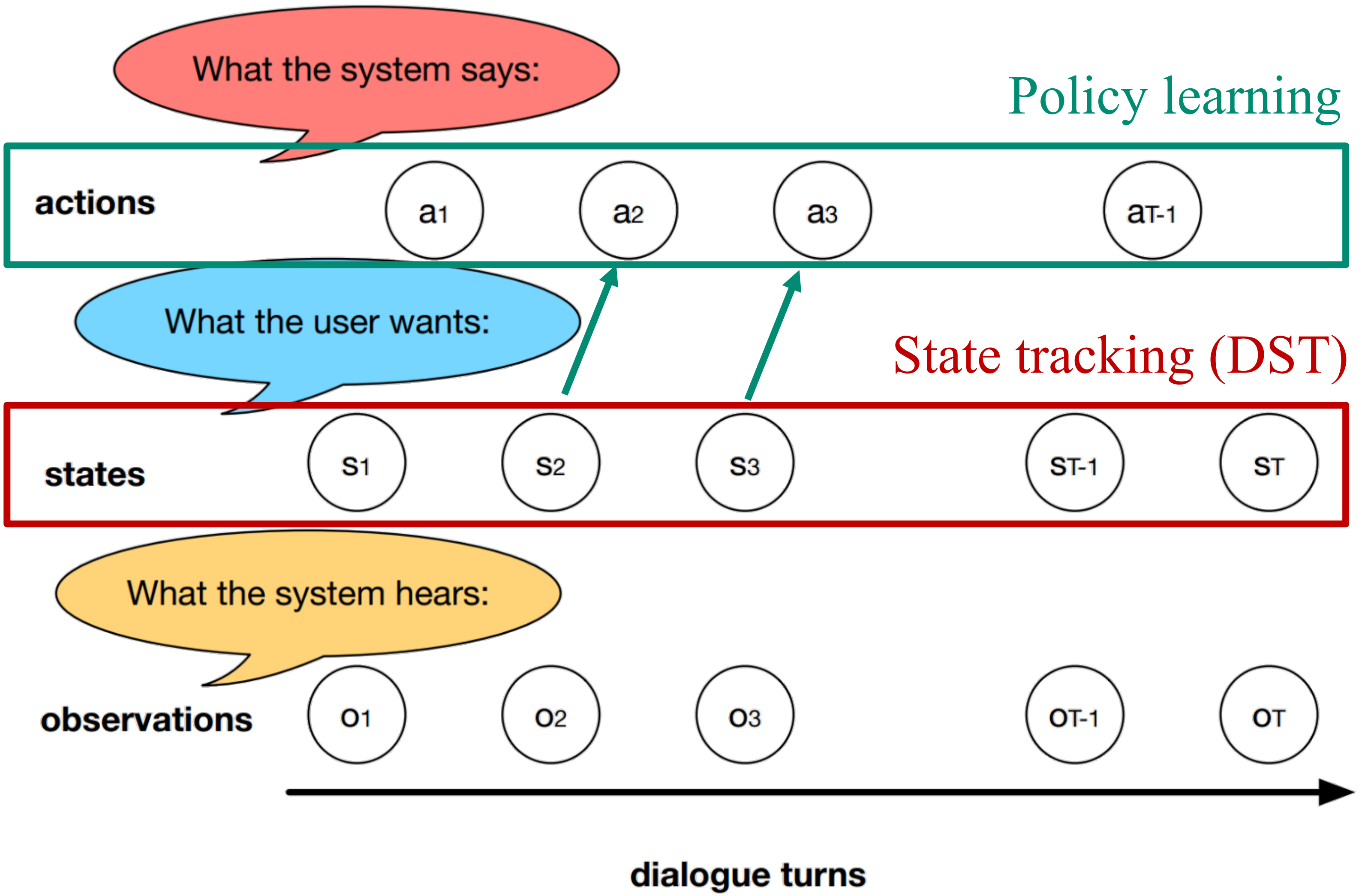
Dialog Manager (policy learner)

State tracking and policy learning



Dialog policy

- Dialog state → Agent act

- Policy execution examples:

`inform(location="780 Market St")`
The nearest one is at 780 Market St

`request(location)`
What is the delivery address?

Simple approach: hand crafted rules

- You have NLU and state tracker
- You can come up with hand crafted rules for policy

Optimizing dialog policies with ML

Supervised learning:

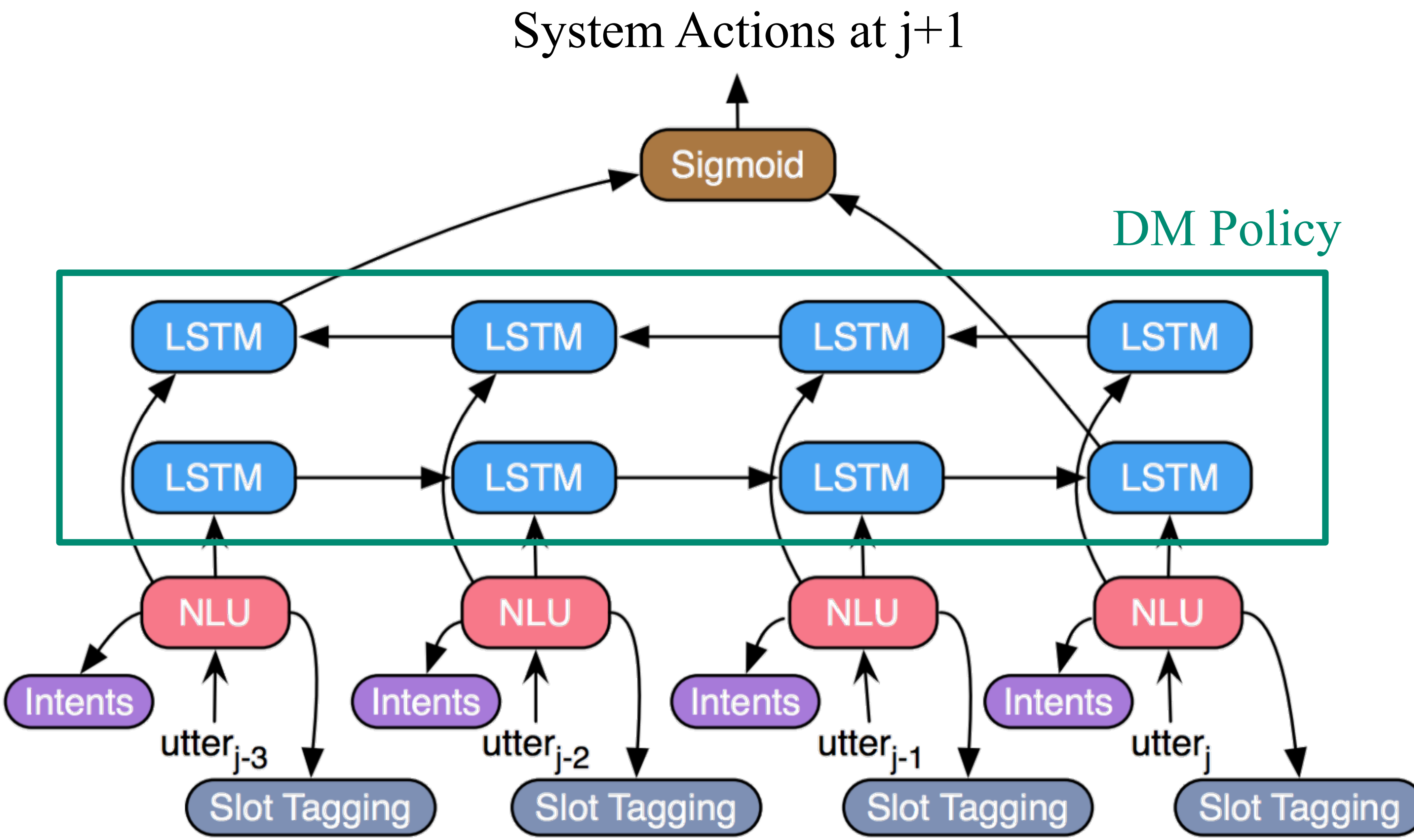
- You train to imitate the observed actions of an expert
- Often requires a large amount of expert-labeled data
- Even with a large amount of training data, parts of the dialogue state space may not be well-covered in the training data

Reinforcement learning:

- Given only a reward signal, the agent can optimize a dialogue policy through interaction with users.
- RL can require many samples from an environment, making learning from scratch with real users impractical
- That's why we need *simulated users* for RL

https://arxiv.org/pdf/1703.07055.pdf

Joint NLU and DM



https://arxiv.org/pdf/1612.00913.pdf

Joint NLU and DM results

Model	DM	NLU
Baseline (CRF + SVMs)	7.7	33.1
Pipeline-BLSTM	12.0	36.4
Joint Model	**22.8**	**37.4**

Frame level accuracies on DSTC 4
(it counts only when the whole frame parse is correct)

https://arxiv.org/pdf/1612.00913.pdf

Summary

- Dialog policy can be done by hand crafted rules
- Or in a supervised way
- Or in a reinforcement learning way