Language Commands for a Virtual Agent

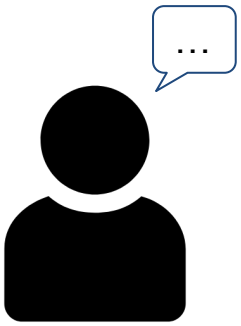
Minh Ngoc Vu
Mentor: *Dr. Jesse Thomason*
GLAMOR lab
[http://glamor.rocks/]

INTRODUCTION

- If you hear someone ask you to “Go to the shrink and pick up the pug” in the kitchen, you can reason that they probably said “sink” and “mug” based on scene context.

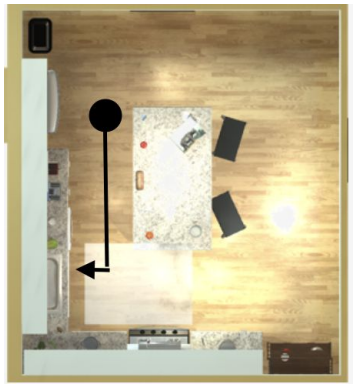
- We’d like to create a virtual agent that can make these corrections and execute spoken commands.

spoken commands



processing…

go to the sink

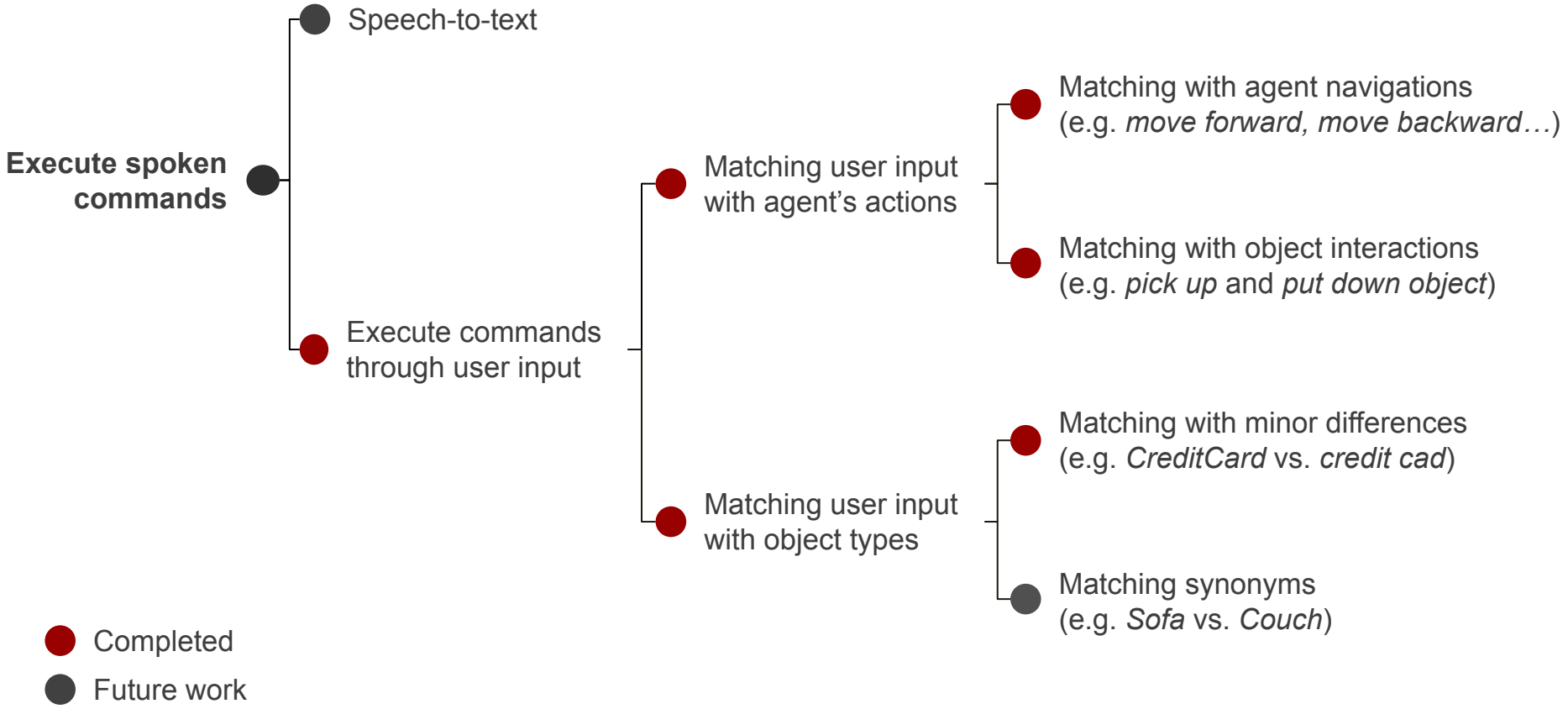


pick up the mug

RESEARCH QUESTION:

- Can we build a command understanding system that recovers from ASR and typographical errors?
  (e.g., “Go to the shrink” → “Go to the sink”
  “Pick up the fjork” → “Pick up the fork”)

RESEARCH PROGRESS



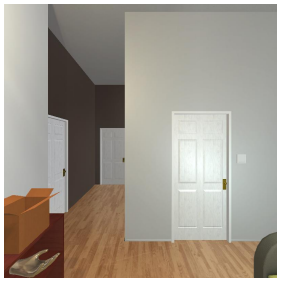
AGENT NAVIGATION

Input: *Move forward*

Input: *Move backward*

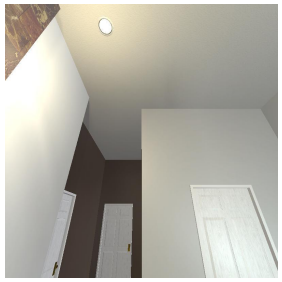
Input: *Look down*

Input: *Move right*



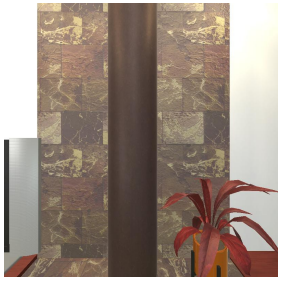
Input: *Move left*

Input: *Look up*



Input: *Turn right*

Input: *Turn left*



Input: *Turn right and turn right*

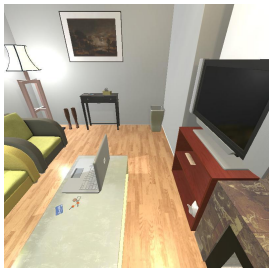
AGENT - OBJECT INTERACTION

Types of Objects Interactions:

- Pick up X
- Put down X
- Put X on Y

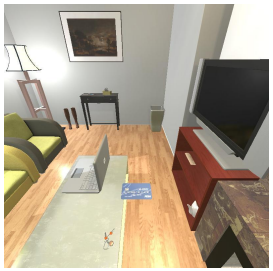
(X and Y are objects)

Input: *Pick up the keychain*

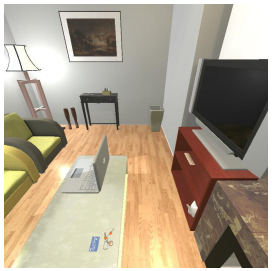


Input: *Put down the keychain*

Input: *Pick up the credit card*



Input: *Put the credit card on the coffee table*



CURRENT TASK:

“Go to” action
- takes the agent to a specific location:

- Get object’s position in the environment
- Get positions that the agent can reach
- Create a navigation graph that takes the agent to the target in possible ways.

FUTURE TASKS:

Using A* Search algorithm to have the agent calculate the shortest path

Speech recognition

Execute spoken commands

Matching synonyms

PROGRESS

Problems:

- Find the position of object in sentence.
- Help the virtual agent understand spelling mistake.
- The virtual agent needs to calculate the shortest path.

Solutions:

- Template-based semantic parsing.
- String matching using Levenshtein distance.
- Create a navigation graph using NetworkX Directed Graph.
- Using A* Search Algorithm to find the shortest path *(future work)*

CONCLUSION

- **Achieved Goal**: The virtual agent is able to be controlled through user input.
- **Future Goal:** Implement speech recognition to control the virtual agent through spoken commands.

ACKNOWLEDGEMENTS

- Dr. Jesse Thomason
- Robotics REU - NSF
- GLAMOR lab