



UTTERANCES	ANNOTATIONS
USER: Hey, could you show me some of your more expensive jackets?	REQUEST:GET, slots: {type: jacket, price: expensive}
ASSISTANT: Sure, the grey and white jacket in the bottom left corner of the wall, the light blue jacket in the middle on the front table, and the grey jacket directly behind it are all in the expensive price range.	INFORM:GET, objects: [11, 8, 0]
USER: What about something with good customer reviews in a size M?	INFORM:REFINE, slots: {type: jacket, price: expensive, size: M, customerReview: good}
ASSISTANT: What do you think of the grey jacket to the right of the light blue one I mentioned? It is size M and has good customer reviews.	INFORM:GET, objects: [9]
USER: I might do that one, but let's see some shirts first. Which of those on the wall do you think would best match my wardrobe?	REQUEST:GET, slots: {type: shirt}
ASSISTANT: How do you like the grey shirt in the middle on the top row?	INFORM:GET, objects: [3]
USER: Actually, I changed my mind, I want to look at jackets again. Do you have anything else similar to the two middle jackets on the tables?	INFORM:GET_SIMILAR, slots: {type: jacket}, objects [0, 8]
ASSISTANT: Sure, the black jacket two garments from the right on the top row of the wall and the red and white jacket four garments from the right are both similar.	INFORM:GET, objects: [4, 0, 2]
USER: I think the grey one on the back table is my favorite of everything you've shown me, put that one in my cart.	REQUEST:ADD_TO_CART, objects: [0]
ASSISTANT: Sure thing, I've added it to your cart.	CONFIRM:ADD_TO_CART, objects: [0]

Figure 7: **Dataset Example.** Dialog labels include intent, slots, and multimodal coreferences. The bounding boxes for the objects mentioned in the dialog (and their corresponding object IDs) are shown for readability.