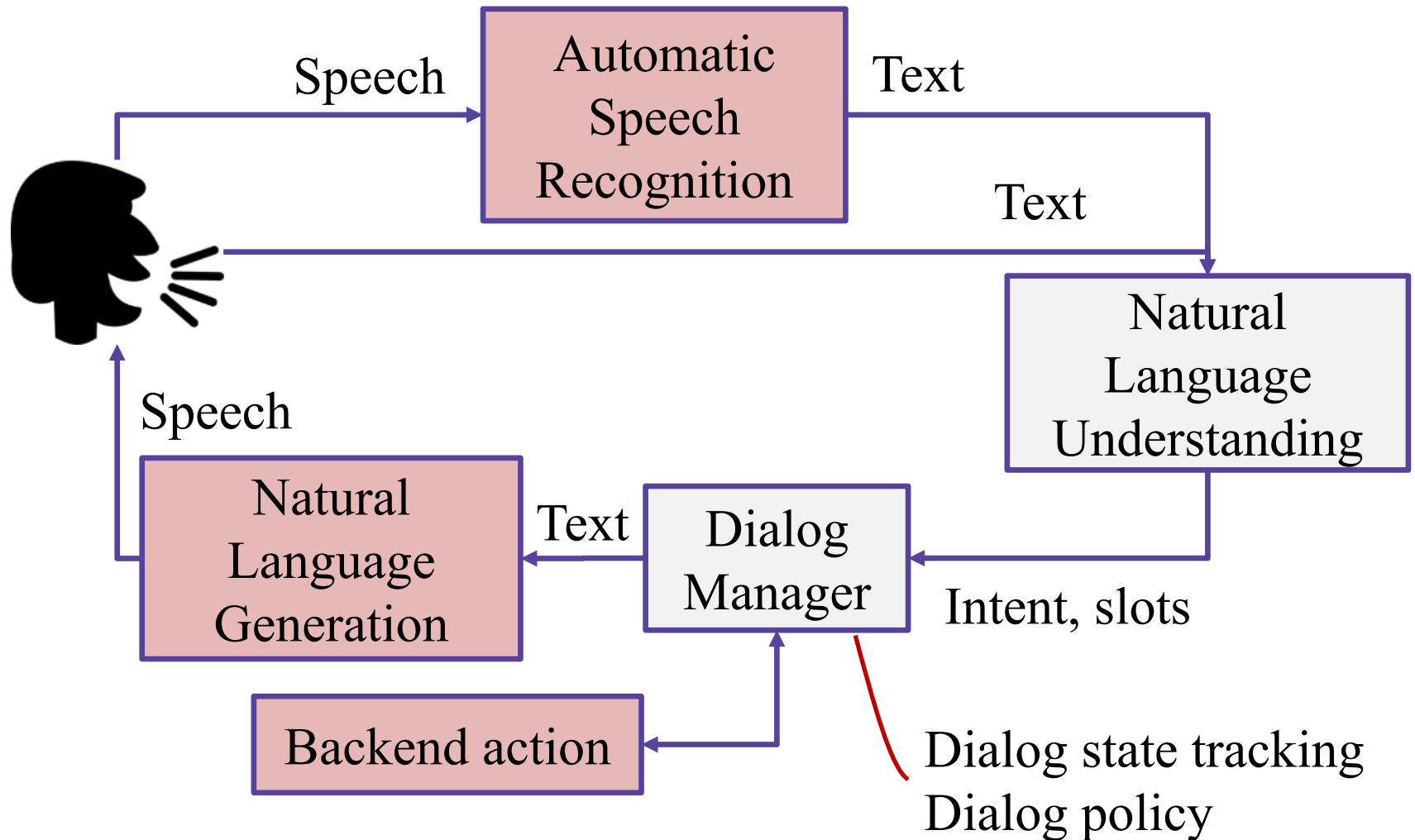


Final remarks

Task-oriented dialog system overview



NLU and DM

- You can train slot tagger and intent classifier in NLU separately
- Or jointly...
- You can train NLU and DM separately
- Or jointly...
- You can use hand-crafted rules sometimes (e.g. dialog policy)
- But learning from data actually works better

Evaluation

- **NLU:**
 - Turn-level metrics: intent accuracy, slots F1
- **DM:**
 - Turn-level metrics: state tracking accuracy, ...
 - Dialog-level metrics: task success rate, reward, ...

The importance of NLU

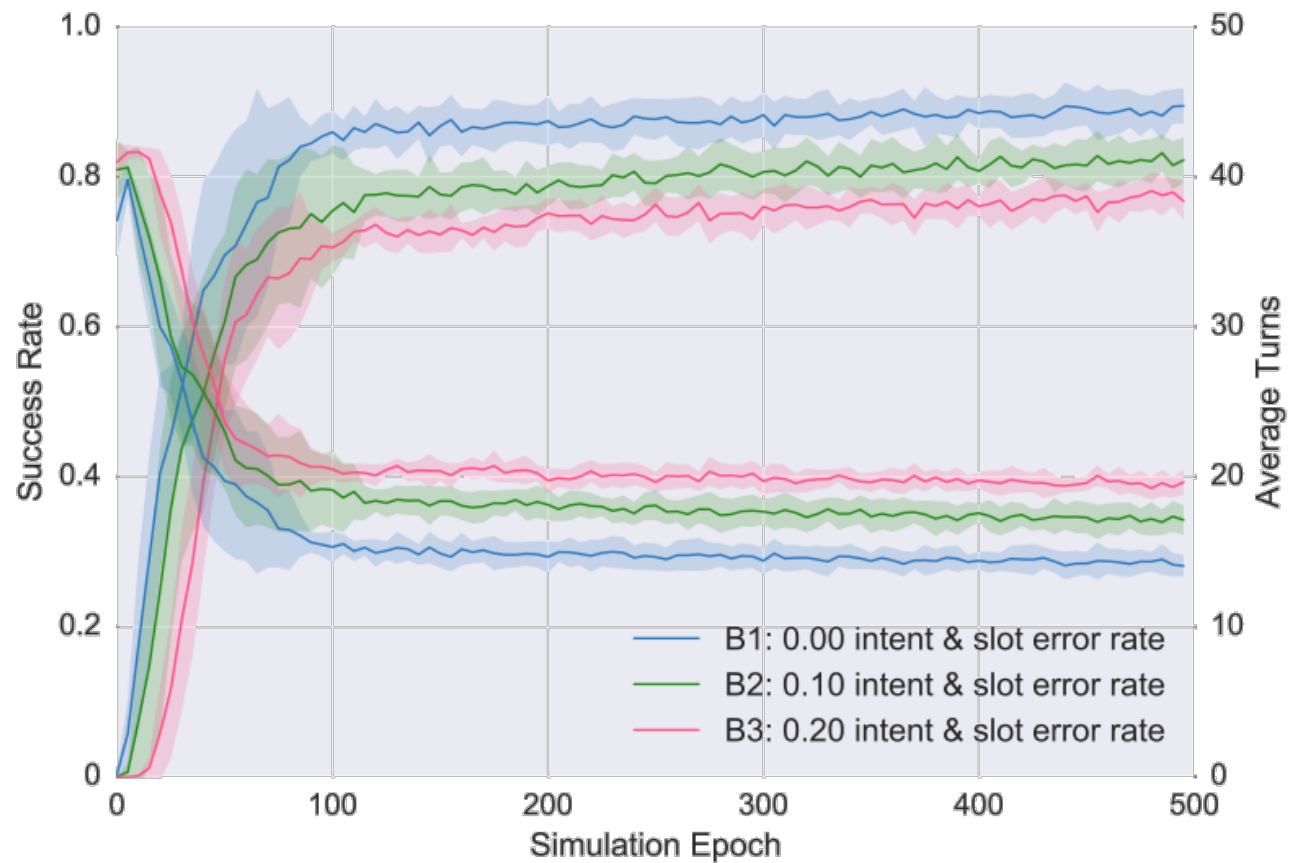
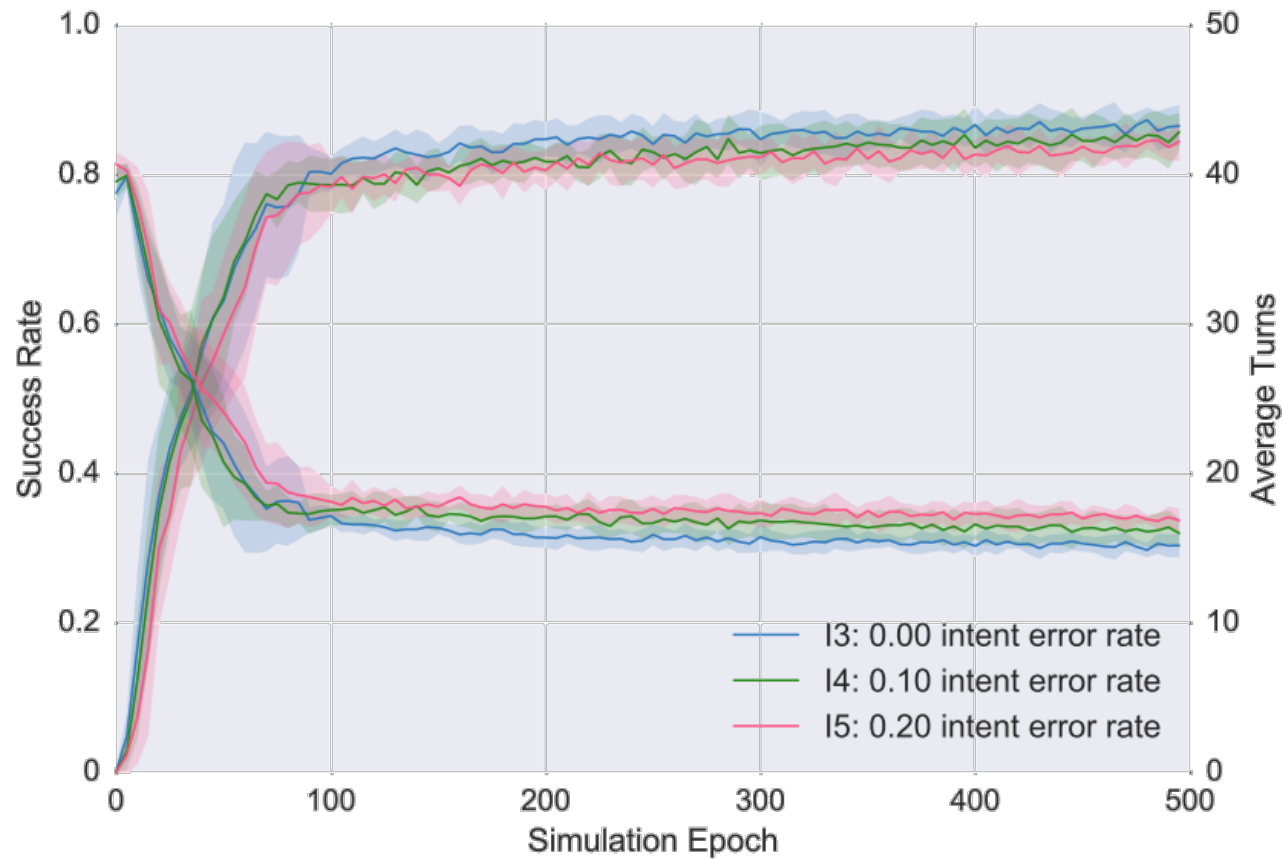


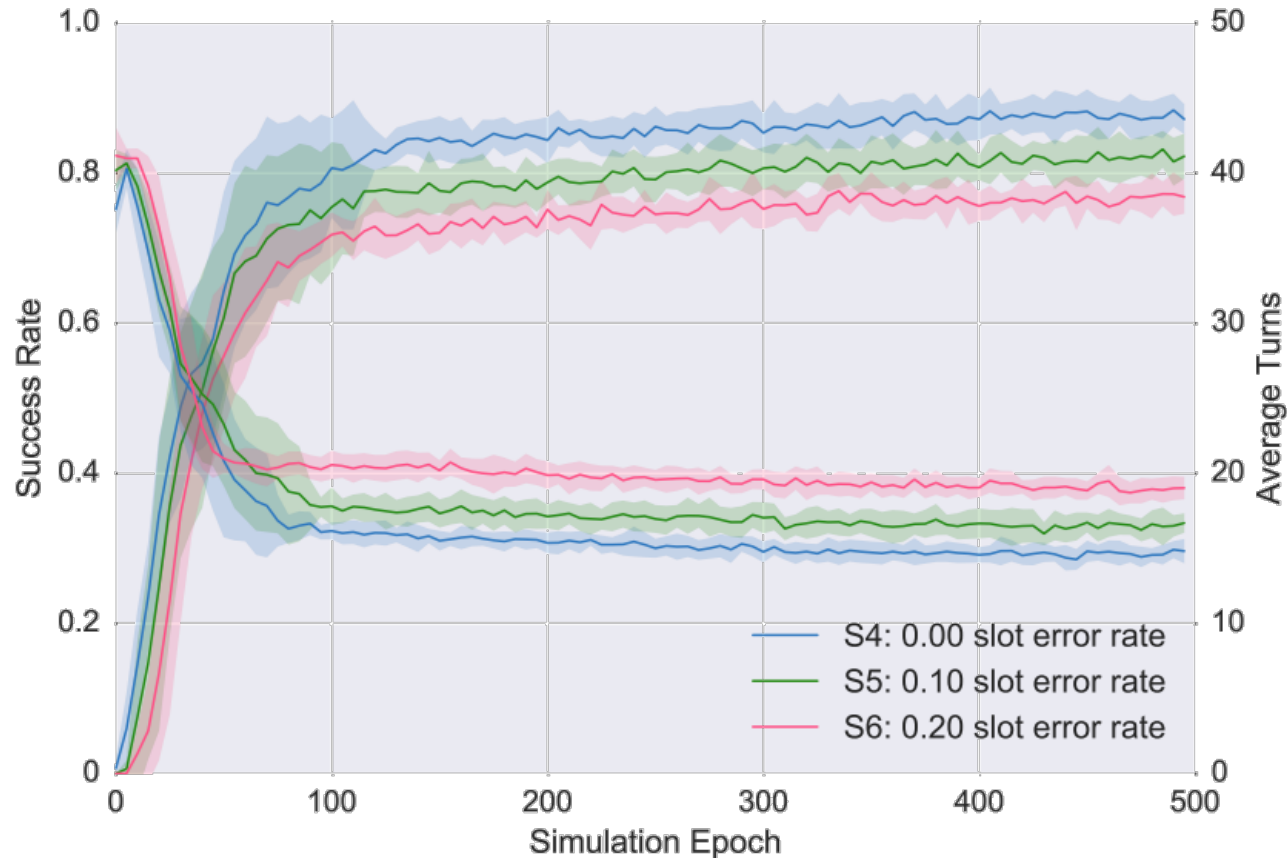
Figure 1: *Learning curves for different NLU error rates.*

What is more important: slots or intent?



(b) Intent Error Rate Analysis

What is more important: slots or intent?



(d) Slot Error Rate Analysis

Looks like we need to concentrate on slot tagger

Summary

- That's it for the week!
- Good luck with your final project!