

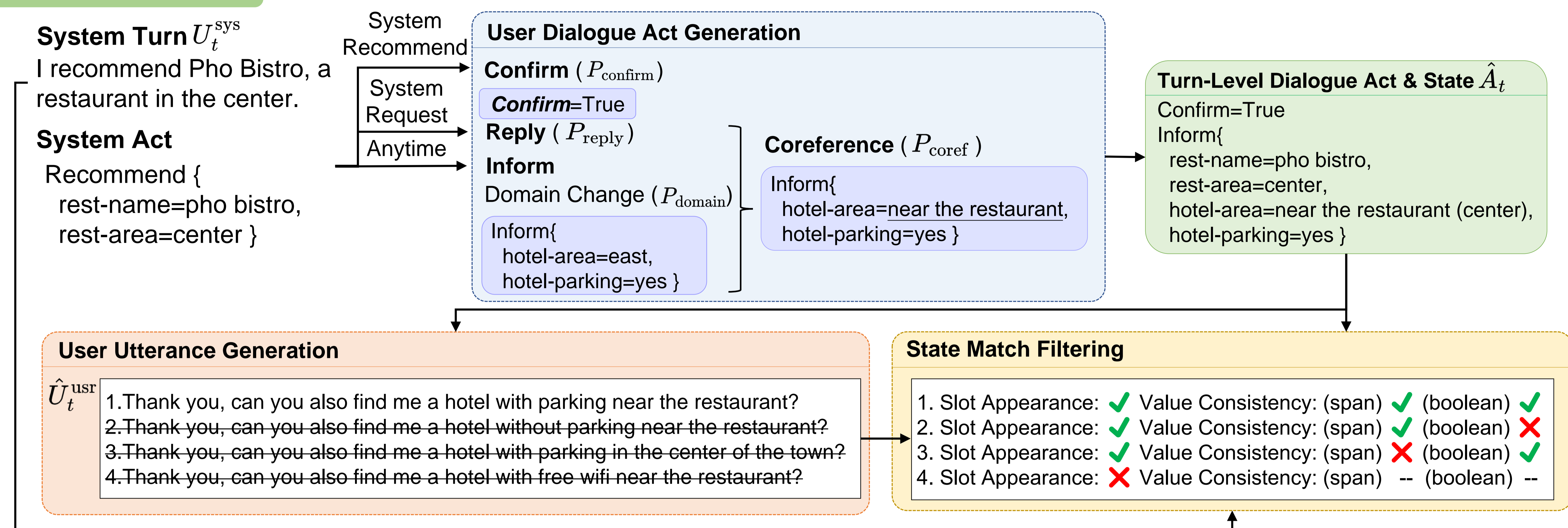
## Summary

- CoCo (Li et al., 2020) improved DST by data augmentation for better **robustness** and **generalizability**.
- Only a simple type of utterances is considered for augmentation.
- Our goal is to augment user utterances with **diverse dialogue acts** for better generalization capability.
- Our proposed CUDA achieves SOTA performance and better robustness on MultiWOZ 2.1.

[System<sub>1</sub>]: Hello, how can I help you?  
 [User<sub>1</sub>]: I need to find a restaurant in the center.  
 [System<sub>2</sub>]: I recommend Pho Bistro, a popular restaurant in the center.  
 Recommend { restaurant-name=pho bistro, rest-area=center }  
 [User<sub>2</sub>]: No, it needs to serve British food and I'd like a reservation for 18:00.  
 Confirm=False, Inform{ rest-area=center, rest-food=British, rest-time=18:00 }  
 [VS-User<sub>2</sub>]: No, it needs to serve Chinese food and I'd like a reservation for 17:00.  
 Confirm=False, Inform{ rest-area=center, rest-food=Chinese, rest-time=17:00 }  
 [CoCo-User<sub>2</sub>]: No, it should serve Chinese food and I need to book a table for 2 people.  
 Confirm=False, Inform{ rest-area=center, rest-food=Chinese, rest-people=2 }  
 [CUDA-User<sub>2</sub>]: Thank you, can you also find me a hotel with parking near the restaurant?  
 Confirm=True, Inform{ rest-area=center, rest-name=pho bistro, hotel-area=center, hotel-parking=yes }

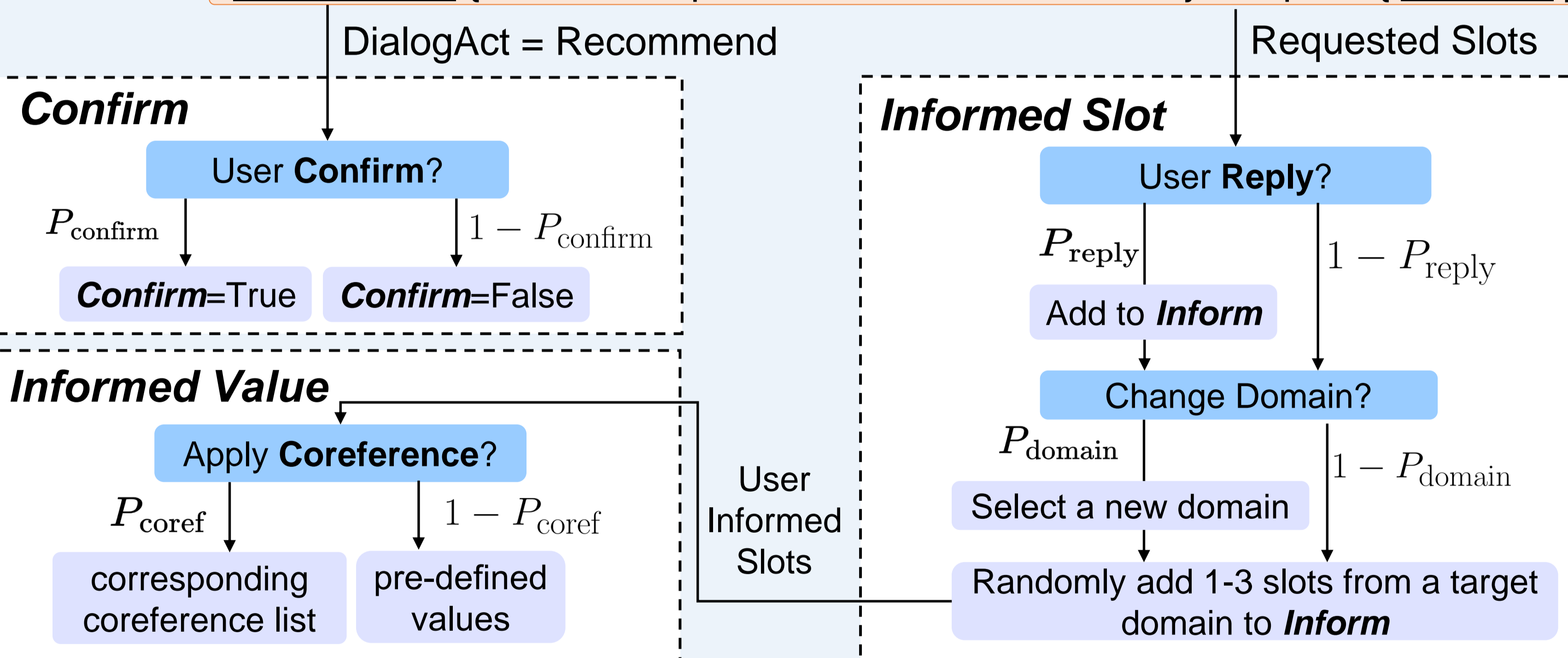
Confirm    Domain Change    Coreference    Boolean

## Framework



## User Dialogue Act Generation

**System Turn** I recommend Pho Bistro in the center. What time do you plan to book the table?  
 Recommend { rest-name=pho bistro, rest-area=center }; Request { rest-time }



- Goal: simulate *reasonable* user behavior
- Method: use a random process to augment more diverse behaviors
- User behaviors:
  - **Confirm**: The user confirms the system's recommendation.
  - **Reply**: The user replies what the system requests.
  - **Inform**: The user actively informs desired slot values to the system.

## User Utterance Generation

- Goal: generate utterances aligned with the given behaviors
- Method: T5 fine-tuning for conditional generation (beam search for diversity)

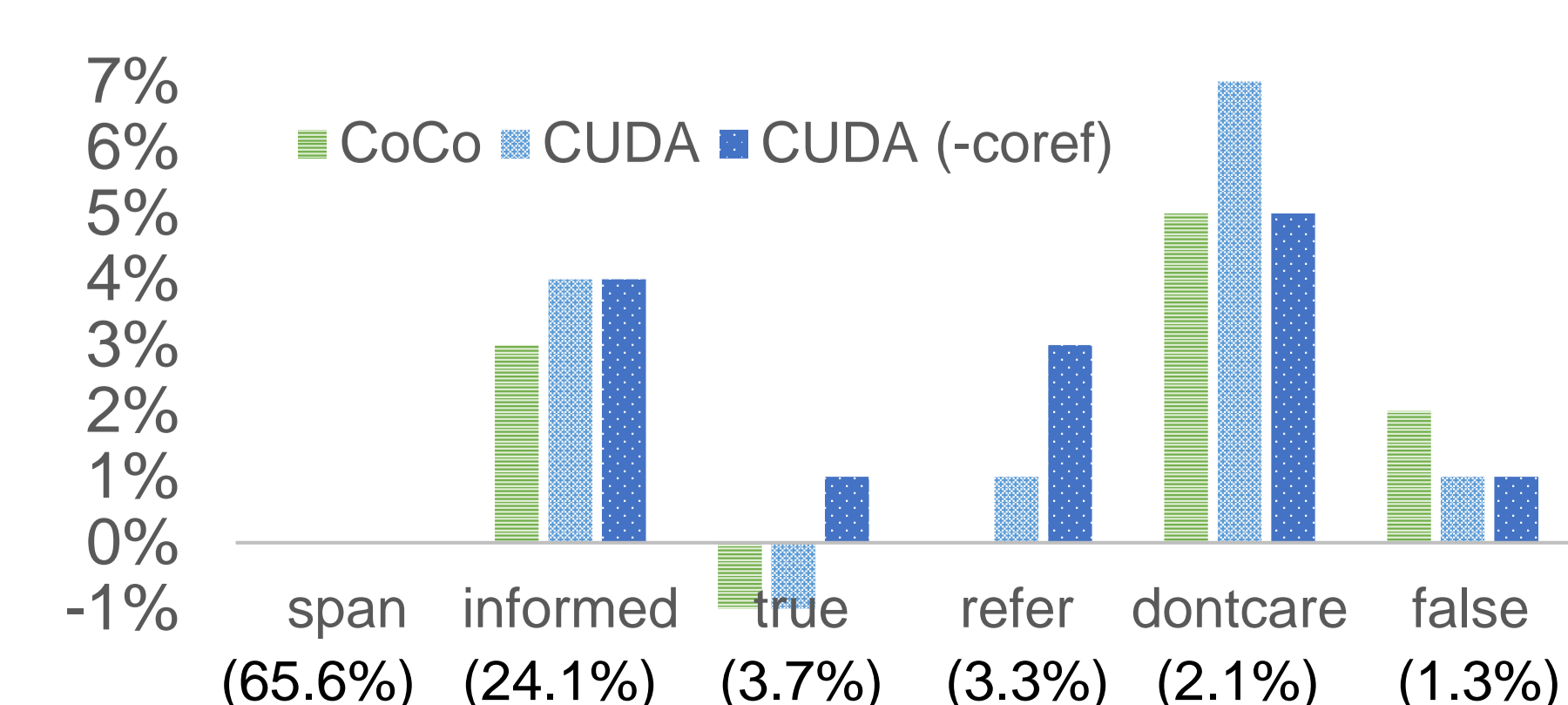
## State Match Filtering

- **Slot appearance** classifier
  - ✓ Goal: check if the given slots are included
  - ✓ Method: BERT for multi-label classification
- **Value consistency** filter
  - ✓ Goal: ensure consistent values between states & utterances

## Experiment

MultiWOZ 2.1	TripPy	TRADE
Original	57.72	44.08
CoCo	60.46	43.53
CUDA	<b>61.28</b> ♦	<b>44.86</b> ♦
CUDA (-coref)	<b>62.93</b> ♦	42.98

### Improvement



- CUDA improves TripPy and TRADE results.
- The models trained on our augmented data show better generalization.
- CUDA improves more on *informed*, *refer*, and *dontcare* slots than CoCo.
- CUDA augments diverse user dialogue acts for helping *informed* and *refer*, and the proposed filter ensures value consistency for improving *dontcare*.