HW 2 Writeup

Kennedy Ellison and Enora Rice

What we did in the homework:

We used babbage to train our models due to cost constraints. Our theme was greek mythology so our starting location was Mount Olympus. Everything else was generated by the GPT-3!

Implementation notes:

- We did not include the optional connection parameter for the `get_connections` function. For our game implementation it was not needed.
- We changed the `get_item_property` functions to have the prompts include the property name. When we used the property description as the prompt input the model would often just pick a random property to use instead of predicting true or false on the property that is referenced by the property description. Using the property name instead, the model predicts for the property given.

Feasibility of fully generating text adventure games?

From this homework, it seems like this could be possible, especially with small simple games maybe made for children. There are still times where the phrasing of the generated text is odd or not connected well to the other rooms or the theme. Generally, it is surprisingly good. I think the game could be improved by more continuity in the finetuning process so that each model could benefit from the data used to train each other model. In addition, using the more costly models like davinci, would definitely improve the believability of the game. Lastly, though connections in our game only go in one direction because the model was not trained on the game format in any way. It might be interesting to train GPT-3 on human made games in the light_environment data format to see what it came up with.

Evaluation:

```
Recall for is_gettable is:  0.8307692307692308
Precision for is_gettable is:  0.864
Recall for is_weapon is:  0.6428571428571429
Precision for is_weapon is:  0.7105263157894737
Recall for is_surface is:  0.46938775510204084
Precision for is_surface is:  0.6764705882352942
Recall for is_container is:  0.625
Precision for is_container is:  0.5882352941176471
Recall for is_wearable is:  0.7142857142857143
```

```
Precision for is_wearable is:  0.8333333333333334
Recall for is_drink is:  0.36363636363636365
Precision for is_drink is:  0.6666666666666666
Recall for is_food is:  0.35714285714285715
Precision for is_food is:  0.7142857142857143
```