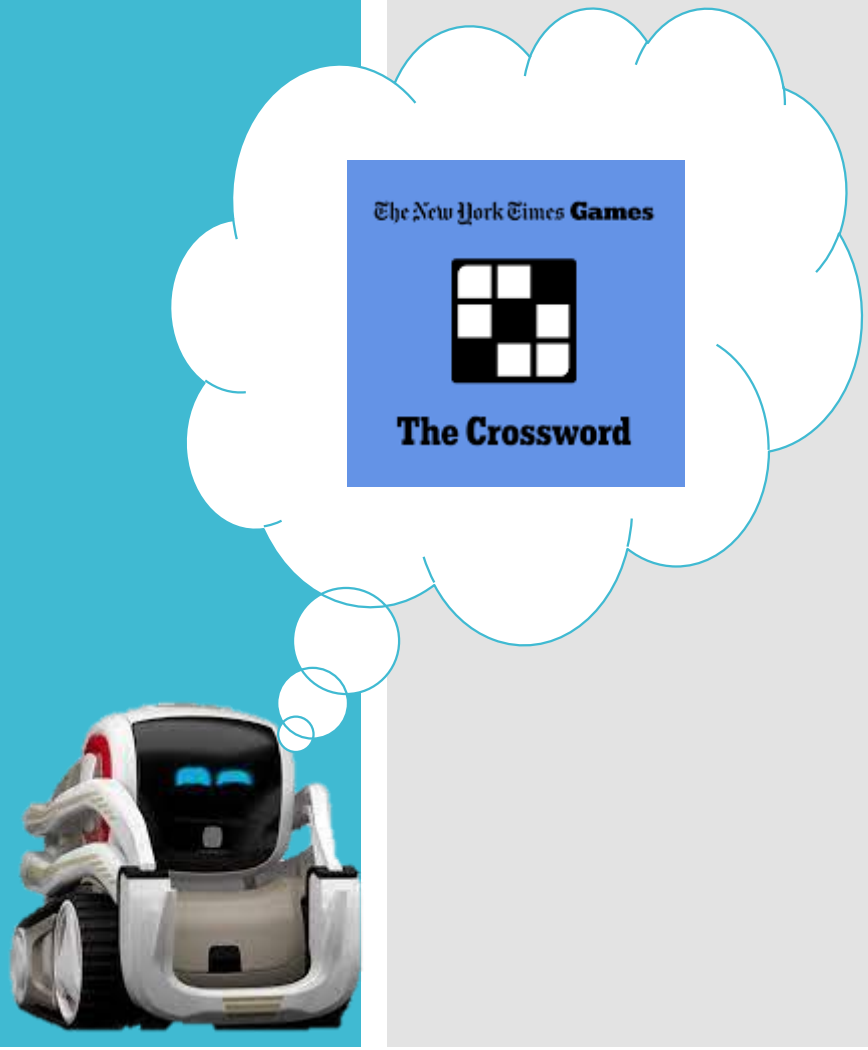
# Cozmo's Crossword

Cognitive Robotics Final Project

Hugo Martinez



# The Problem?

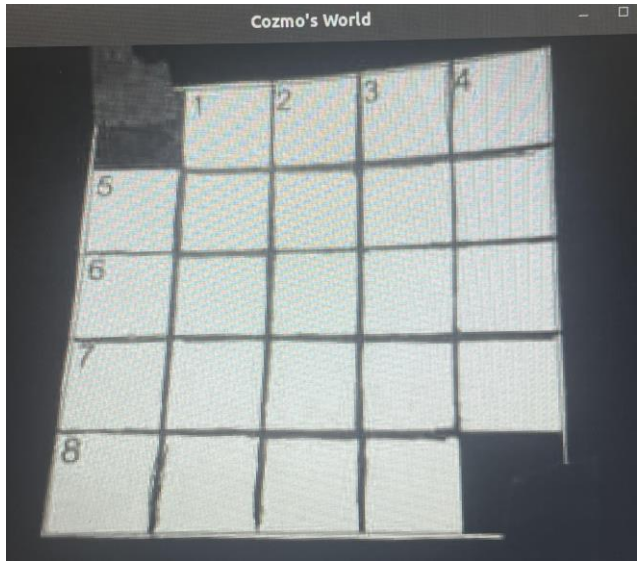
- Solving the crossword can be too hard sometimes.
- Googling solutions, is too easy.
- I am currently obsessed with the crossword and could not think of another topic.

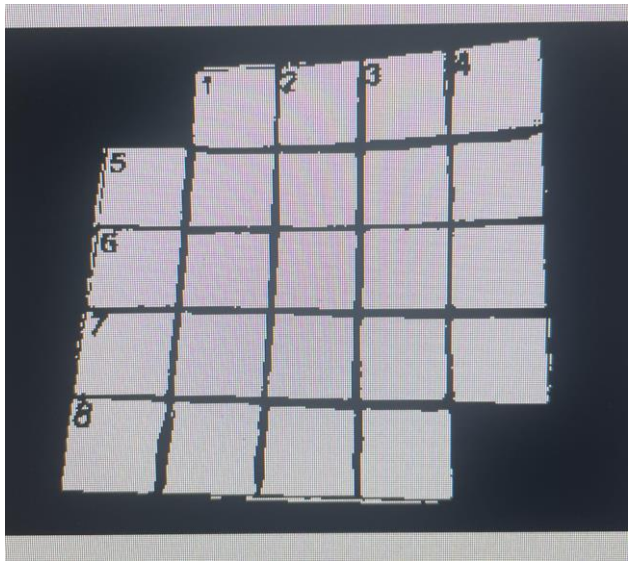
# What does it take for Cozmo to solve a crossword?

- Step 1: Read the puzzle board:
  - Cozmo needs to be able to see some physical crossword board and "read" it.
- Step 2: Understand and remember the puzzle:
  - This means Cozmo needs to understand things like where the clues are and how long each clue is (i.e. what does 1-Down mean and how long is the answer)
  - Cozmo needs to remember what we have answered so far and how that impacts future clues
- Step 3: Listen for clues and solve them:
  - We want Cozmo to be listening to the user and respond with an answer to a provided clue

# Step 1: Computer Vision

- Cozmo starts by looking at a printed out, blank crossword puzzle, preferably against a dark background
- Cozmo uses OpenCV's connected components (with stats) tool to look for the number of white squares it can find
  - With a good background, Cozmo needs no other information to identify the squares (and can parse any sized crossword). However, in a more realistic scenario, I have coded in some minimum and maximum area requirements in order to make sure that Cozmo is able to identify the squares effectively
- Using statistics from the connected components, we recreate the board in python



# Step 2: Python algorithms

- The process of making Cozmo understand the board was all just python
- The method by which clues are numbered is standardized, and I used this method to allow Cozmo to number all of the clue names and locations on his own board representation
  - This just means that Cozmo fully understand references to things like 2-Down or 4-Across)
- Additionally, Cozmo will be storing the board, as well as any words that you decide to keep in his representation of the board.
  - This lets Cozmo have all of the necessary information available about the clue, like length and known letters
  - This means that all the user has to do is read off clues, making it a significantly more "realistic" experience

# Step 3: Hear transition and ChatGPT

- Using Cozmo's built in Hear transition, I am listening for any response, and processing the input string to see if it is a valid clue for the board and do some slight normalization.
  - I have also built in the choice to either say the clues verbally or use text messages
- To actually respond, Cozmo first checks his board for all of the info that he can gather about the clue we are looking at (length, known letters) and makes a small string with information
- Then we send the clue we got from the user as well as the known information about that clue to ChatGPT and get back a response from there.
- Cozmo then speaks out the clue to the user, asking for their opinion on whether or not it should be kept.

# The preamble sent to ChatGPT

```
preamble = ''' You and I will be solving a crossword puzzle, I will provide you
        information about the crossword clues in the following format:
        Clue: (This will be a crossword style clue)
        Information: (This will be in the format of n spots, where n is
        the length of the word. Each spot will either be a lettter
        or an underscore. If it is a letter, that is a letter that we know
        is already in the word. If it is an underscore, it could be any
        letter)

        After the clue and information is provided, please provide a
        possible solution. Always provide a guess, even if there is not a
        lot of information.

        When you respond, only send back the answer itself in all caps.

        Here is an example prompt:
        Clue: Studied carefully, with "over"
        Information: P _ _ _ _

        Example:
        PORED '''
```

# Results

- Generally went well!
- Strengths:
  - ChatGPT is good, but not too good at solving clues
    - It does improve with more information given which is exactly what I had hoped it would do
  - When the computer vision goes right, the background python as well as the state machine worked well
- Weaknesses:
  - The camera
    - Had to pivot to printed picture with a dark background and very thick lines
  - Error handling
    - I did not get too much time to make sure Cozmo handles errors effectively and naturally
  - Hearing is a bit spotty
    - Was unable to listen for more specific queues, having to work with the raw string is a little difficult

# Future Extension

- Working with a computer screen
  - Ideally, I had wanted to be able to show Cozmo the NYT Crosswords from my phone, as that was the most realistic way it would happen if I really had a little crossword solving robot with me at all times
- Identifying the puzzle itself
  - I had wanted to do something along the lines of shape matching to have Cozmo be able to pick out the crossword puzzle (like in a newspaper)
- More natural ChatGPT responses
  - I wanted to let ChatGPT have a little bit more free reign with its responses, to really make it be like two friends chatting
- Take user input for possible crossword solves
  - This was the biggest miss in my project, I really wanted to have it go back and forth but at times it was difficult to just get Cozmo to hear correctly and I just did not have a chance to fully implement this
- Add a celebration at the end
  - I realized at the last second that Cozmo should do a little animation when you finish solving but couldn't get it to work