

# Multimodal Coreference Resolution

# Multimodal Coreference Resolution



Hi, do you have short  
sleeve t-shirts?



UNIVERSITY OF  
CAMBRIDGE

Alejandro Santorum Varela  
Svetlana Stoyanchev, Kate Knill

**TOSHIBA**

# Multimodal Coreference Resolution



Hi, do you have short sleeve t-shirts?

Yes, we do. What color do you prefer?



UNIVERSITY OF  
CAMBRIDGE

Alejandro Santorum Varela  
Svetlana Stoyanchev, Kate Knill

**TOSHIBA**

# Multimodal Coreference Resolution



Hi, do you have short sleeve t-shirts?

Yes, we do. What color do you prefer?



White. Similar to the one next to the black hoodie.



UNIVERSITY OF  
CAMBRIDGE

Alejandro Santorum Varela  
Svetlana Stoyanchev, Kate Knill

**TOSHIBA**

# Multimodal Coreference Resolution



Hi, do you have short sleeve t-shirts?

Yes, we do. What color do you prefer?



White. Similar to the one next to the black hoodie.



UNIVERSITY OF  
CAMBRIDGE

Alejandro Santorum Varela  
Svetlana Stoyanchev, Kate Knill

**TOSHIBA**

# Multimodal Coreference Resolution



Hi, do you have short sleeve t-shirts?

Yes, we do. What color do you prefer?



White. Similar to the one next to the black hoodie.



dialogue history



UNIVERSITY OF  
CAMBRIDGE

Alejandro Santorum Varela  
Svetlana Stoyanchev, Kate Knill

**TOSHIBA**

# Multimodal Coreference Resolution



Hi, do you have short sleeve t-shirts?

Yes, we do. What color do you prefer?



White. Similar to the one next to the black hoodie.



objects data



dialogue history



UNIVERSITY OF  
CAMBRIDGE

Alejandro Santorum Varela  
Svetlana Stoyanchev, Kate Knill

**TOSHIBA**

# Multimodal Coreference Resolution



Hi, do you have short sleeve t-shirts?

Yes, we do. What color do you prefer?



White. Similar to the one next to the black hoodie.



objects data



scenes and visual data



dialogue history



UNIVERSITY OF  
CAMBRIDGE

Alejandro Santorum Varela  
Svetlana Stoyanchev, Kate Knill

**TOSHIBA**



# Multimodal Coreference Resolution



Hi, do you have short sleeve t-shirts?

Yes, we do. What color do you prefer?



White. Similar to the one next to the black hoodie.



objects data



scenes and visual data

dialogue history



UNIVERSITY OF  
CAMBRIDGE

Alejandro Santorum Varela  
Svetlana Stoyanchev, Kate Knill

**TOSHIBA**

# Multimodal Coreference Resolution



Hi, do you have short sleeve t-shirts?

Yes, we do. What color do you prefer?



White. Similar to the one next to the black hoodie.



objects data



coreference resolution

dialogue history

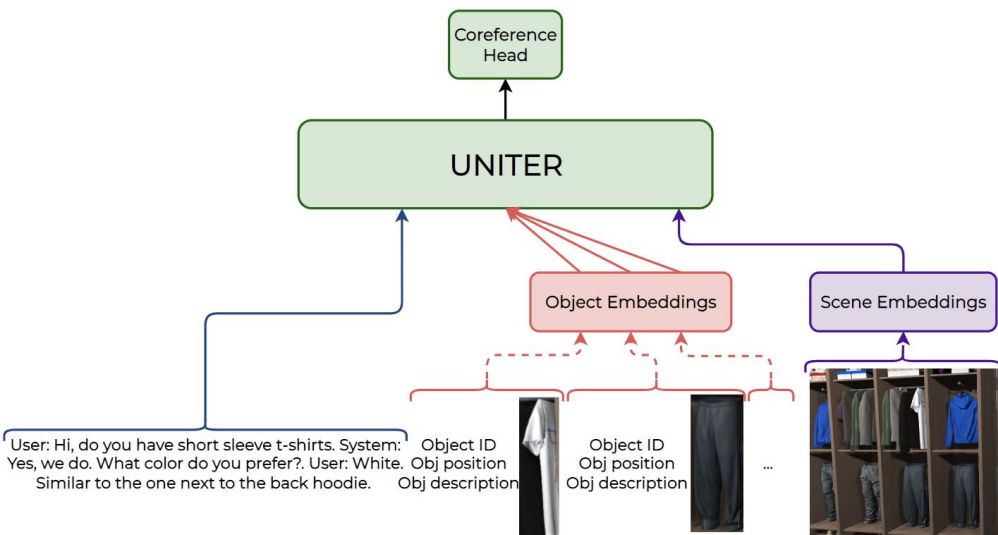


UNIVERSITY OF  
CAMBRIDGE

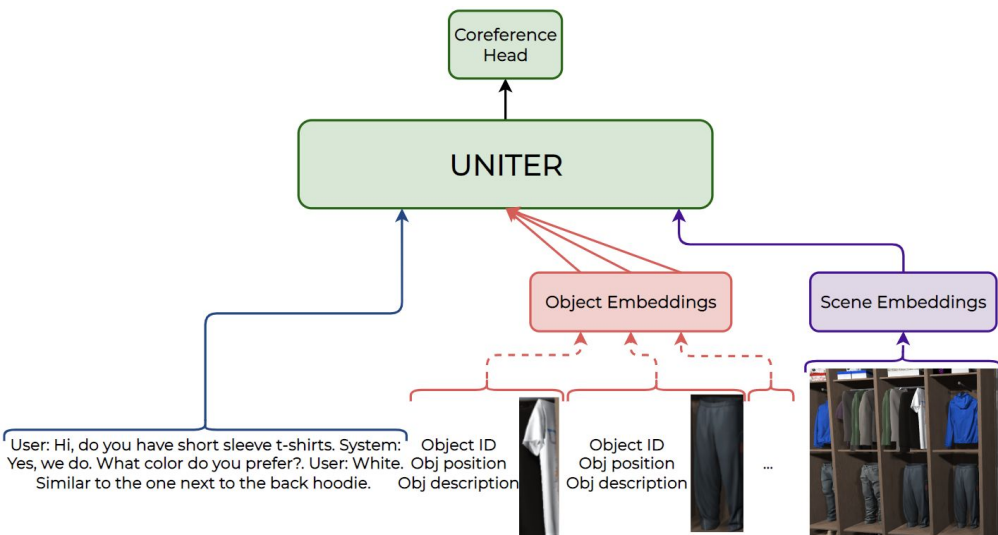
Alejandro Santorum Varela  
Svetlana Stoyanchev, Kate Knill

**TOSHIBA**

# Two state-of-the-art reviewed models

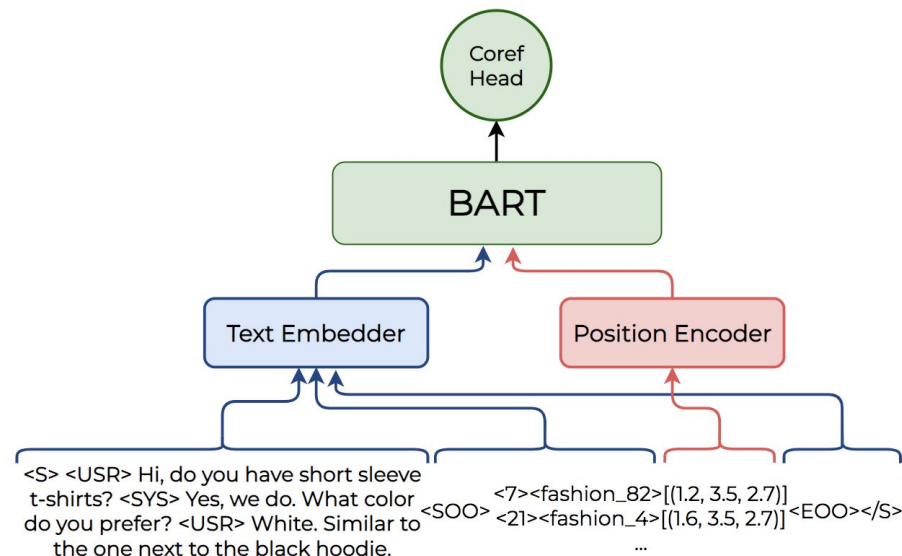
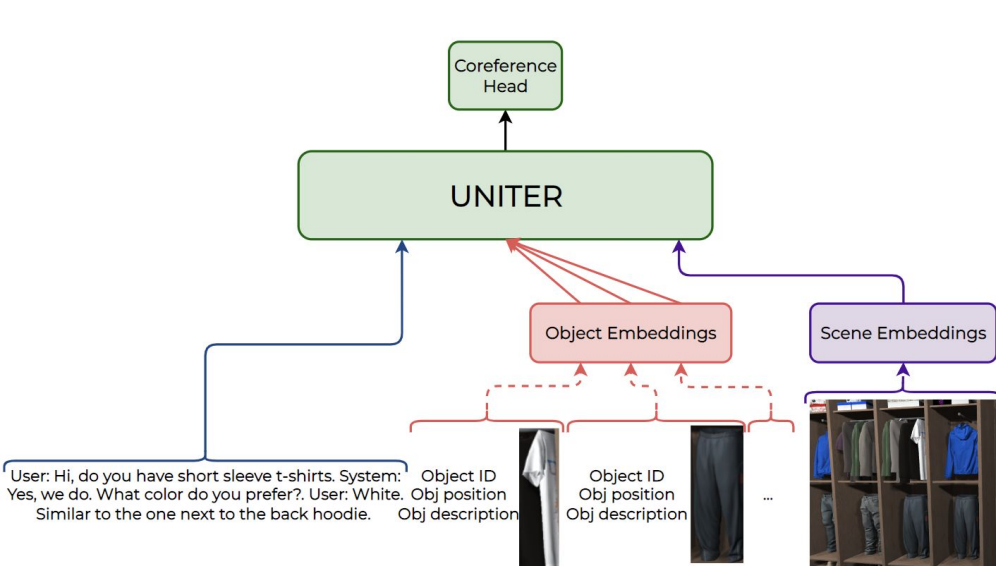


# Two state-of-the-art reviewed models



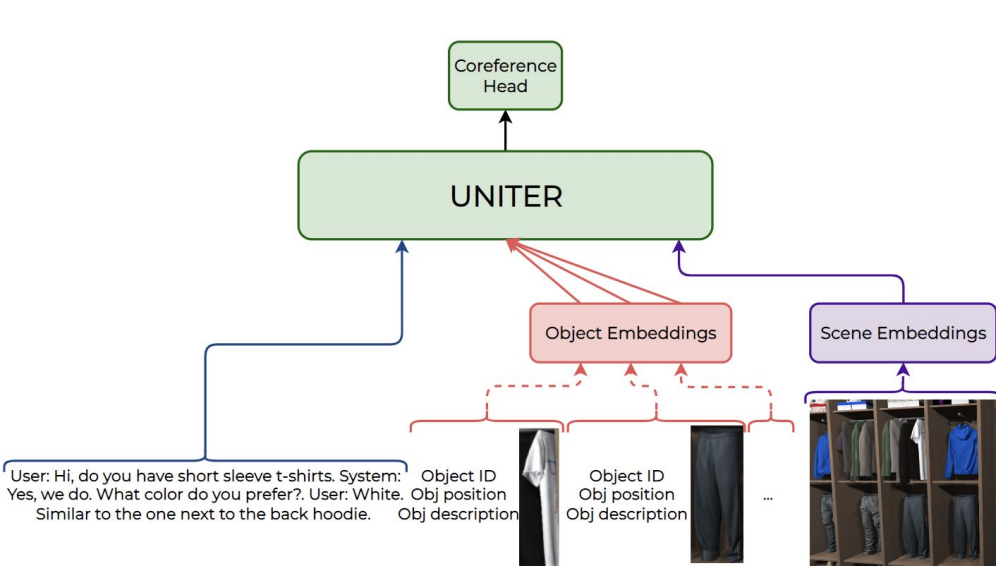
- Final reported performance: 67.4% F1-score
- Reference: Yichen Huang et. al. UNITER-Based Situated Coreference Resolution with Rich Multimodal Input. *Computing Research Repository*. 2021.

# Two state-of-the-art reviewed models

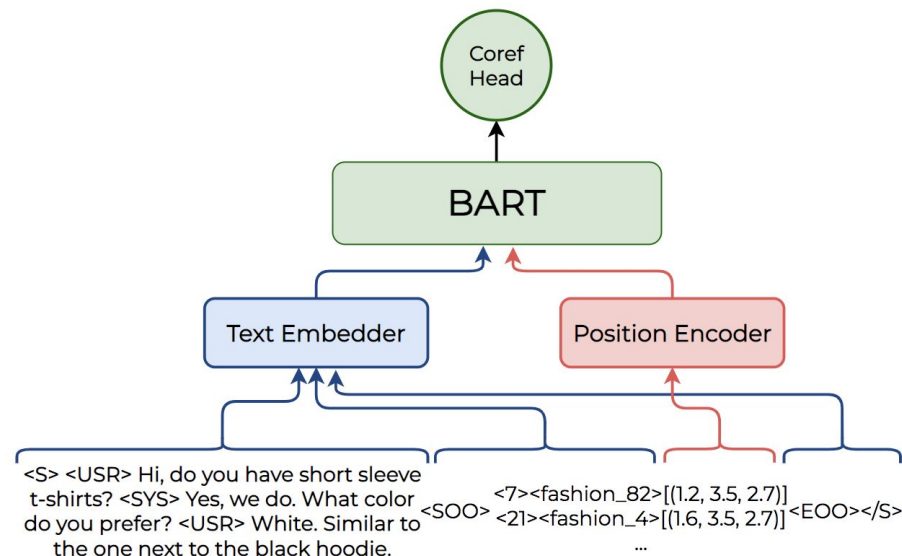


- Final reported performance: 67.4% F1-score
- Reference: Yichen Huang et. al. UNITER-Based Situated Coreference Resolution with Rich Multimodal Input. *Computing Research Repository*. 2021.

# Two state-of-the-art reviewed models



- Final reported performance: 67.4% F1-score
- Reference: Yichen Huang et. al. UNITER-Based Situated Coreference Resolution with Rich Multimodal Input. *Computing Research Repository*. 2021.



- Final reported performance: 74.3% F1-score
- Reference: Haeju Lee et. al. Tackling Situated Multi-Modal Task-Oriented Dialogs with a Single Transformer Model. *Association for Computational Linguistics*. 2021.