

Table 2: All 43 test cases and associated challenge types.

Query	Personalization	Intent resolution	Device resolution	Persistence	Command chaining
Turn on the tv			✓		
What channel is playing on the tv?			✓		
Turn on the light by the bed			✓		
What is the current phase of the dish washing cycle?			✓		
Lower the volume of the tv by the light			✓		
What is the current temperature of the freezer?			✓		
Is the tv by the credenza on?			✓		
Play something for the kids on the tv by the plant		✓	✓		
Play something funny on the tv by the plant		✓	✓		
Set up lights for dinner		✓	✓		
Set the lights in the bedroom to a cozy setting	✓	✓	✓		
Darken the entire house		✓	✓		✓
Turn off all the lights that are dim			✓		✓
I am getting a call, adjust the volume of the tv		✓	✓		
Dishes are too greasy set an appropriate mode in the dishwasher.		✓			
Put something informative on the tv by the plant.	✓	✓	✓		
Change the lights of the house to represent my favourite hockey team. use the lights by the tv, the dining room and the fireplace.	✓	✓	✓		✓
I am going to sleep. change the bedroom light accordingly.	✓	✓	✓		
Turn off all the tvs and switch on the fireplace light			✓		✓

Continuing to the next page