

# Dialogue State Generation: Transcending Slot Schemas for Domain-General State Inference

James D. Finch, Boxin Zhao, Jinho D. Choi

Emory University

# Dialogue State Tracking

1 Welcome to Walkworth Inn! How can I assist today?



2 Hi, I'd like to make a reservation for this upcoming Monday.



3 Of course! How many people will be in your party, and for how many nights would you like to book?



4 It's me, my wife, and 2 kids, and also my brother and his wife. We're staying for 5 days. What's the cheapest booking you can give us?



5 Let me calculate the most cost-effective option for your party.  
For a group of six adults and two children, we could book Room 122 and Room 156.  
This arrangement would provide a total of three beds, which should be sufficient for your party.  
The total cost for both rooms for the duration of your stay would be \$1815.



# Dialogue State Tracking

1 Welcome to Walkworth Inn! How can I assist today?

2 Hi, I'd like to make a reservation for this upcoming Monday.

book day

3 Of course! How many people will be in your party, and for how many nights would you like to book?

4 It's me, my wife, and 2 kids, and also my brother and his wife. We're staying for 5 days. What's the cheapest booking you can give us?

num guests

price range

num nights

5 Let me calculate the most cost-effective option for your party.  
For a group of six adults and two children, we could book Room 122 and Room 156.  
This arrangement would provide a total of three beds, which should be sufficient for your party.  
The total cost for both rooms for the duration of your stay would be \$1815.

total cost

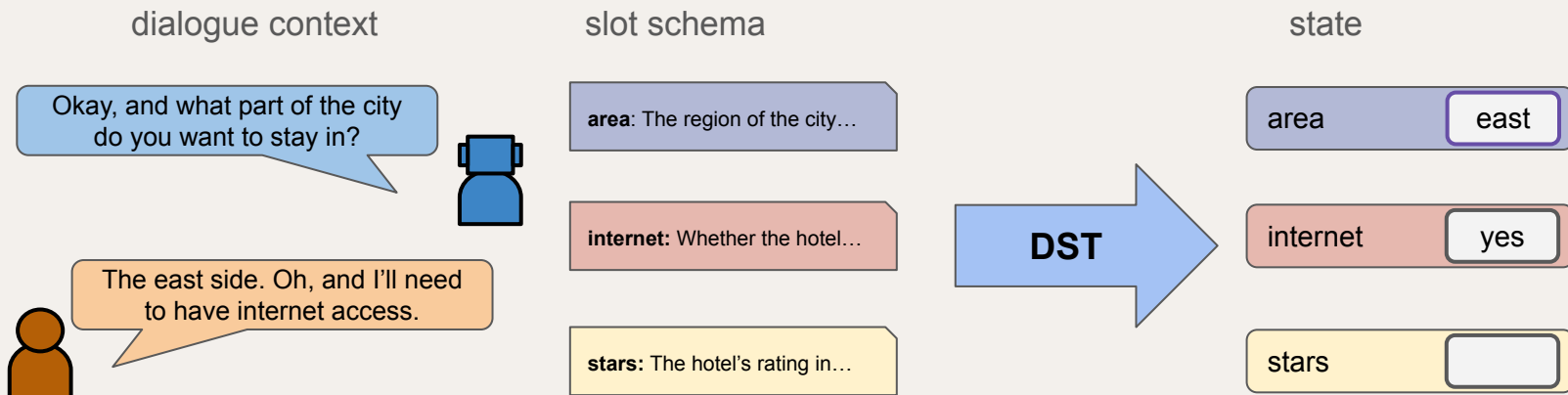
num guests

book rooms

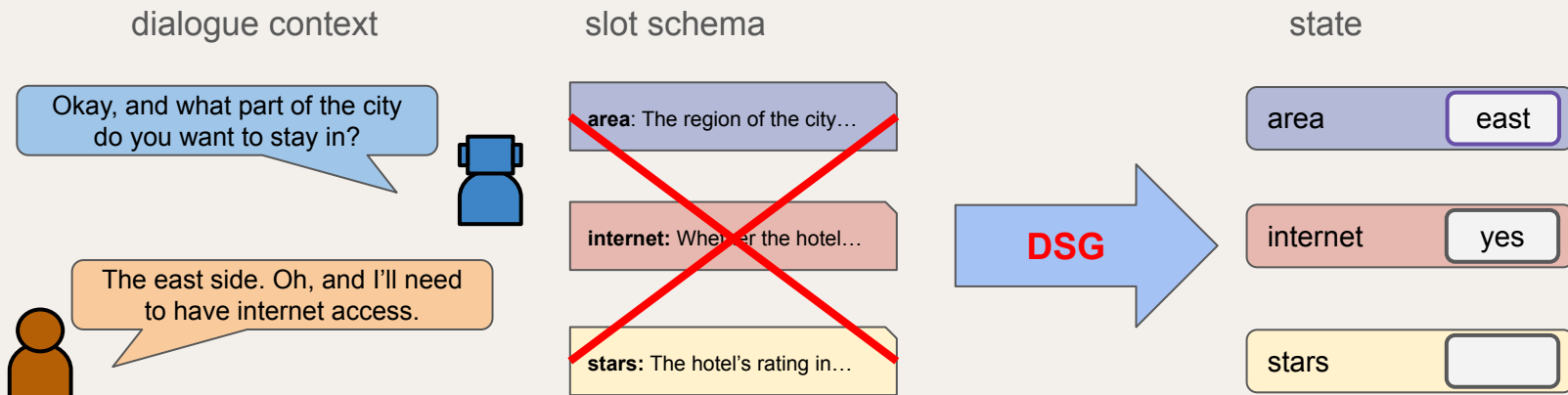
# Limited DST Domain Diversity

- Slots are different for every application
- DST data limited in diversity
- Tackle domain generalizability via *Dialogue State Generation*

	Domains	Slot Types	Dialogues
<b>MultiWOZ</b>	7	24	8,438
<b>SGD</b>	16	214	16,142



# Dialogue State Generation



# Dialogue State Generation using ILLMs

Dialogue Context

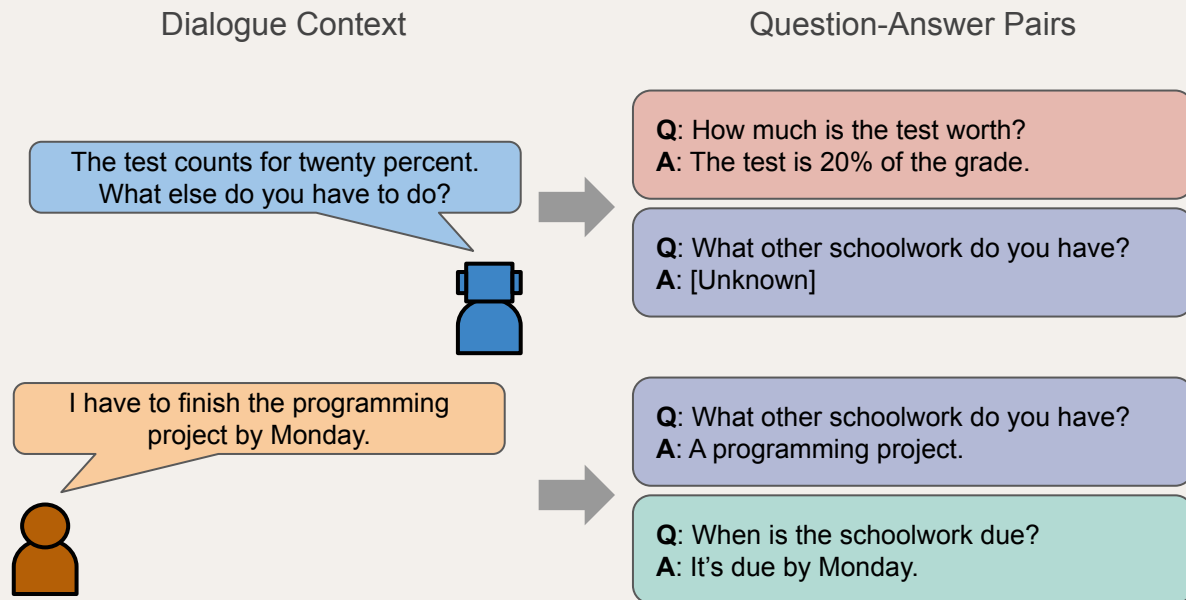
The test counts for twenty percent.  
What else do you have to do?



I have to finish the programming  
project by Monday.

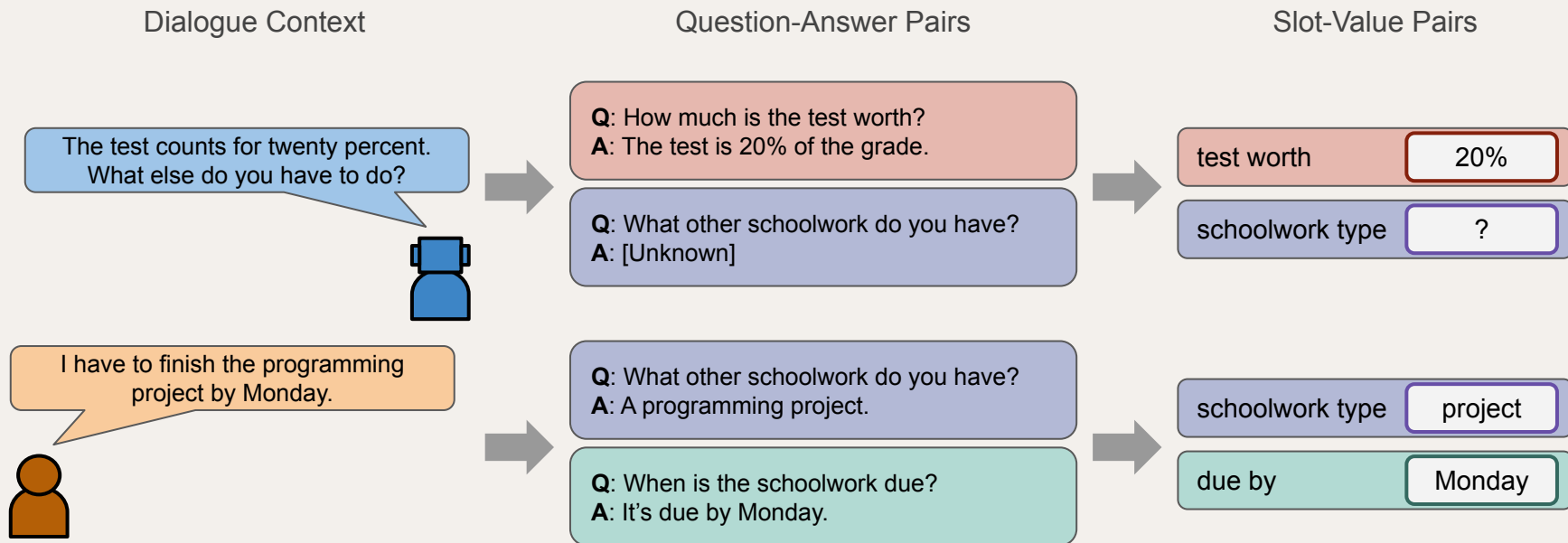


# GPTPipe: DSG with LLMs





# GPTPipe: DSG with LLMs



# DSG5K: Domain-Diverse Dialogue Dataset

## Scenarios

"Please list everyday scenarios..."

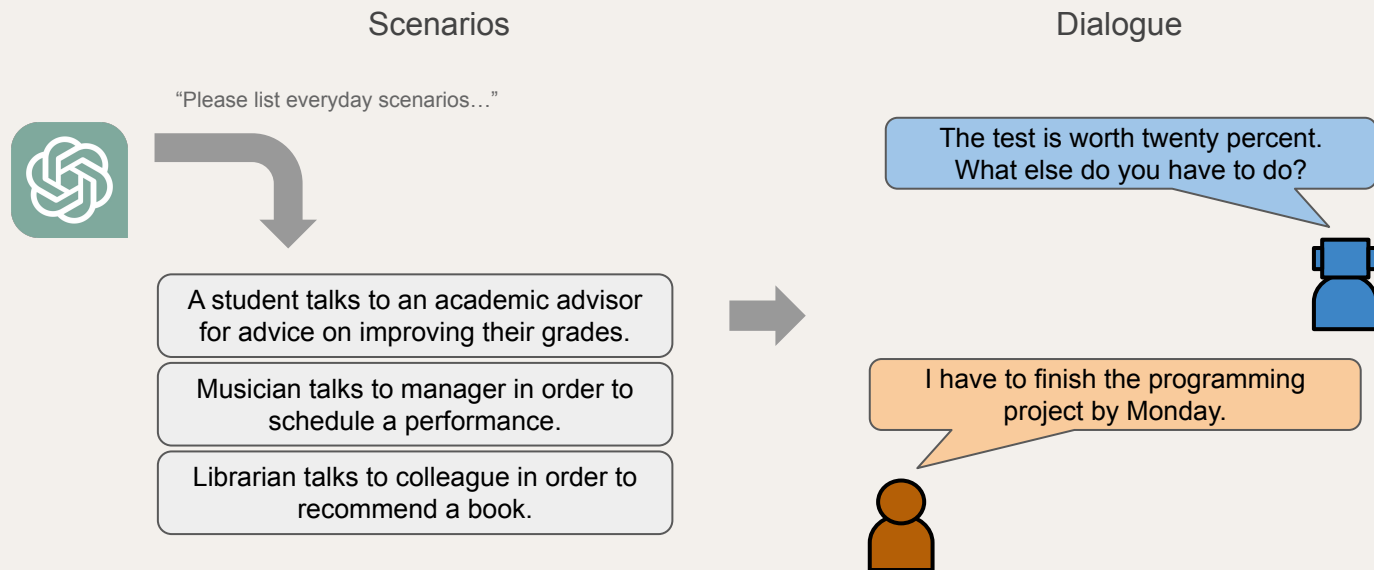


A student talks to an academic advisor  
for advice on improving their grades.

Musician talks to manager in order to  
schedule a performance.

Librarian talks to colleague in order to  
recommend a book.

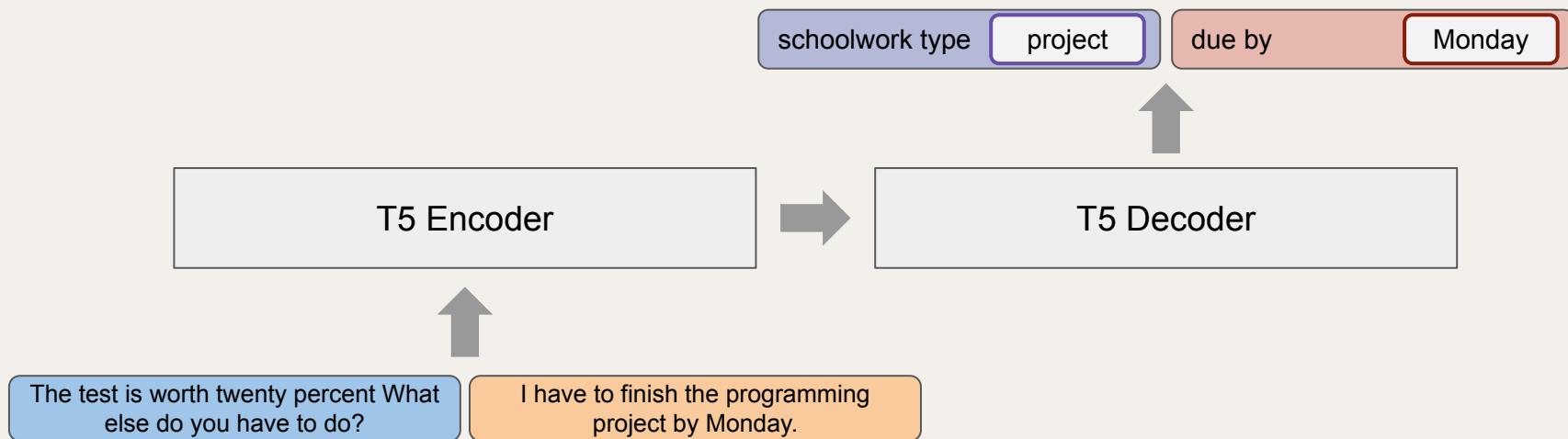
# DSG5K: Domain-Diverse Dialogue Dataset



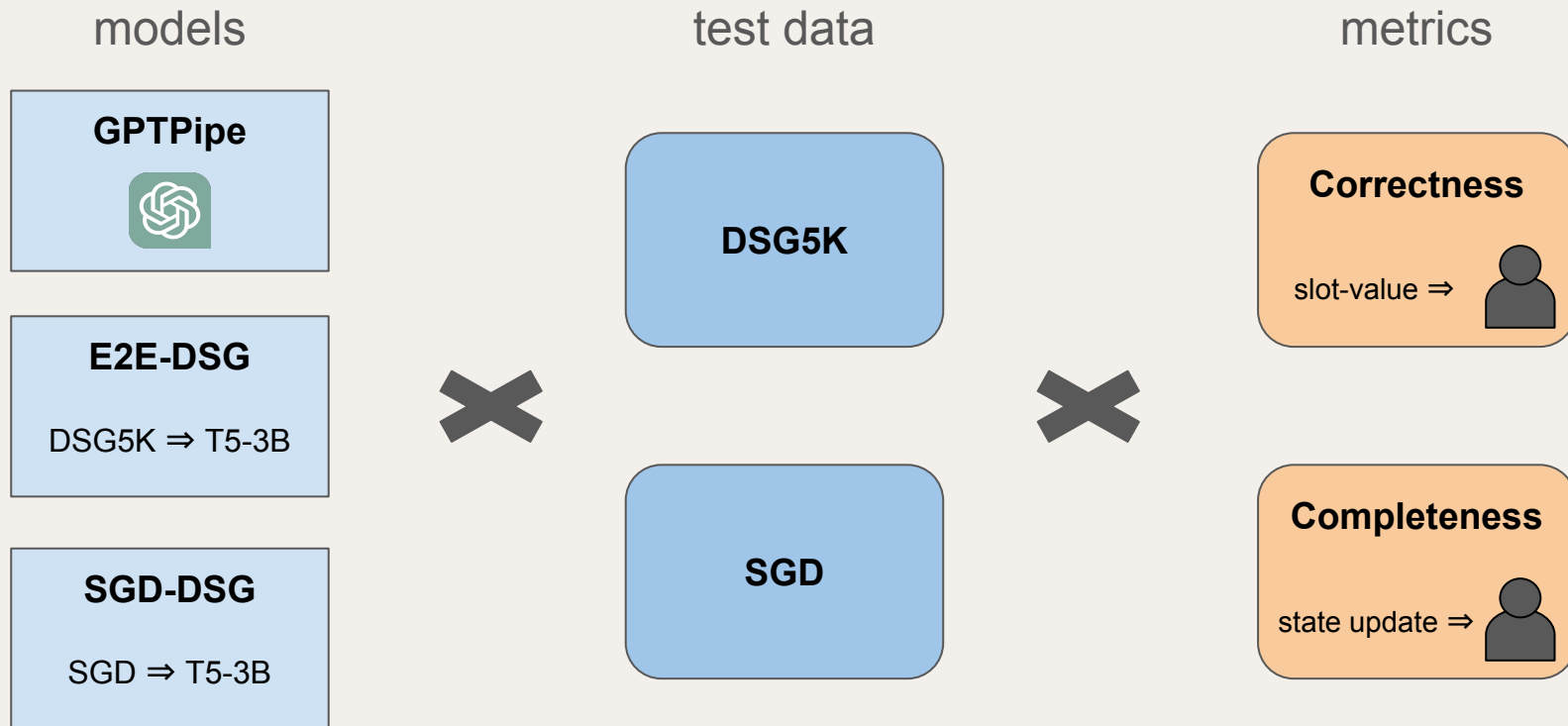
# DSG5K: Domain-Diverse Dialogue Dataset

- 1,000 scenarios
  - mix of task-oriented and “role-oriented”
- 5 dialogues per scenario
- 5,000 dialogues total
- Automatically annotated with GPTPipe

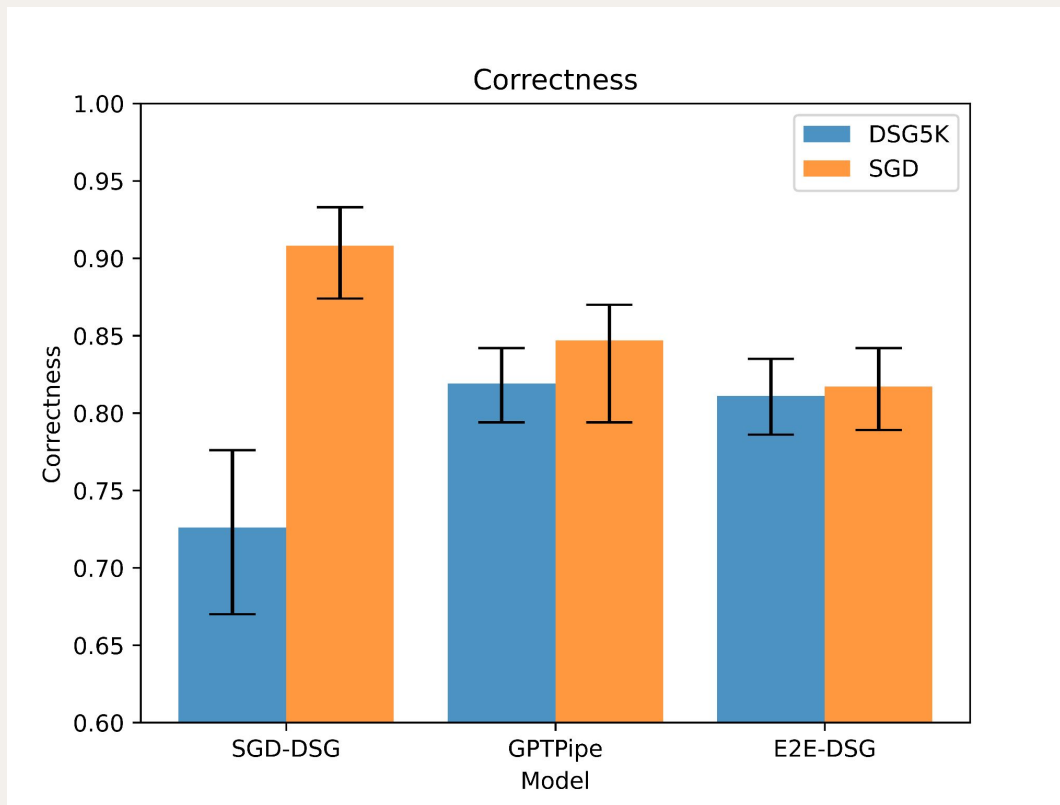
# End-to-End Model



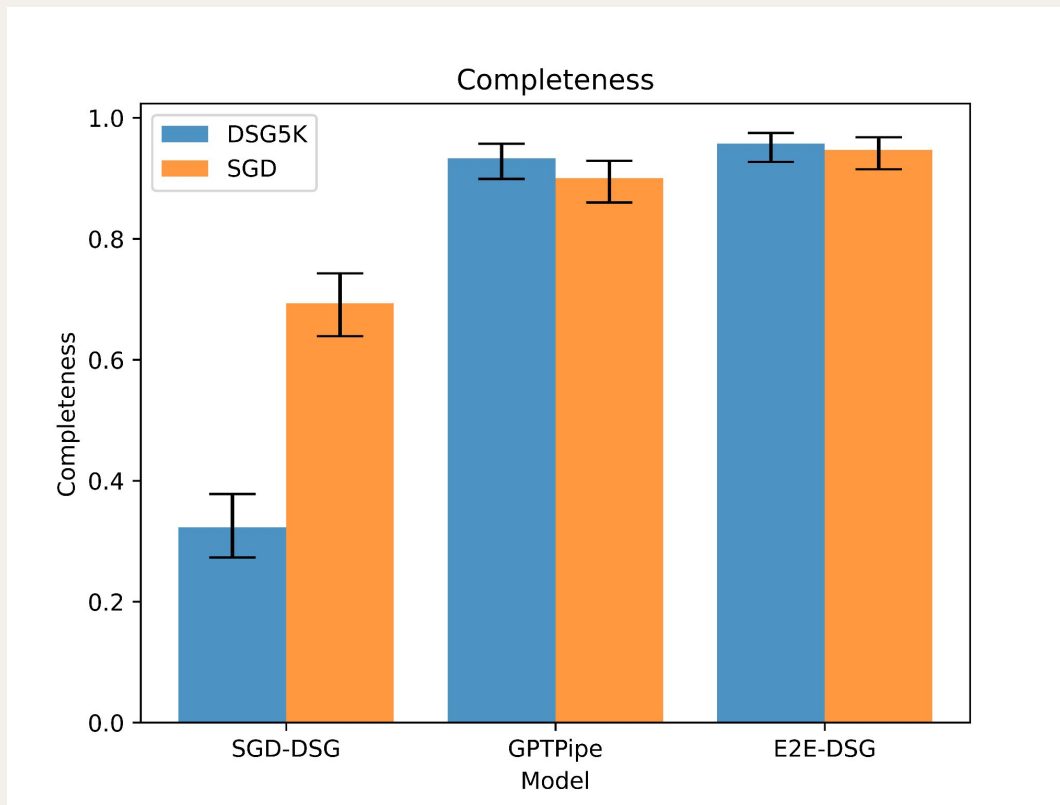
# Experiments



# Results - Correctness

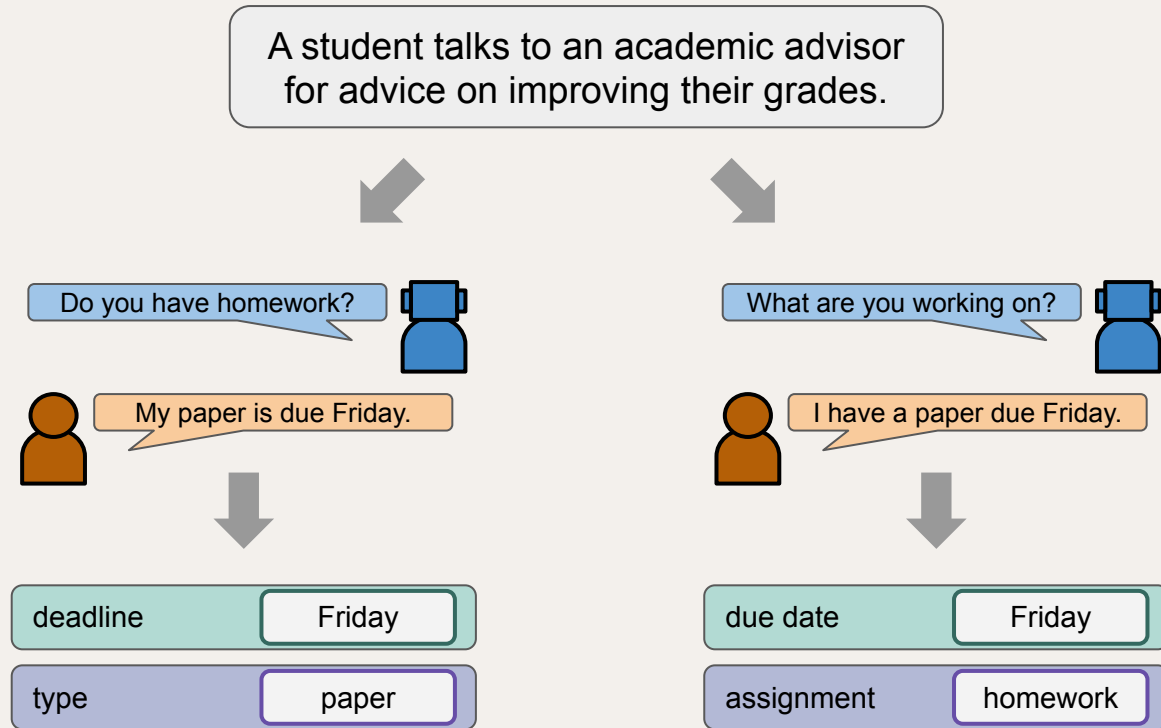


# Results - Completeness





# Limitation: Schema Inconsistency



# Conclusion

- Domain-general inference of dialogue state
- Limitation: schema consistency
- Future work: resolve schema consistency
- Use DSG for generating diverse DST training data

# References

- Budzianowski, Paweł, Tsung-Hsien Wen, Bo-Hsiang Tseng, Iñigo Casanueva, Stefan Ultes, Osman Ramadan, and Milica Gašić. 2018. “MultiWOZ - A Large-Scale Multi-Domain Wizard-of-Oz Dataset for Task-Oriented Dialogue Modelling.” In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 5016–26. Brussels, Belgium: Association for Computational Linguistics. <https://doi.org/10.18653/v1/D18-1547>.
- Rastogi, Abhinav, Xiaoxue Zang, Srinivas Sunkara, Raghav Gupta, and Pranav Khaitan. 2020. “Towards Scalable Multi-Domain Conversational Agents: The Schema-Guided Dialogue Dataset.” *Proceedings of the AAAI Conference on Artificial Intelligence* 34 (05): 8689–96. <https://doi.org/10.1609/aaai.v34i05.6394>.
- Raffel, Colin, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. “Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer.” *The Journal of Machine Learning Research* 21 (1): 140:5485-140:5551.