Harrison

A Helpful Archeology Interpreter

Final Project Submission for 15-494 Cognitive Robotics
By Brayden Kohler (braydenk)

The Problem

Stratigraphy is a complicated thing. It requires you to keep mapping out relative positions of objects at multiple layers throughout your site and noting in what ways they relate to each other.

Further, synthesizing all of your collected data into a nice interactable page and making a proposed report of what may have happened is a task that can take some time.

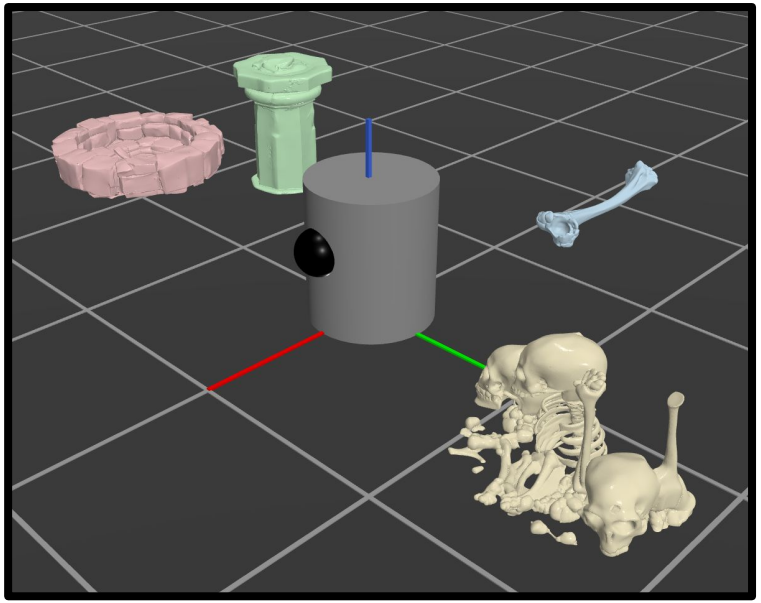
What Harrison Does

Harrison does each of those things for us:

1) Harrison surveys the site, building an internal world model at multiple levels in the site.

2) Harrison moves in the direction of the largest gap to avoid hitting objects and stores positional relationships between artifacts and archeological notes by interacting with GPT to reason about the site.

3) Harrison builds an interactable Harris matrix DAG and makes a final report about what he thinks happened at this site.



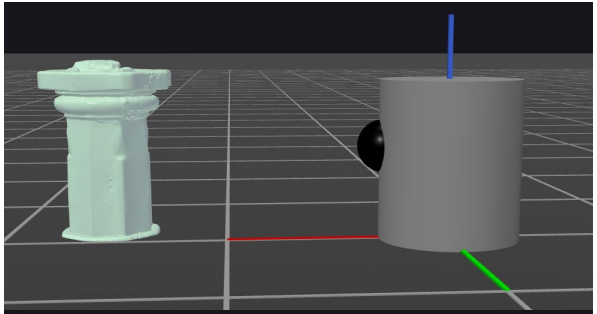
Demo: https://youtu.be/1ekp3fuSnME

The Layered World Map (Most Interesting)

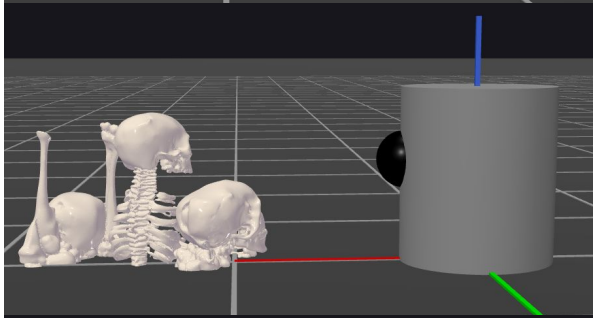
The robot sometimes makes poor estimates of objects far away and of april tags facing awkward directions or moves objects it thinks are missing.

Based on the reasonable assumption that a layer of the site doesn't change, we maintain an extra internal set of world models for each layer and use it to update the robot's true world model to have only the most accurate estimates and be able to switch between multiple contexts very easily

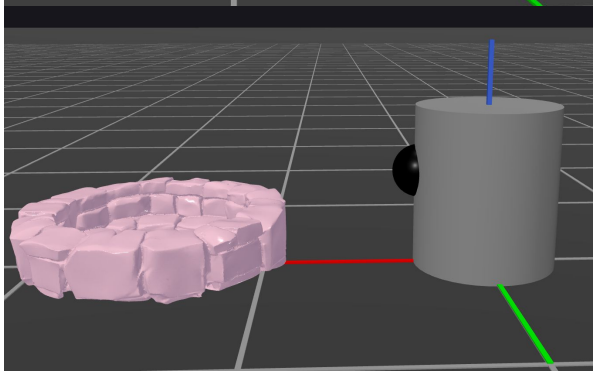
Layer 0 :



Layer -1 :



Layer -2 :



What Worked

- Few-Shot Learning helped make GPT ask good questions in general
- Finding directions to move that minimize the likelihood of knocking a fragile artifact, by moving towards large gaps in its world map

What Didn't Work

- GPT sometimes struggles with being aware of what object relationships it should ask about.
- Complicated survey techniques where the robot would move close to objects to take a good look at it.
- Surveying quickly

Future Extensions

Make the dialogue possible to be more oral and conversational.

Automatically combine position data and user answers to reduce number of questions asked.

Save the layered world map to visualize it all at once in 3D later on

Instead of using AprilTags to represent objects, create actual small models of artifacts, and use computer vision to not just determine position of objects, but also features about the object itself (color, material, etc.)