End-to-End Learning of Task-Oriented Dialogs

Bing Liu

PhD Candidate, Carnegie Mellon University

Advisor: Prof. Ian Lane

Email: liubing@cmu.edu

Web: http://bingliu.me

Dialog Systems

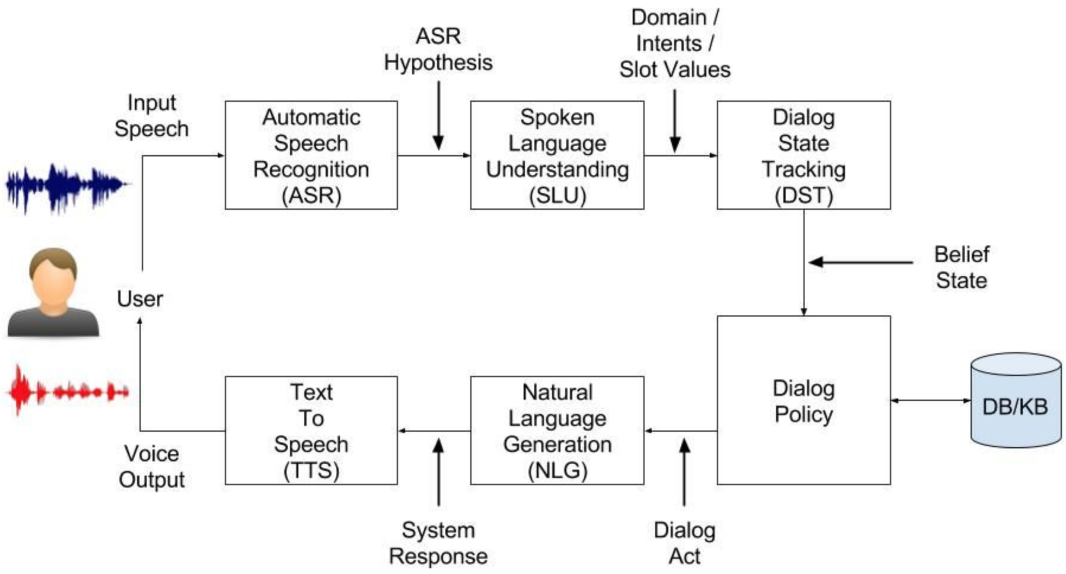
- ❖ Dialog system, or conversational agent, is a class of intelligent system that interacts with users in natural language form via speech or text.
  - ➢ Personal assistants (e.g. Amazon Alexa, Google Assistant, Apple Siri, etc.)
  - ➢ Voice command in vehicle and smart home
  - ➢ Customer service
  - ➢ Chat for entertainment
  - ➢ Psychotherapy

Dialog Systems

- ❖ Task-Oriented Dialog System
  - ➢ Chat to complete tasks
  - ➢ Usually with a user goal in a specific task domain, e.g. movie search, flight booking

- ❖ Chit-Chat Bot
  - ➢ Designed for casual chat, entertainment, and companionship
  - ➢ Open domain, usually does not focus on a particular task

Current Dialog System Architecture



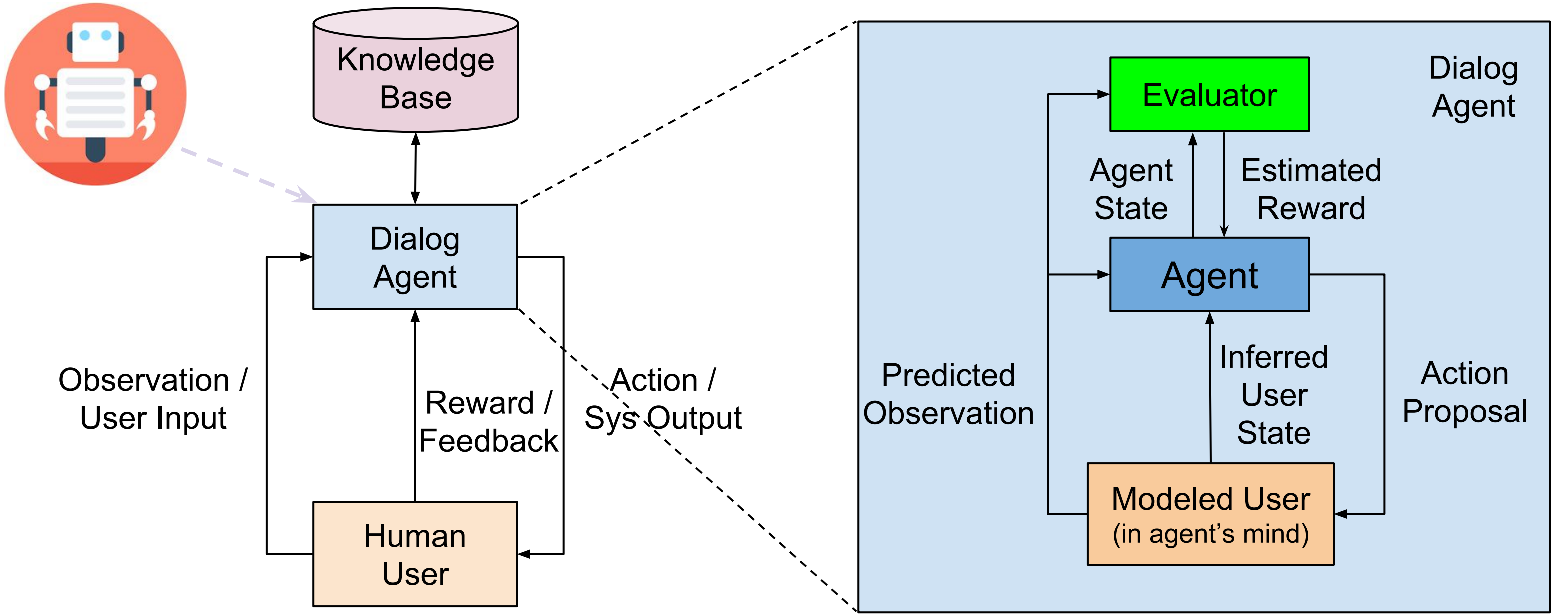
Limitations:

- Highly handcrafted / Many complex rules
- Credit assignment challenge
- Misaligned optimization targets

Thesis Statement

Can we learn *end-to-end* task-oriented dialog system effectively through *interaction* with users?

Overview of Proposed Learning Framework



Challenges / Research Questions

- **Learn End-to-End Dialog from Corpora**
  - How to understand user's natural language request? (*Interspeech 2016, SIGDIAL 2016, NIPS Workshop 2017*)
  - How to reason over long-term dialog context and model task-oriented dialog end-to-end? (*Interspeech 2017*)

- **Learn from Real Interactions**
  - How to learn interactively from human teaching and corrections? (*NAACL 2018a*)
  - How to learn interactively from human feedback? (*NIPS Workshop 2017, AAAI 2018, NAACL 2018b*)

- **Learn from Simulated Interactions**
  - How to model user dynamics, and how to train the user model iteratively with the dialog agent? *(in progress)*
  - How to learn by integrating real and simulated experiences? *(in progress)*

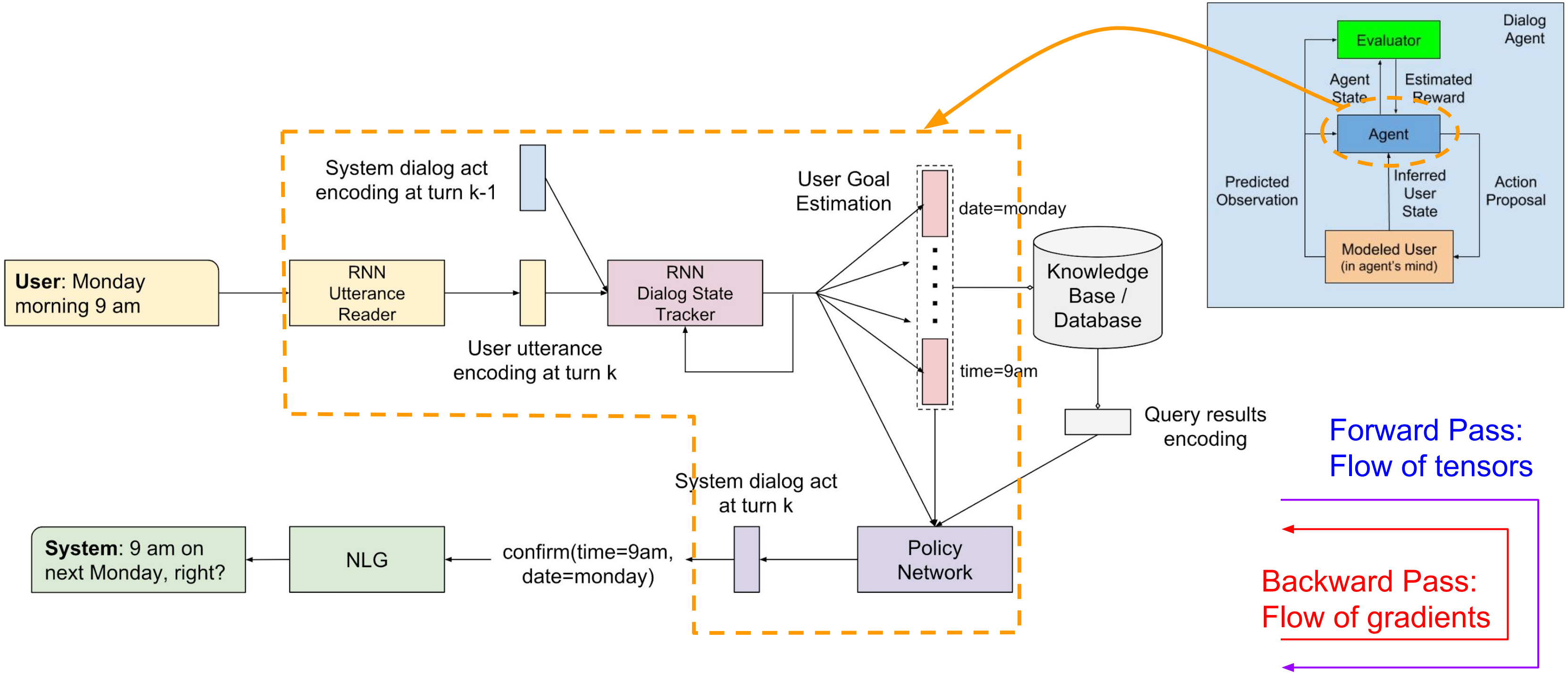
Proposed Dialog Learning Methods

- **Learning End-to-End Task-Oriented Dialog via User Interaction**
- Learning from Simulated Experiences

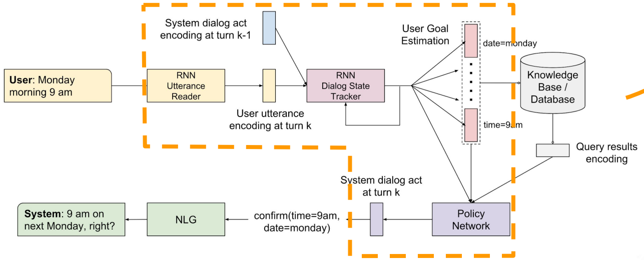
Task-Oriented Dialog Modeling

- Goal: Design a system that takes actions in order to complete a task with users and maximize user satisfaction.
- Sequential Decision Making problem:
  - State: Dialog context ← Modeled with hierarchical LSTM
  - Action: System prompt or response ← Policy network
  - Reward: Task completion ← User feedback and/or Adversarial Reward
- Hybrid Learning Strategy:
  - Supervised pre-training on dialog corpora
  - Interactive learning with human-in-the-loop

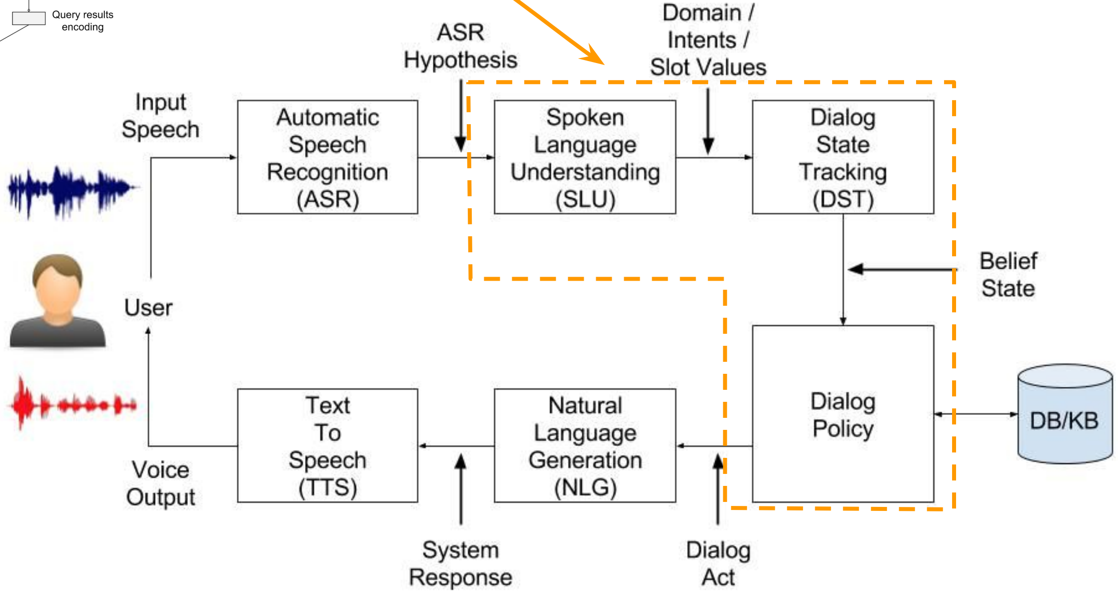
End-to-End Task-Oriented Dialog Modeling



End-to-End Task-Oriented Dialog Modeling

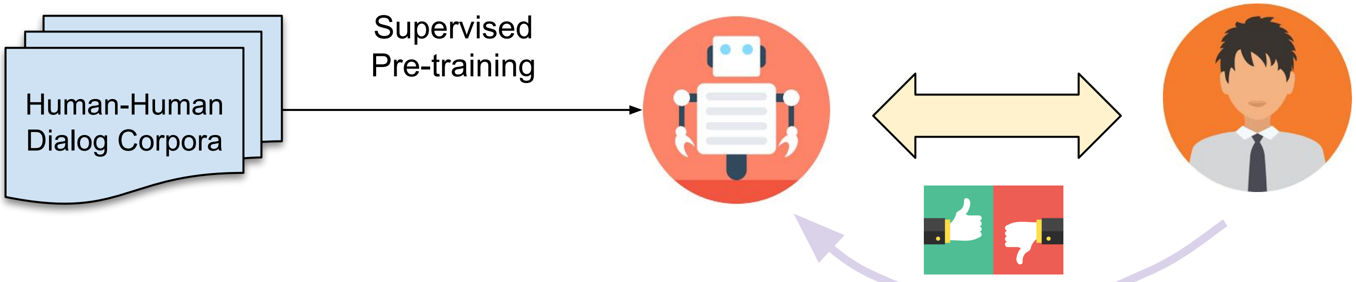


End-to-End Modeling of
SLU, DST, and Dialog Policy



Dialog Learning with Human-in-the-loop

❖ Supervised pre-training + Interactive learning with user feedback



Supervised training objective: linear interpolation of cross-entropy losses:

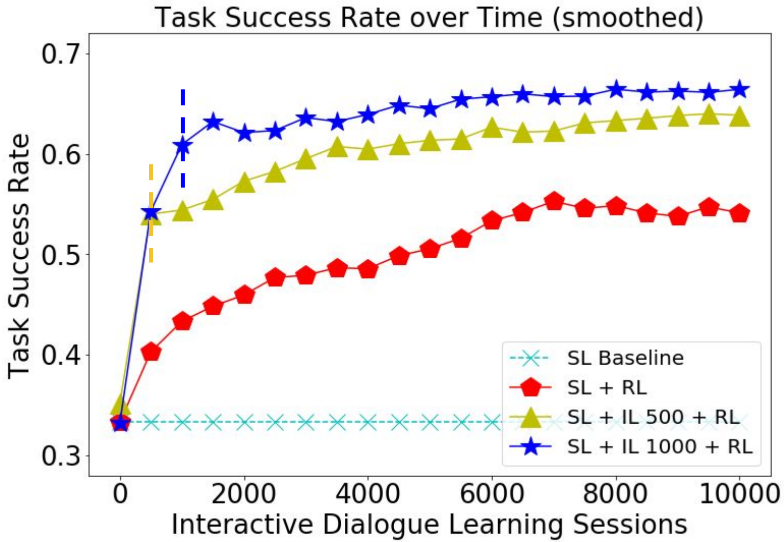
- ❑ User goal estimation, and
- ❑ System action prediction

Interactive learning with RL:

- ❑ Binary feedback as dialog reward
- ❑ E2E optimization with REINFORCE

System Evaluation

❖ Interactive evaluation in a movie booking task domain[1]



SL: Supervised learning model
IL: Imitation learning with human teaching
RL: Reinforcement learning with feedback

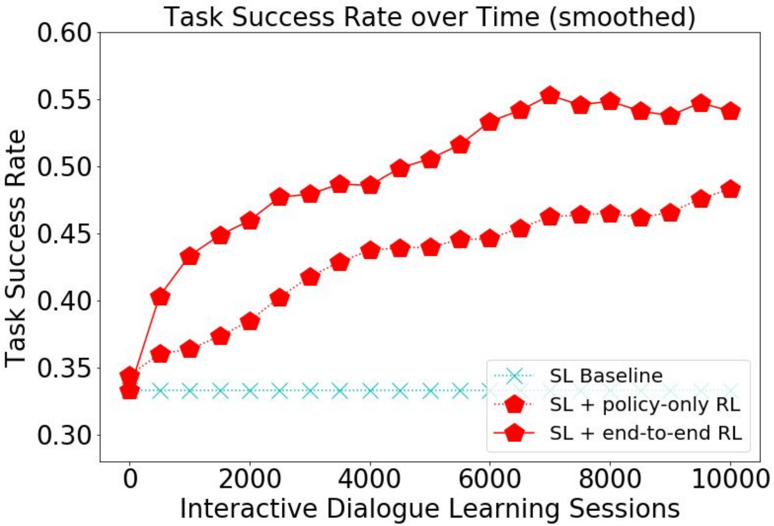
Table: Human evaluation results. Mean and standard deviation of crowd worker scores (1-5)

Model	Score
SL	3.987 ± 0.086
SL + IL 1000	4.378 ± 0.082
SL + IL 1000 + RL	4.603 ± 0.067

[1] Bing Liu, Gokhan Tur, Dilek Hakkani-Tur, Pararth Shah, and Larry Heck, "*Dialogue Learning with Human Teaching and Feedback in End-To-End Trainable Task-Oriented Dialogue Systems*", in NAACL 2018.

End-to-End Model Optimization with RL

❖ Experimental results on task success rate

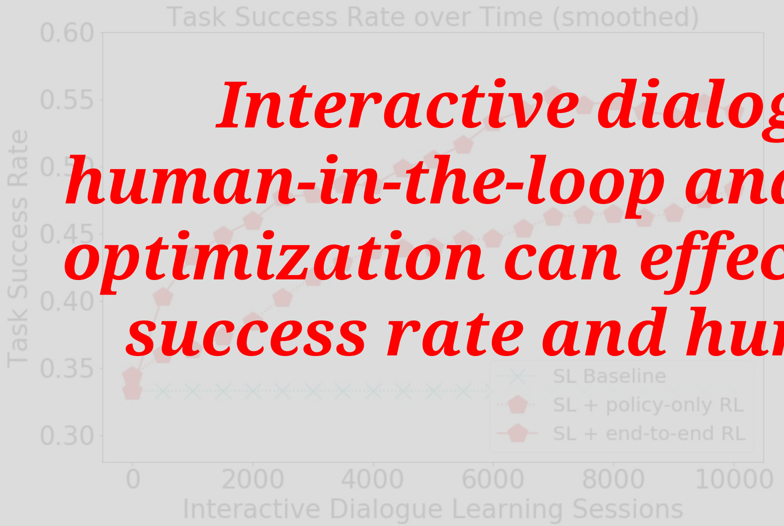


SL: Supervised learning model
RL: Reinforcement learning with feedback

Bing Liu, Gokhan Tur, Dilek Hakkani-Tur, Pararth Shah, and Larry Heck, "Dialogue Learning with Human Teaching and Feedback in End-To-End Trainable Task-Oriented Dialogue Systems", in NAACL 2018.

End-to-End Model Optimization with RL

❖ Experimental results on task success rate



SL: Supervised learning model

RL: Reinforcement learning with feedback

Interactive dialog learning with human-in-the-loop and end-to-end system optimization can effectively improve task success rate and human user ratings.

Bing Liu, Gokhan Tur, Dilek Hakkani-Tur, Pararth Shah, and Larry Heck, "Dialogue Learning with Human Teaching and Feedback in End-To-End Trainable Task-Oriented Dialogue Systems", in NAACL 2018.

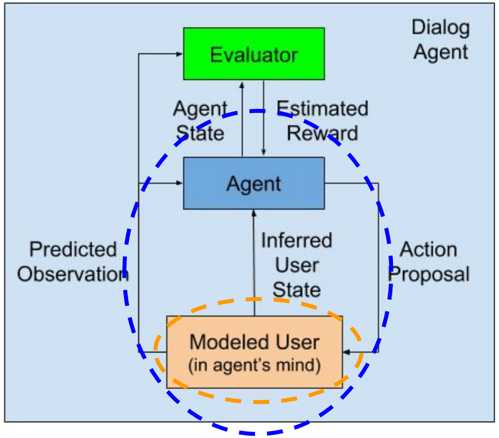
Proposed Dialog Learning Methods

- ❖ Learning End-to-End Task-Oriented Dialog via User Interaction
- ❖ **Learning from Simulated Experiences**
  - ➢ Modeling user dynamics
  - ➢ Co-training of dialog agent and modeled user

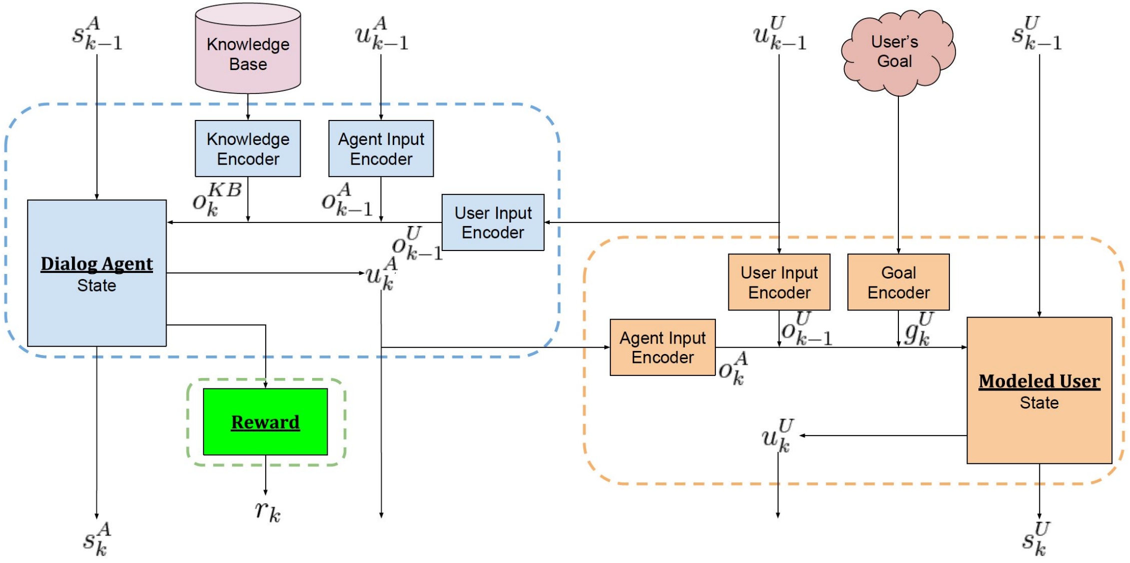
User Modeling

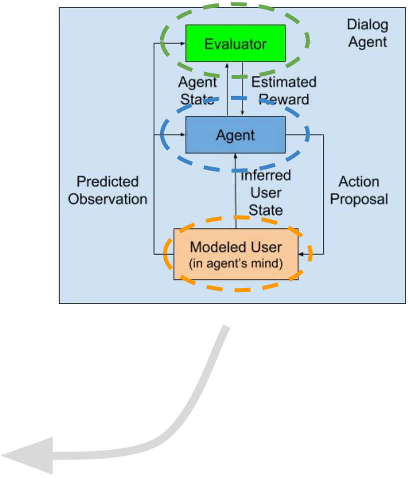
- ❖ Querying (i.e. interacting with) human user for dialog model training can be slow and inefficient.
- ❖ Previous works use *rule-based user simulator* → Fixed and can quickly become a bottleneck for dialog agent training.
- ❖ Building a reliable user model is not trivial, often as difficult as building a dialog agent.

Our Solution → Learn a basic user model, and continuously to improve it with the dialog agent



Iteratively Optimizing Dialog Agent and Modeled User

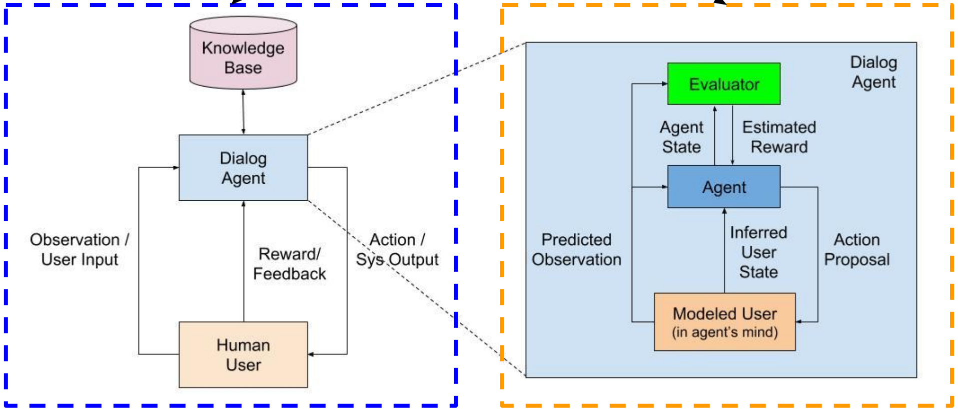




Integrated Learning from Real and Simulated Interactions

- ❖ Learning from real dialogs by querying users is slow and sample inefficient
- ❖ Learning from simulated interactions is limited by the modeled user capacity

Solution → Integrated learning from real and simulated experiences



Conclusions

- ❖ We present an end-to-end trainable system for task-oriented dialogs.
- ❖ We design a hybrid learning framework with:
  - ➢ Offline learning from fixed dialog corpora
  - ➢ Interactive learning from human demonstration and feedback
- ❖ We propose an integrated learning strategy by learning from both real and simulated experiences.

Thanks!

Q & A