957.
- R. Bellman.
  A markovian decision process.
  Journal of Mathematics and Mechanics, vol. 6, pp. 679–684, 1957.
- A. Boularias, J. Kober, and J. Peters.
  Relative entropy inverse reinforcement learning.
  Journal of Machine Learning Research - Proceedings Track, 15:182–189, 2011.
- M. Gasic, F. Jurcicek, B. Thomson, K. Yu, and S. Young.
  On-line policy optimisation of spoken dialogue systems via live interaction with human subjects".
  In *Proc. of ASRU 2011*, Hawaii (USA), 2011.
- E. Levin and R. Pieraccini.
  Using markov decision process for learning dialogue strategies.
  In *Proc. ICASSP'98, Seattle (USA)*, 1998.
- J. Schatzmann, K. Weilhammer, M. Stuttle, and S. Young.
  A survey of statistical user simulation techniques for reinforcement-learning of dialogue management strategies.
  Knowledge Engineering Review, vol. 21(2), pp. 97–126, 2006.
- R. Sutton and A. Barto.
  Reinforcement Learning: An Introduction.
  The MIT Press, 3rd edition, March 1998.