Analyzing Robot Adaptation in New Tasks and Environments

Ruth Berkun, ruth3usa@gmail.com
Arcadia High School, Class of 2022
USC Viterbi Department of Robotic Engineering, SHINE 2020

Introduction

To advance robots, engineers strive to make robots more adaptable to the kinds of real-life environments humans often encounter. In a kitchen environment, for example, a robot may not have access to all the tools it would expect to use, and it may be commanded to carry out tasks it has never seen implemented before. Our research aims to develop an algorithm to choose the best available tool for a given task (ex: "cut banana"). My role in all of this is to generate training data from online datasets and human feedback so the robot can learn what tools are best for certain types of actions.

Objective and Impact of Research

Our research will help develop robots to assist humans in a cooking environment, where certain tools are not available and where commands without mention of a tool (ex: "cut the banana") are often used by humans. Eventually, we hope our research can be expanded to help other robots generalize learned manipulation action plans to new tasks and environments.

Hard Skills Used/Learned

- Python
  - Basics (strings, ints, dictionaries, lists, lopps, functions, searching)
  - Used Anaconda and Jupyter Notebook
  - Reading and writing JSON files
  - Libraries (NLTK, JSON, tqdn, math, re, pandas)
- Confusion matrices
- Google Forms and Sheets API

What I Did Part 1: Reading Recipe Dataset

First, I had to extract three words combinations ("**trigrams**") from a dataset of sentences from recipes. The first word would be a tool, the second word would be an action, and the third word would be an object.

For example, given the following lists of
Tools: knife, spoon, fork
Actions: cut, stir
Objects: meat, banana, soup, milk
some possible trigrams include
knife cut meat
spoon stir soup
etc.
Given the sentence: *"Put the meat on a chopping board and cut it with a knife"*
we could extract the trigram "knife cut meat", because knife is one of our tools, cut is one of our actions, and meat is one of our objects, and all three words appear in the sentence.

```
1  {"knife cut meat": 352, "spoon mix soup": 2, "oven bake butt
   heat flour": 178, "stove heat sauce": 110, "bowl stir beef":
   bake fish": 86, "cup stir lemon": 331, "cup stir lemon juice
   oil": 222, "plastic wrap wrap dough": 58, "pan heat sauce":
   tomato": 459, "cup stir water": 922, "cup stir red pepper":
```

First few lines of output

I used the NLTK (Natural Language Toolkit) library to make sure I didn't skip verbs and nouns that appeared in different forms in the sentences than they did in the lists (ex: "cutting" would count as an instance of "cut").
In my program, I used a dataset, which was in the form of a single JSON object containing over 1 million recipes, as input, and returned a JSON file with a dictionary containing each trigram found and how many times each trigram was found.

What I Did Part 2: Collecting and Interpreting Human Feedback

My user study contained 2 types of questions in a Google Form, and was sent to 12 people.

What is the best tool to (verb) (object)?
a. tool 1
b. tool 2
c. tool 3
d. tool 4

For this type of question I used the python pandas library to make a confusion matrix in table form comparing the tools humans chose and the tools the robot chose:

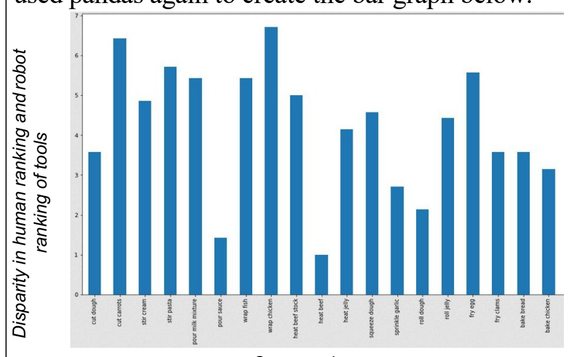
Robot choice

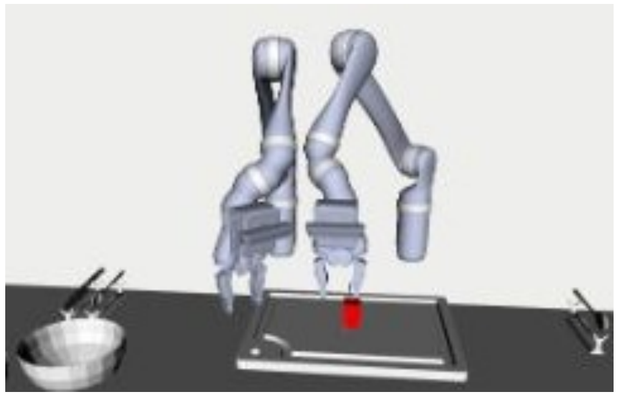
Human choice	knife	food processor	pizza cutter	spoon
knife	7	0	0	4
food processor	0	0	0	0
pizza cutter	0	0	0	0
spoon	0	0	0	0

Part of the tool confusion matrix (the entire matrix is 31x31)

Rank the following tools (1=best, 4=worst) in terms of how useful they are for (verb) (object).
a. tool 1
b. tool 2
c. tool 3
d. tool 4

To calculate disparity between human and robot rankings, I summed the absolute value of the difference of rankings for each of the four tools, then used pandas again to create the bar graph below:





3D Simulation of robot picking up a knife

References and Relation to STEM coursework

The previous study that this study is largely built off of:
Zhang, H., & Nikolaidis, S. (2020). Robot Learning and Execution of Collaborative Manipulation Plans from YouTube Cooking Videos. Retrieved from https://arxiv.org/abs/1911.10686

This relates to my past STEM CS coursework from APCSA, building off of basic programming principles such as searching and sorting. What I learned this summer about computer memory, APIs, and different libraries will help me create more diverse and effective programs in the future.

Acknowledgements

I would like to thank Professor Nikolaidis, my PhD mentor Hejia Zhang, and my SHINE partner Nikitas Klapsis for helping me in my research journey. I also want to thank my center mentor, Ashley, and Dr. Mills for checking in with me and making sure I wasn't confused about SHINE. Lastly, I would like to thank the rest of the SHINE team and participants -- without you, none of this would have been possible.