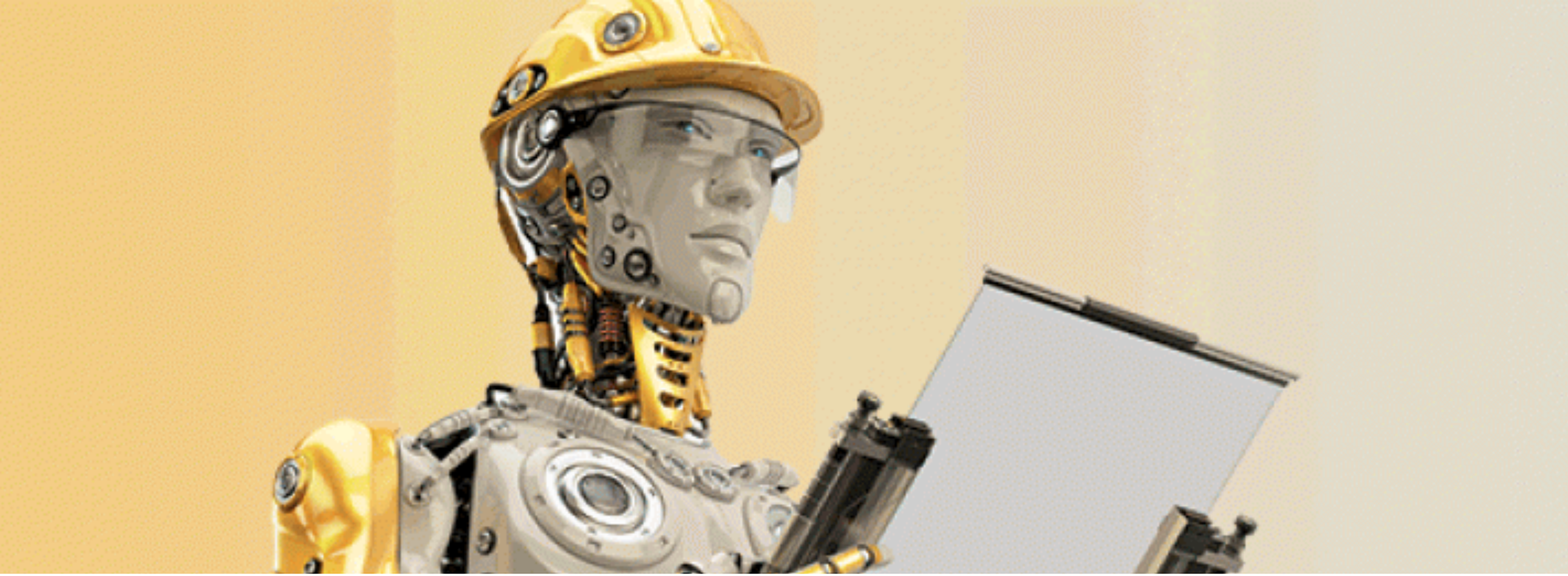


VRND CAPSTONE – ROBO REPLACEMENTS

Eric Zavesky - May-June 2018

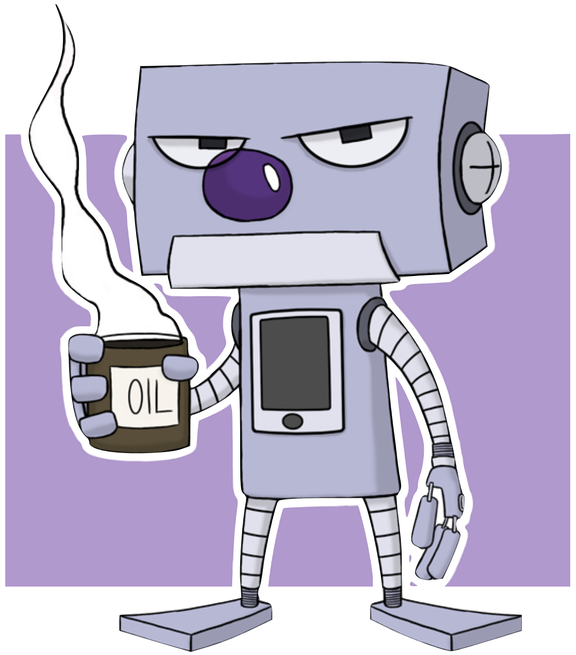
OVERVIEW

- This capstone will explore the emotive side of robotic automation -- mostly focused on humor, curiosity, and surprise!
- The player's task will be to teach a robot (upper half of a robot avatar) how to accomplish a few simple tasks by demonstration and vocal commands.
  - Doing/Listening, Seeing, Speaking
- The focus of this application will be the gameplay and robot interaction, as opposed to extensive graphical interactions.

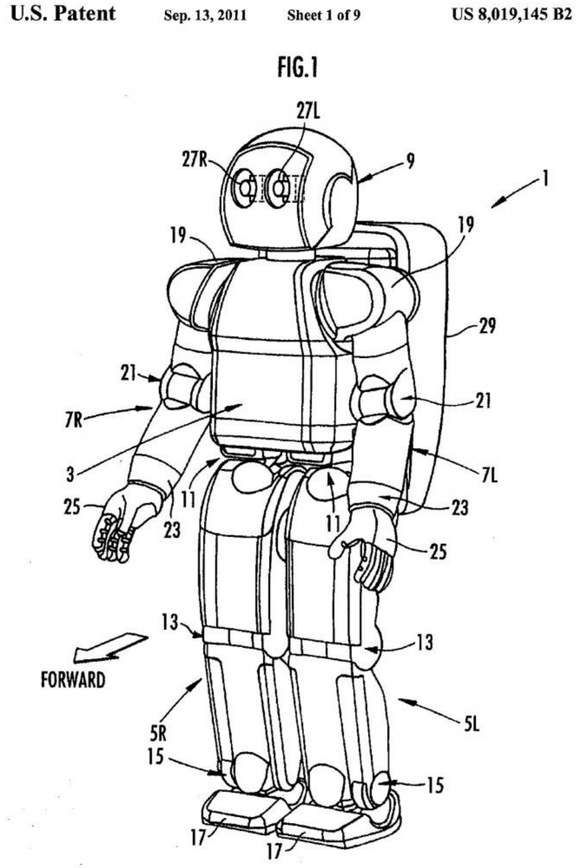
Image: http://www.mhlnews.com/technology-automation/automation-education-training-first-generation-automated-robotics-operators

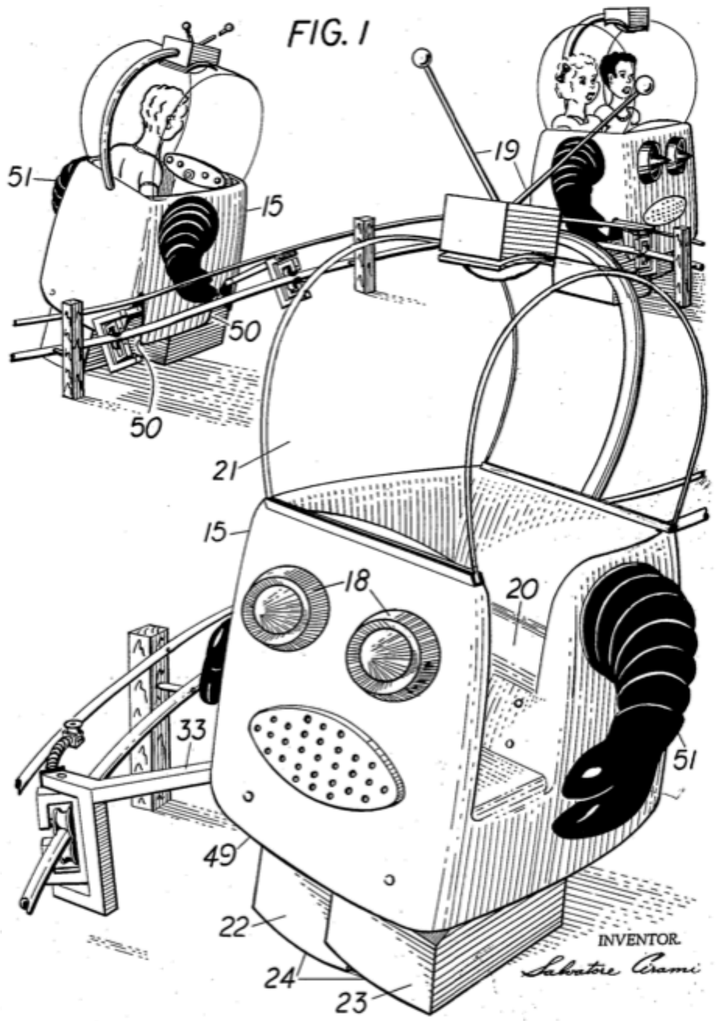
GAMEPLAY

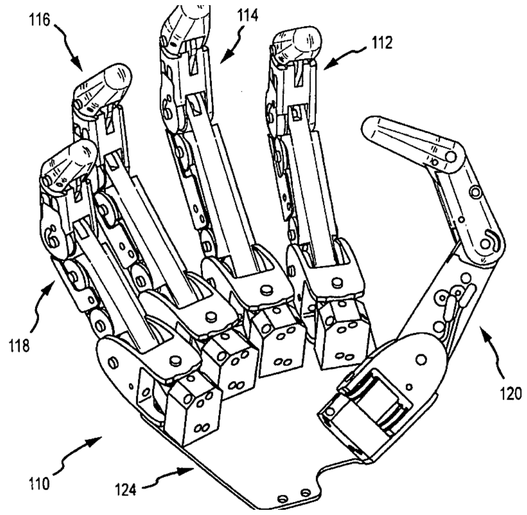
- Task Review
  - Doing/Listening, Seeing, Speaking
  - Robot interjects satirical, comical commentary, a la dialog from games like "The Lab" or "RoboRecall"

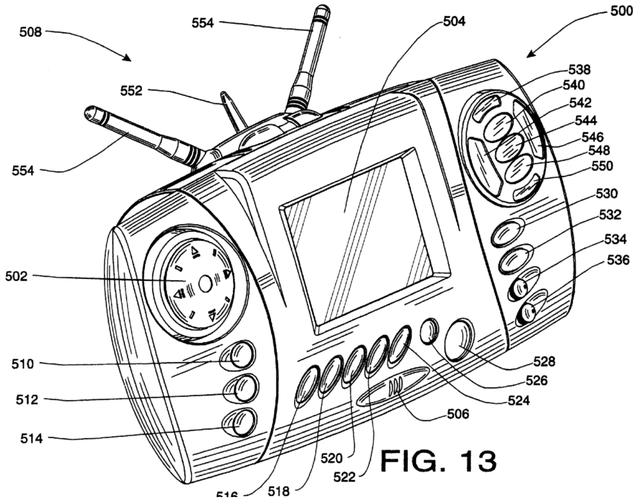


- Doing Task
  - goal: using simple interaction mechanics, move to make a sandwich
  - Robot copies user movements
  - User orders task
- Seeing Task
  - goal: teach the robot how to see different items (house, car, flower)
  - user places item on table, robot classifies them online (AI/ML component)
- Listening/Speaking Task
  - goal: user helps robot with speech recognition to complete a "MadLib"
  - ideally pre-programmed animations will bring in media that adjusts according to a few options presented to the user

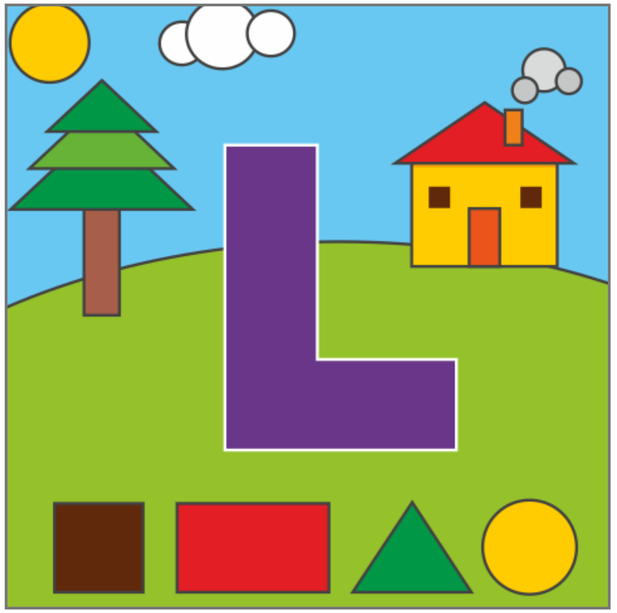








IMPLEMENTATION BREAKDOWN

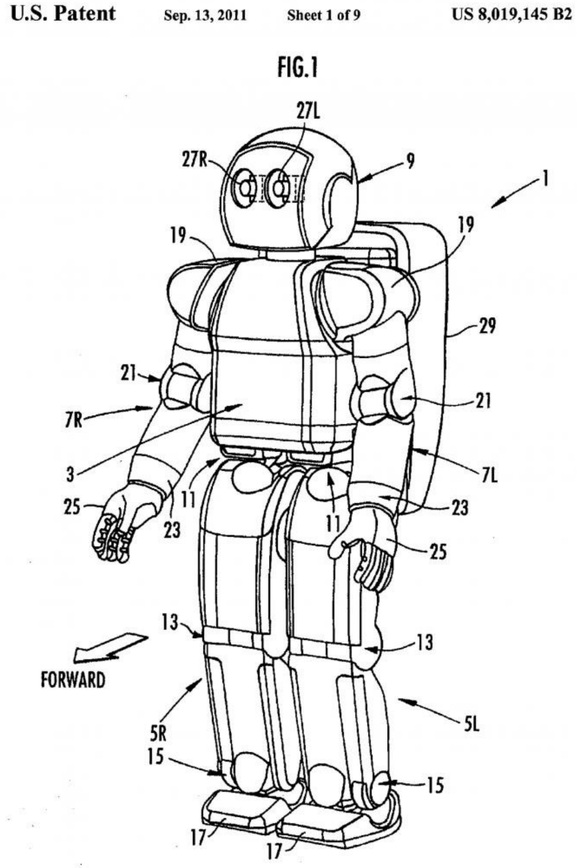


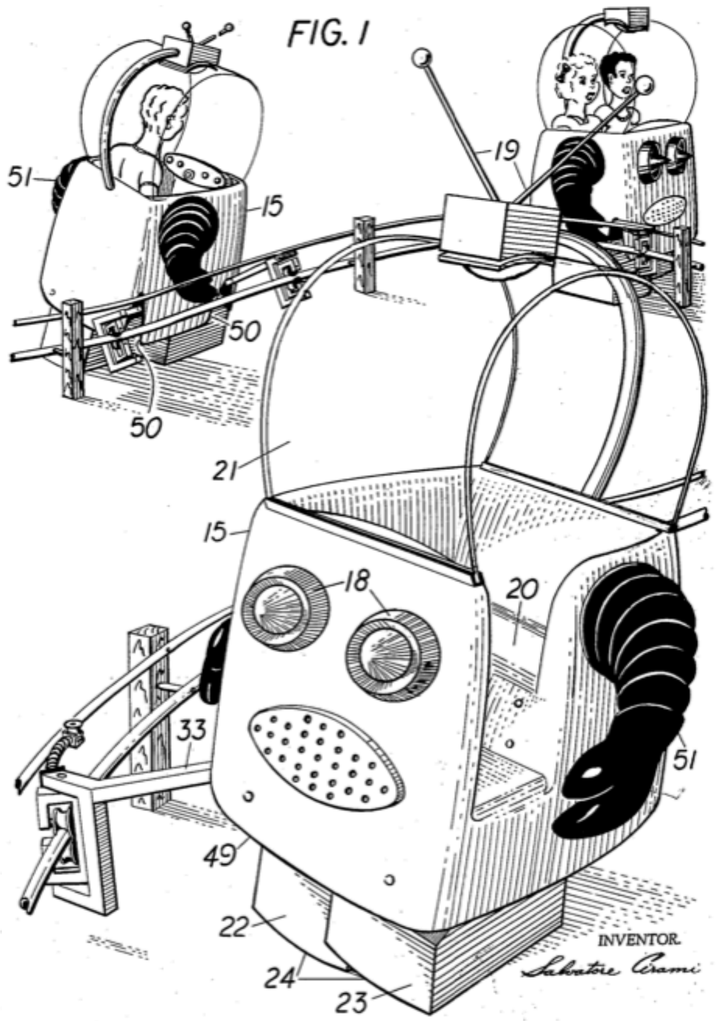
FEATURES AND DEPENDENCIES: MODELS & ANIMATIONS

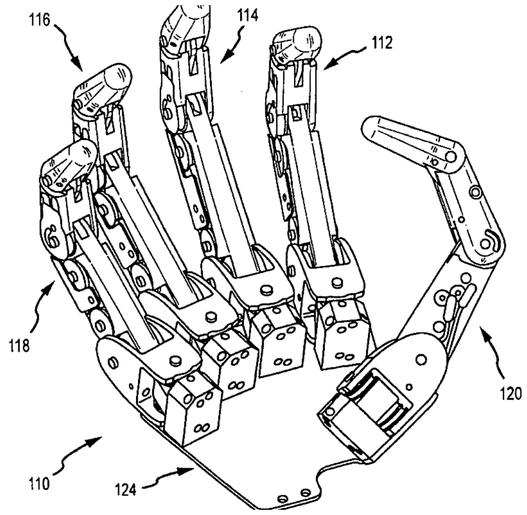
- Robotic avatar
- Doing task
  - food models: bread, jelly, butter
  - tool models: plate, knife
  - Simple kitchen environment
    - possibly reusing models from Udacity early models and starter kits
- Seeing task
  - simple shapes created in Unity
  - simple camera or simple office setting
- Listening task
  - news desk + image/video assets
  - animations for a few tasks to present objects

FEATURES AND DEPENDENCIES: SOFTWARE & PROGRAMMING

- Text to speech with funny robot voice
- Doing
  - Link kinematics between robot hand and human hand for testing
  - Simple memory/copier for task
  - Ordering process for task
- Seeing
  - Capability to create shapes and placement
  - Simple nearest neighbor for placement using shape and pallets placement
- Listening/Speaking
  - Create a few words
  - Create recognition model constrained to a few words
  - Compose text as output







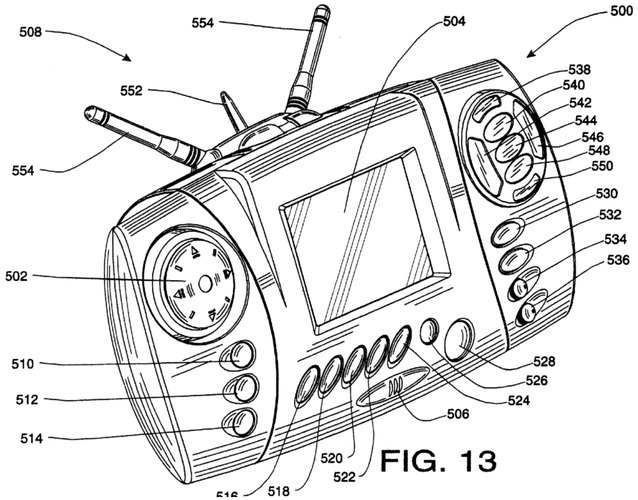


FIG. 13

SOFTWARE & INTERACTIONS

ELIMINATED FOR SCOPE 7/10

FEATURES AND DEPENDENCIES: TEXT TO SPEECH

- Priority: Text Bubbles in OSD
- Achievable: Recorded voice + audio FX with funny robot voice
- Stretch: an automated TTS generation with personality

- http://www.acapela-group.com/
  - would require removal of background
- https://www.cepstral.com/en/demos
  - simple demo, apply dizzy droid or other audio effect for robot sound
- Mac "speak" tool plus a visual effect
  - https://www.youtube.com/watch?v=j_FvFAdlbwo
  - http://www.asktoby.com/#killerringer
- https://assetstore.unity.com/packages/tools/audio/speech-auto-detector-86008
- https://forum.unity.com/threads/rt-voice-run-time-text-to-speech-solution.340046/

FEATURES AND DEPENDENCIES: SPEECH RECOGNITION

ELIMINATED FOR SCOPE 7/12

- Priority: Just using buttons for input
- Stretch: Full speech recognition through Google or Watson

- https://bitbucket.org/Unity-Technologies/speech-to-text
  - https://stackoverflow.com/questions/39611728/how-to-add-speech-recognition-to-unity-project#39613264
- https://assetstore.unity.com/packages/templates/tutorials/vr-watson-speech-sandbox-114015
- https://assetstore.unity.com/packages/tools/ai/ibm-watson-sdk-for-unity-108831
-

PARTIALLY ELIMINATED FOR SCOPE 7/12

FEATURES AND DEPENDENCIES: MACHINE LEARNING

- Priority: Simple shape + KNN classification
- Stretch: (this is a fun requirement and will only be skipped under duress)

- https://github.com/nvnhcmus/kNN - knn method
- https://github.com/pavl0v/RandomForest - simple random forest method
- http://accord-framework.net/docs/html/R_Project_Accord_NET.htm
- http://accord-framework.net/

FEATURES AND DEPENDENCIES: DOING TASK IK SOFTWARE

REDUCED SCOPE 7/10

- Priority: Simple IK following of user and objects
- Stretch: Tasks out of order, different sizes, etc; also more complicated ordering requirements (instead of linearly asking for user help)

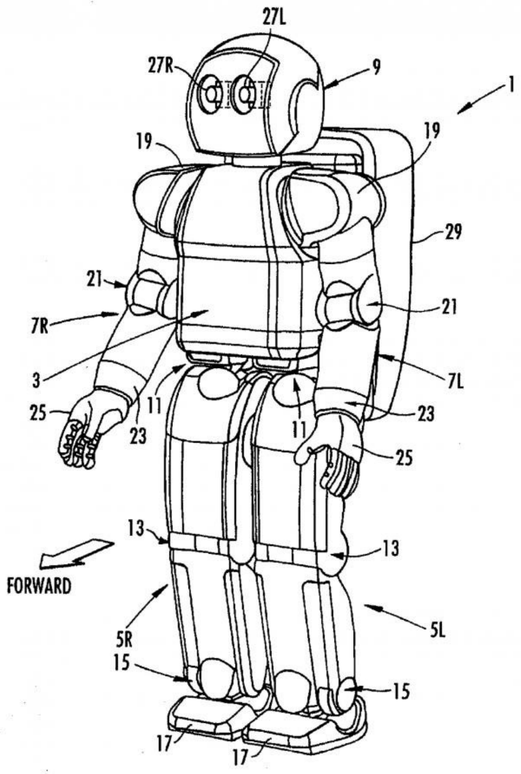
- Discovery of bound avatar with desired actions (e.g. full IK+bones)
- Create method for "doing" task
  - Robot will follow user object with head when something is held
  - Backend script tracks which object is grabbed
    - ~~Backend script captures position in update as driver for robot behavior~~
  - Robot follows own grabbed object when manipulating
- Display will update with task, allowing user to go back or forth with grips
- ~~Push button between tasks to record them as done or restart~~
- ~~Push button when task is completed~~
- Realtime computation of similarity based on object position differences

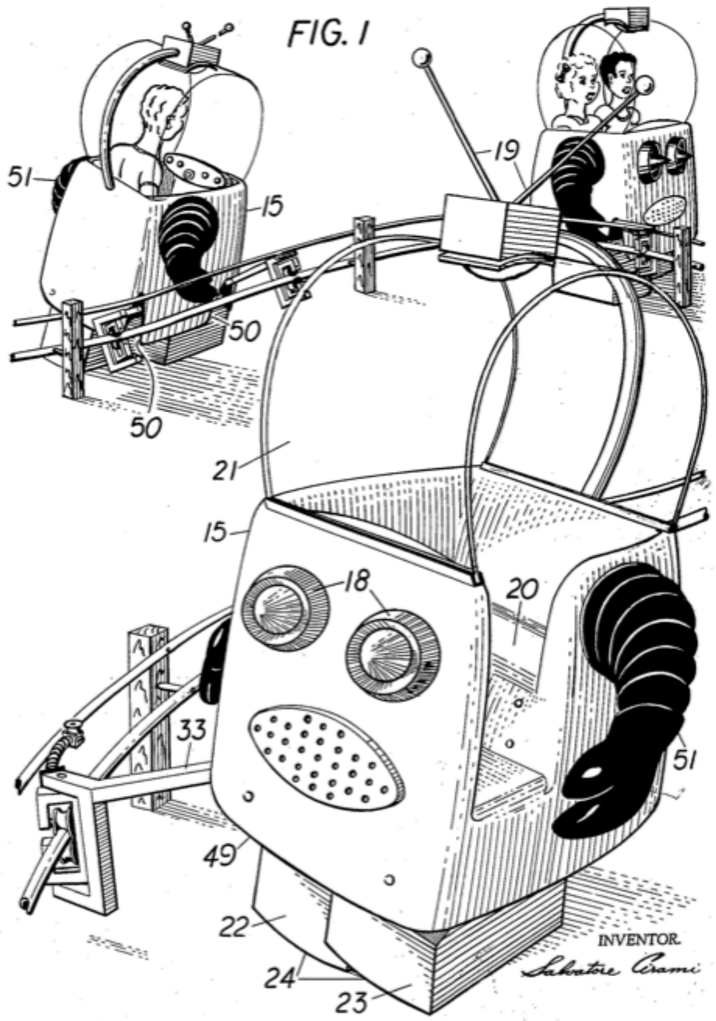
FEATURES AND DEPENDENCIES: LISTENING AND READING

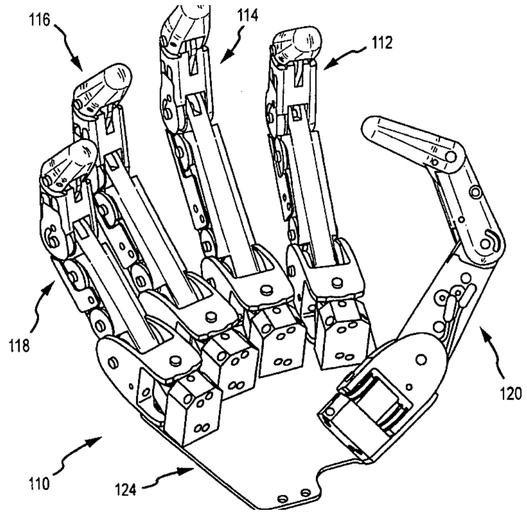
- Priority: OSD for story and text prompts, buttons
- Stretch: Spoken prompts, speech recognition for constrained reply

- Propose the use of animations to accomplish much of robot animation
  - https://assetstore.unity.com/packages/3d/animations/everyday-motion-pack-free-115067
  - CMU motion capture -> FBX (search first, then find which one it is)
    - http://mocap.cs.cmu.edu/search.php
    - https://assetstore.unity.com/packages/3d/animations/huge-fbx-mocap-library-part-2-20282

FIG.1







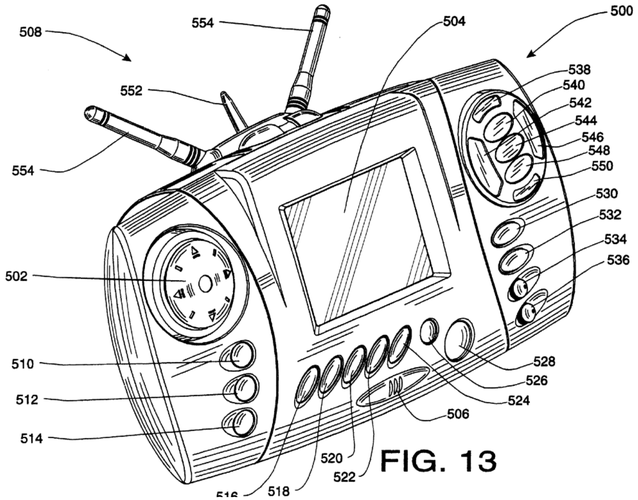
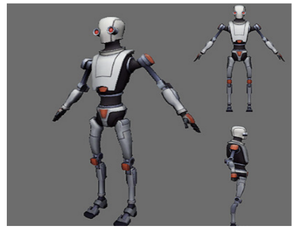


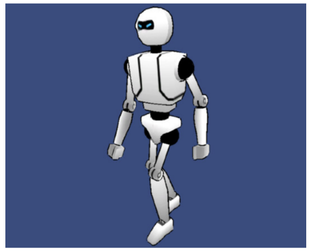
FIG. 13

MODELING

FEATURES AND DEPENDENCIES: KINEMATIC ROBOT

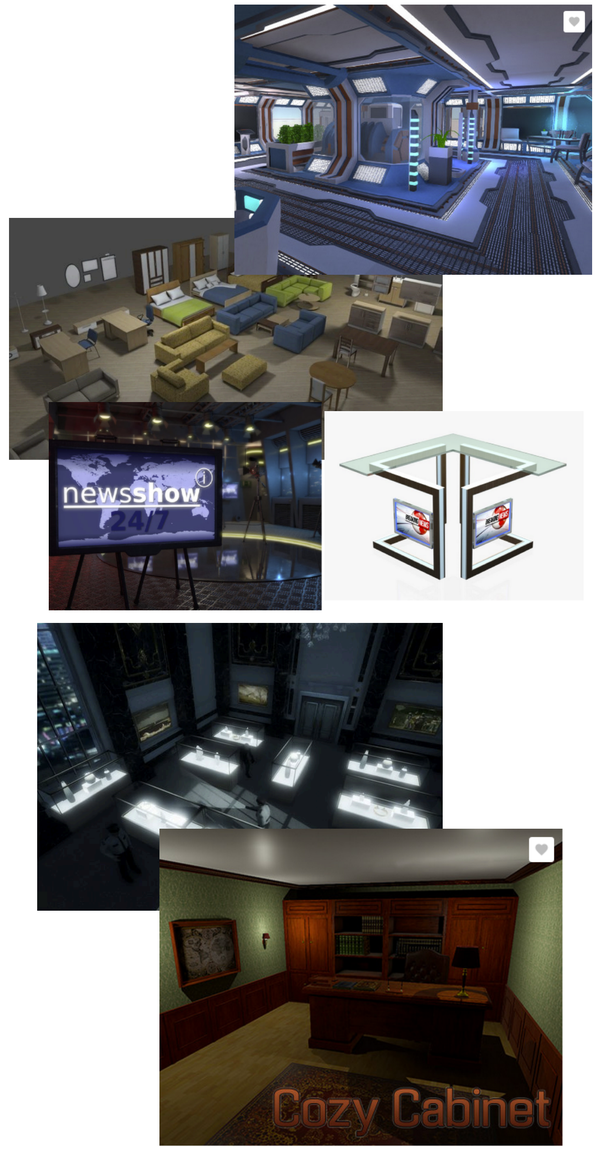
- https://assetstore.unity.com/packages/3d/characters/robots/space-robot-kyle-4696
  - simple kinematic robot
- https://assetstore.unity.com/packages/3d/characters/robots/robot-1-65726
  - larger kinematic robot, may be too scary
- https://assetstore.unity.com/packages/3d/characters/robots/cyber-soldier-52064
- https://assetstore.unity.com/packages/3d/characters/robots/sleek-toon-bot-free-34490
- https://assetstore.unity.com/packages/3d/characters/robots/dummy-animation-64322
- Software IK





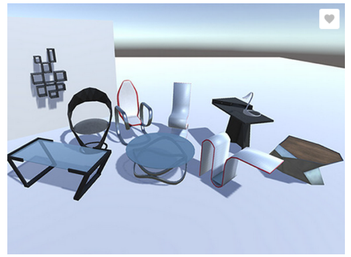
FEATURES AND DEPENDENCIES: ENVIRONMENT MODELS

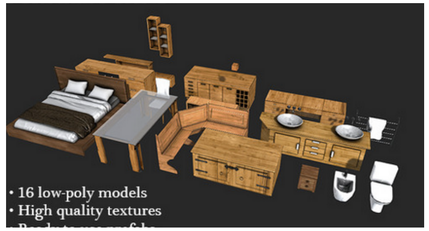
- https://assetstore.unity.com/packages/3d/environments/sci-fi/sci-fi-styled-modular-pack-82913
  - comprehensive space outpost package
- https://assetstore.unity.com/packages/3d/props/furniture/big-furniture-pack-7717
  - some beds, couches, etc
- https://www.turbosquid.com/3d-models/virtual-tv-studio-news-dxf-free/991615
  - https://assetstore.unity.com/packages/3d/environments/virtual-news-studio-91126
  - example of news studio (no model, inspiration)
- https://assetstore.unity.com/packages/templates/tutorials/the-great-fleece-110186
  - museum example
- https://assetstore.unity.com/packages/3d/characters/cozy-cabinet-117482
  - interior "study" room
- Udacity apartment/kitchen model example
- https://assetstore.unity.com/packages/audio/music/absolutely-free-music-4883
  - music!?



FEATURES AND DEPENDENCIES: OBJECT MODELS

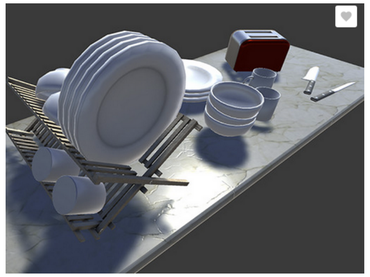
- https://assetstore.unity.com/packages/3d/props/tools/carpenter-s-tools-118471
  - joking tools for cooking
- https://assetstore.unity.com/packages/3d/props/furniture/modern-furniture-pieces-pack-81417
  - additional furniture pieces
- https://assetstore.unity.com/packages/3d/props/furniture/gray-furniture-pack-40580
  - gray furniture
- https://assetstore.unity.com/packages/3d/props/furniture/chalet-style-furniture-31966
  - chalet dark wood furniture
- TV from museum example

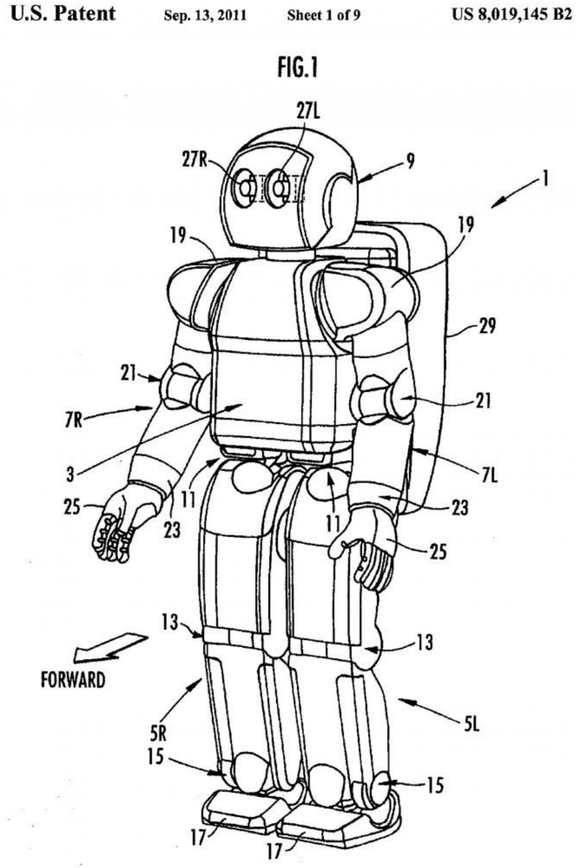


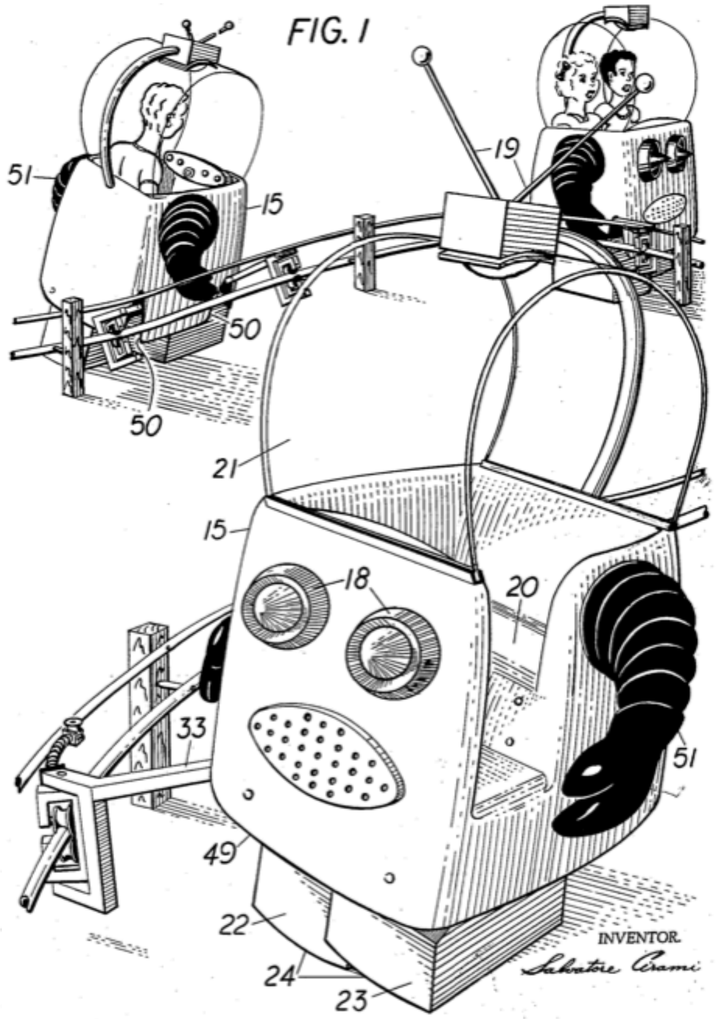


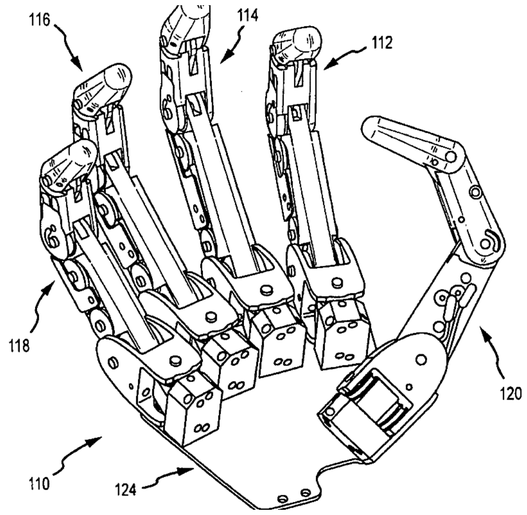
FEATURES AND DEPENDENCIES: OBJECT MODELS

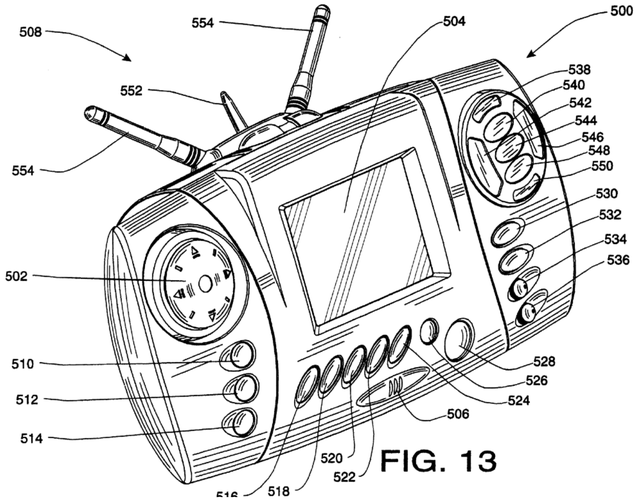
- https://assetstore.unity.com/packages/3d/props/interior/ygs-mugs-96665
  - coffee mugs
- https://assetstore.unity.com/packages/3d/props/interior/kitchen-props-free-80208
  - mugs, plates, toaster
- https://assetstore.unity.com/packages/3d/environments/fantasy/pirate-tavern-113463
  - mugs, plates, etc
- https://assetstore.unity.com/packages/3d/props/interior/props-for-the-classroom-5977
  - classroom objects











USER TESTING AND FEEDBACK

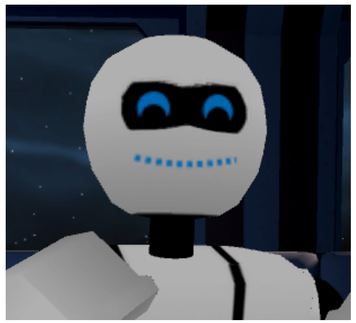
USER PERSONA AND NEEDS

- (child) - ~7 yr old
  - Goals: have fun, learn a little
  - Experience: very little
  - Requirements:
    - Visually interesting for short-attention span
    - Possibly, ability to read prompts

- (adult) - ~30-40 yr old
  - Goals: laugh, quick in+out for game
  - Experience: moderate to high
  - Requirements:
    - Comical
    - Simple but entertaining game play

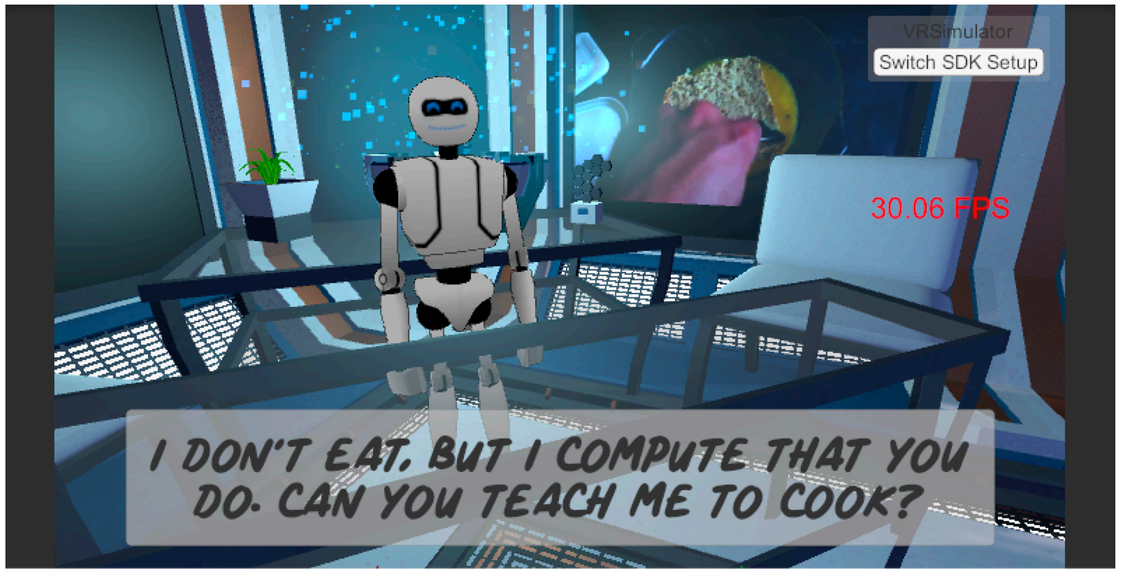
USER TESTING RESULTS (RECORDED IN TIME ORDER)

- 7/6 - Child was tested with several robot appearances. She choose the white "toon" robot for primary use.
- Child user preferred "happy" toon face, so minor material updates were made to source robot.
- Child wanted to let each robot have a different name. This modification can be evaluated as a canvas on the robots, but will likely get scoped out due to time.



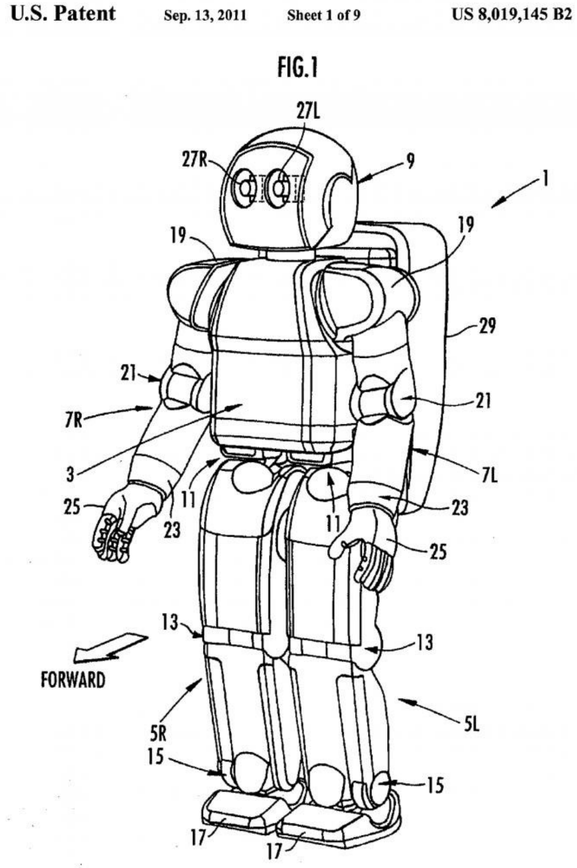
USER TESTING RESULTS (RECORDED IN TIME ORDER)

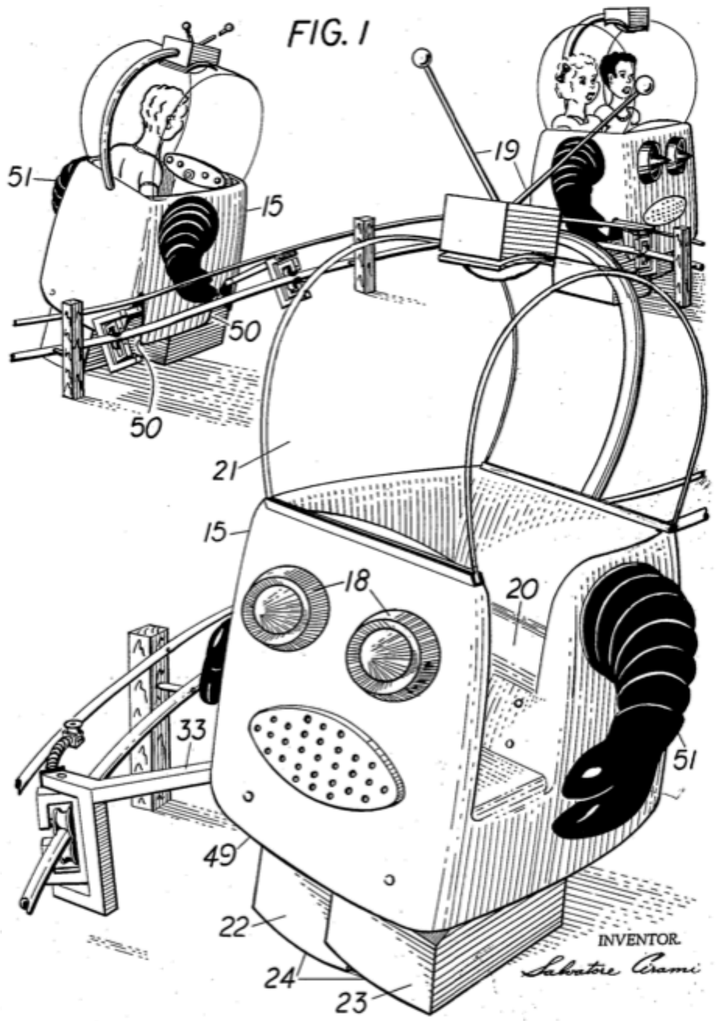
- 7/8 - Adult play tested with dialog and initial comical sense of robots in different scenarios
  - results were very positive and encouraging
  - font and mild sarcasm was appreciated
  - additionally, "cartoon" nature of robot was appreciated and did not detract from the game
- 7/10 - Developed with a child, the reading/ listening game should be fun for most
  - based on simple stories with silly substitution parts
  - simple textual and visual responses should make it easy to appreciate

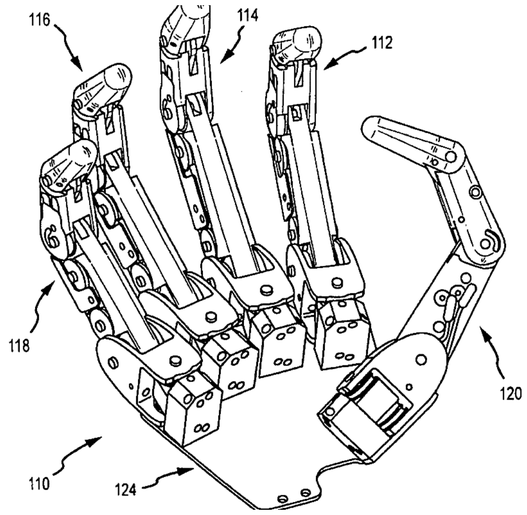


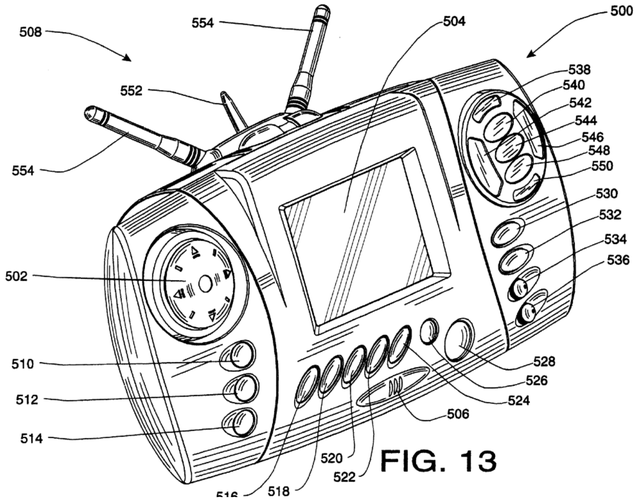
USER TESTING RESULTS (RECORDED IN TIME ORDER)

- 7/12 - testing with cooking/doing stage added a few cute details to the robots, like a chef's hat
- 7/13 - a few profiling rounds found nothing remarkable
  - some of the render textures are expensive but okay
  - no scenes are particularly lagging
  - no excessive script calls, some 4.5k draw calls
  - in testing on development desktop (not VR rig), frame rate did drop to 10FPS or below
-









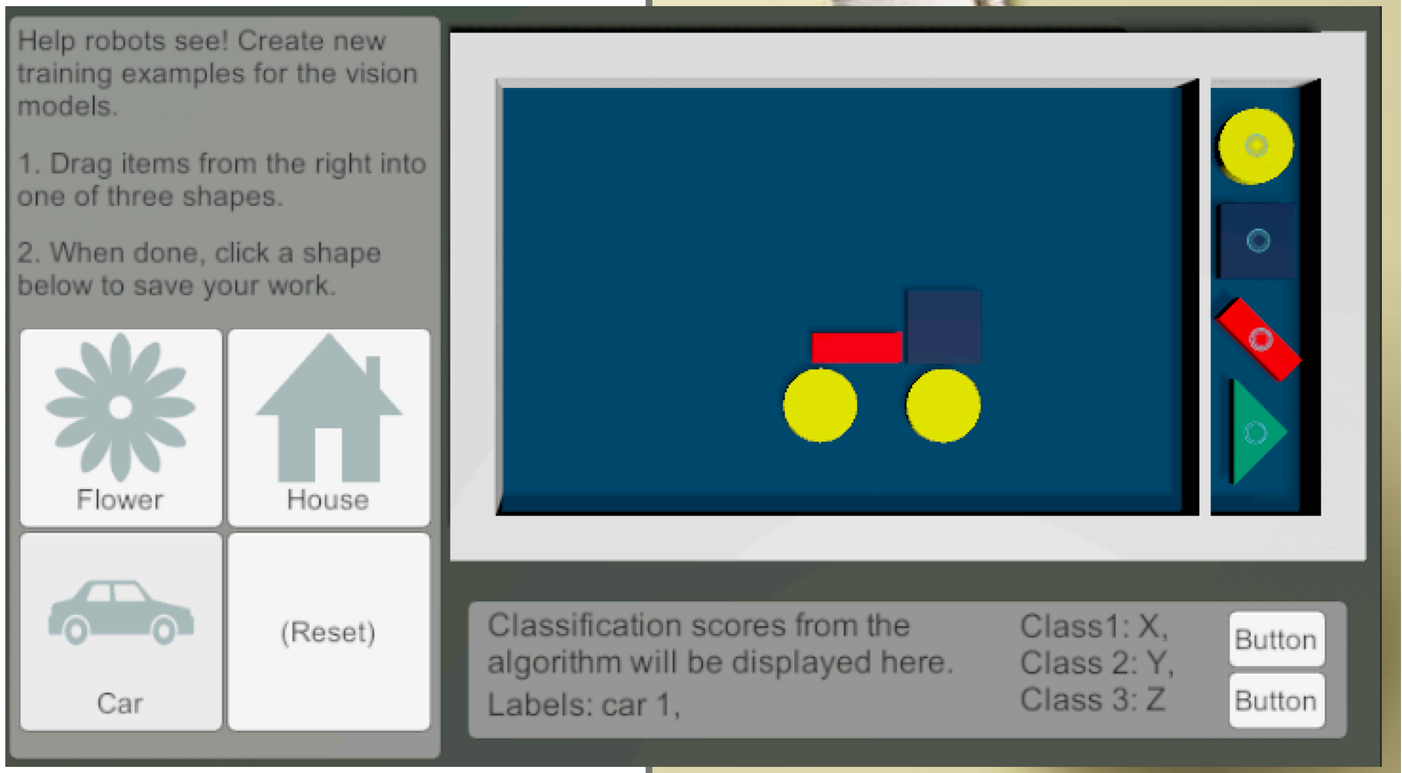
PROGRESS LOG & SCOPING

DELIVERABLE SCOPING (FULL LOG)

- 7/5 - Created design docs (this log) and scoping
- 7/6 - Scene and environment scoping, after discovery of the stylized sci-fi module, it appears that environment creation time can be reduced.
- 7/6 - Kinematics for robotic actions
  - Kinematics for robot more challenging than suspected, a "follow IK" script is good example
  - Deferred: Initially kinematics for repeated actions will loop on one task, but eventually they should be programmatically conditioned
- Text-to-Speech and audio generation for robot speech will be deferred until almost the last point in the game; instead text captions will be used.
  - AudioClip clip2 = Resources.Load<AudioClip>("Sounds/cube_up");
- 7/8 - The dialog system appears to work well for showing robot conversations with triggers.

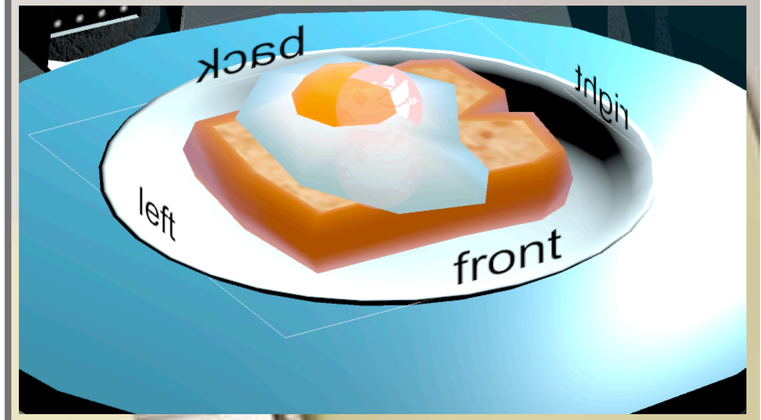
DELIVERABLE SCOPING (FULL LOG)

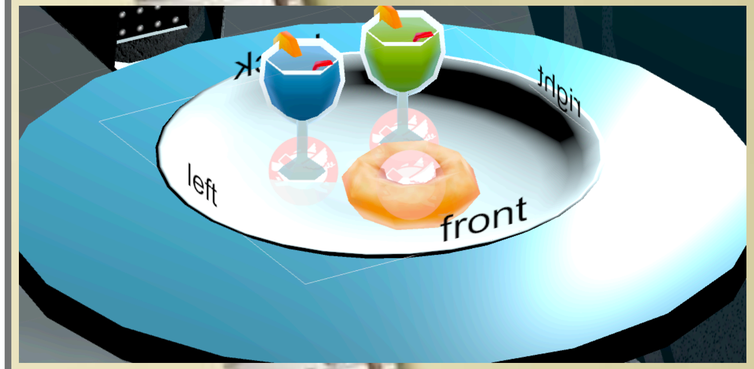
- 7/8 - inspired by 2D drag+drop tutorial, created 2D tool to seed training and evaluation of seeing (primitive object task, below)
  - serialize to data that can be used for training
  - quick evaluation of interaction schemes with VR tools
- 7/10 - listening/reading was completed with simple button interaction as first pass
  - two stories currently included
- 7/10 - move to simplified IK task
  - make robot mirror person
  - goal is to create an object with robot instead of human alone
  - reduced scope of record+playback



DELIVERABLE SCOPING (FULL LOG)

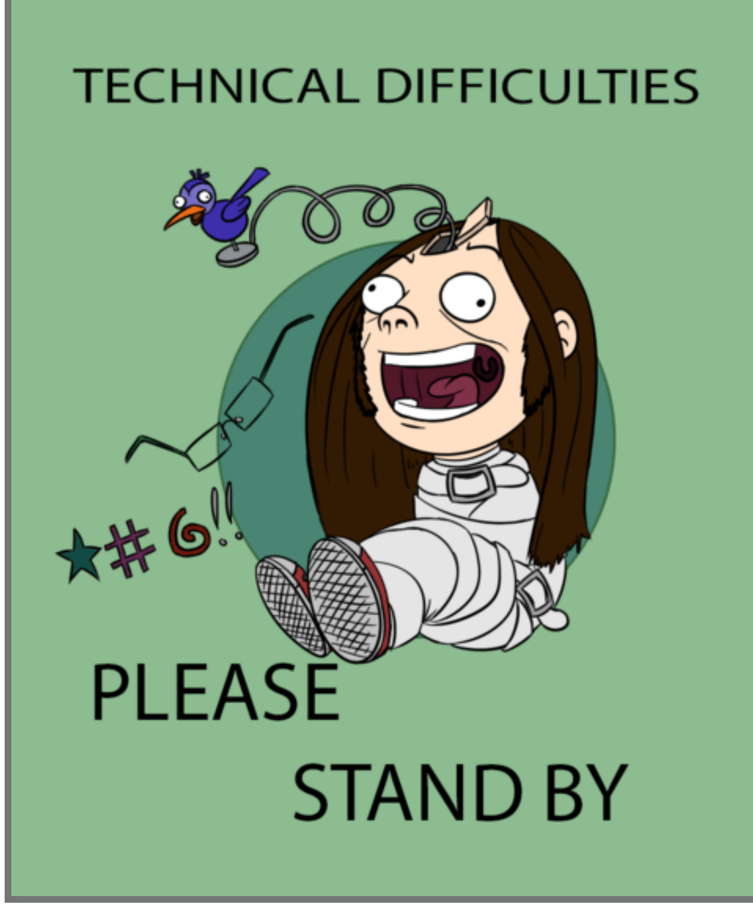
- 7/11 - 2 stories for reading/listening is sufficient
- 7/11 - doing task has one food building right now, need to work on sync with robot IK interaction
- 7/12 - speech recognition libraries may be a non-starter because they require additional builds or hooks for streaming
- 7/12 - reduced complexity of food to mirror
  - turned off physics (although cool) because it made matching too hard and pieces were always moving
  - reduced recipes to be simple 2-3 piece items for evaluation
- 7/12 - finished recording and assigning robot voices
  - completed idle robot "sentry" via IK animation as filler and additional testing within Unity; possibly expand to Boids Flocking?





DELIVERABLE SCOPING (FULL LOG)

- 7/13 - closing notes
  - it was disappointing to not have any home-grown ML, particularly given the effort of collecting data; some flocking code may be investigated, but "seeing" task was incomplete
  - speech recognition was fun to look at but in the end there wasn't a viable solution that was as out-of-the-box as necessary
  - TTS was replaced with reasonable speech + audio processing; additional tools were available for other effects, but not worth the time to evaluate or utilize
  - overall, for just a 6 day period, the final result is quite pleasing

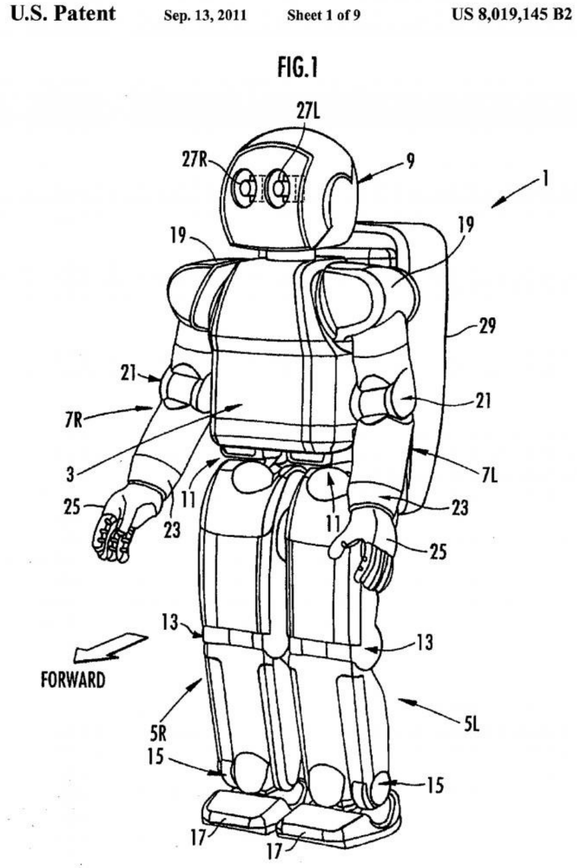


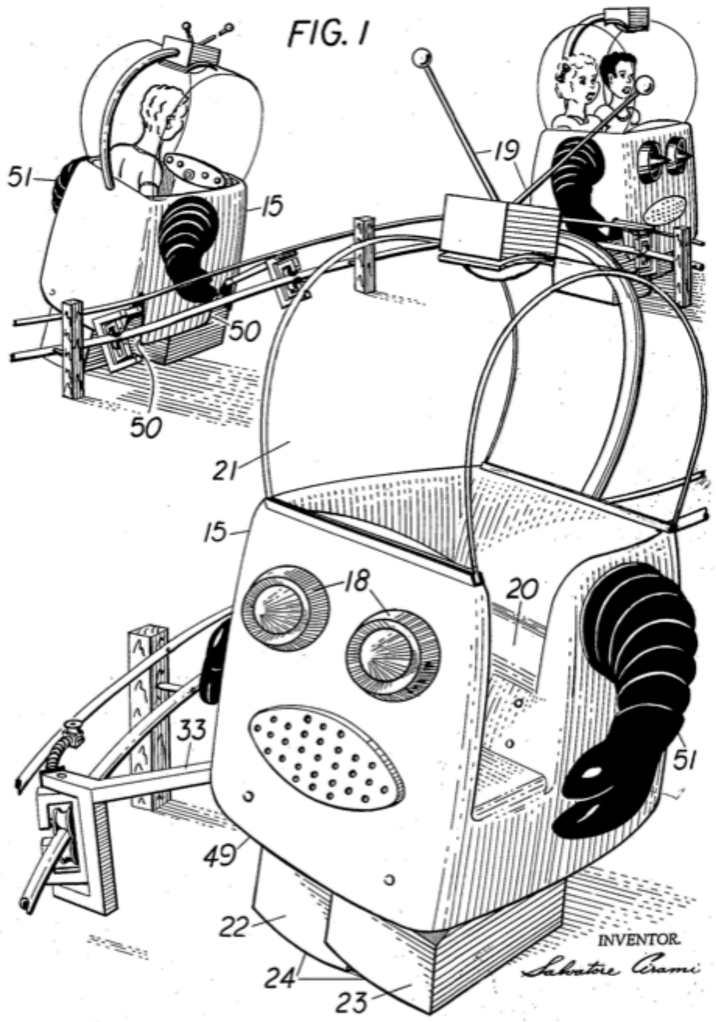
DELIVERABLE SCOPING (FULL LOG)

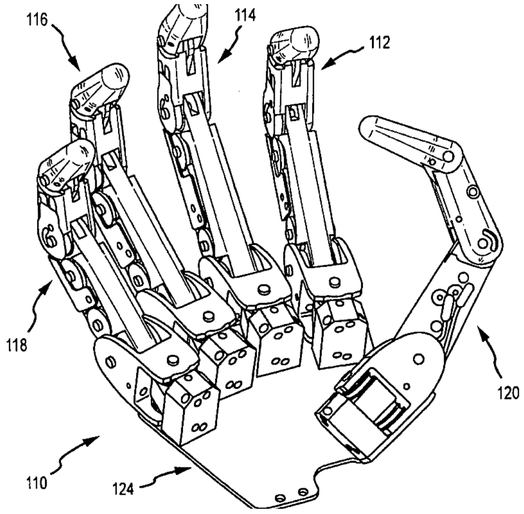
- 7/13 - building and runtime evaluations
  - additional effort was required to get the frame rate to tolerable levels; luckily this banked on some prior experience
    - turning most lights to baked
    - reducing the effects of some particle filters
    - changing some objects to static instead of dynamic
  - other optimizations that could be pursued
    - most slow-downs occurred in the doing (cooking) stage; these may be improved by modifying the camera hover script or making the cameras static entirely.
    - shadow casting was not closely checked for most objects in the scene
    - hunting for emissive materials

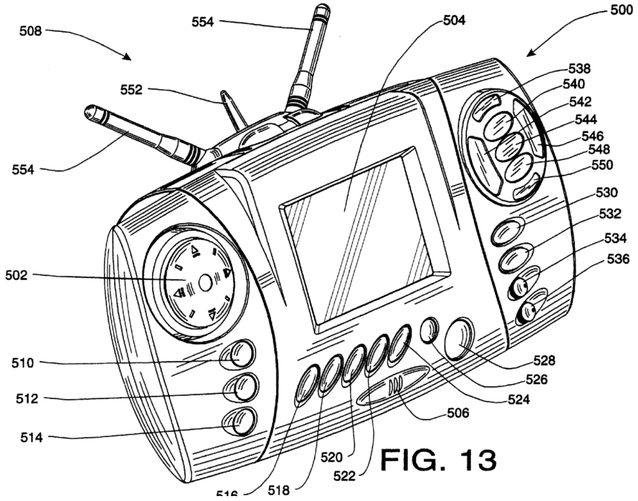
RUNNING TODO (FOR TRACKING)

- Primitives for seeing task
  - basic shapes, snapping deposit location
  - display of status via GUI on adjacent screen
  - full method for feature generation and scoring of seeing task
  - mode to "REPLAY" shapes for a database
  - stretch: mode to upload/retrieve shapes from online database
- Add some gaming capability (like stars in OSD, etc) for accomplishments and time to completion.
- Stretch goal: enhanced dialog state system
- (possible) Improve story telling task - add 1-2 more stories
- Stretch goal: add speech recognition for listening task







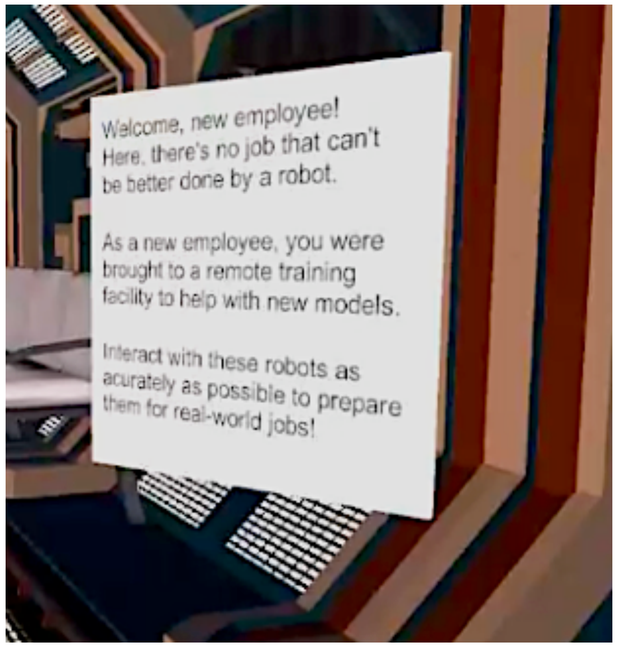


FINAL VR GAME WALKTHROUGH

INTRODUCTION

https://youtu.be/Kobba1wtcrI

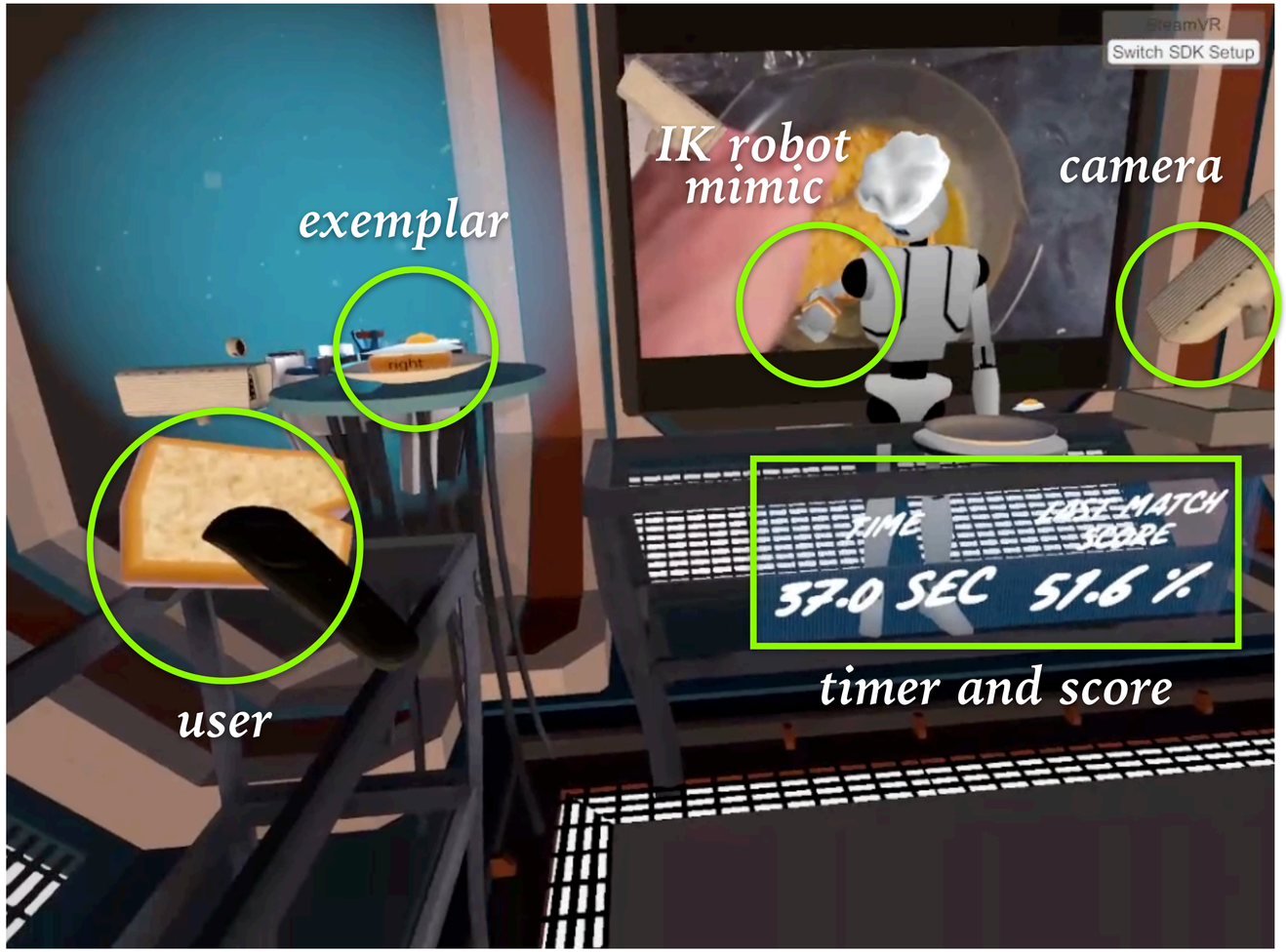
- Upon entry, the user is introduced to the fictional company of Robo Replacements.
- The company has different task stations that the user must help with as a new employee.
- This environment has basic user interactions that change the narrative of each robot task.



DOING/COOKING (1/2)

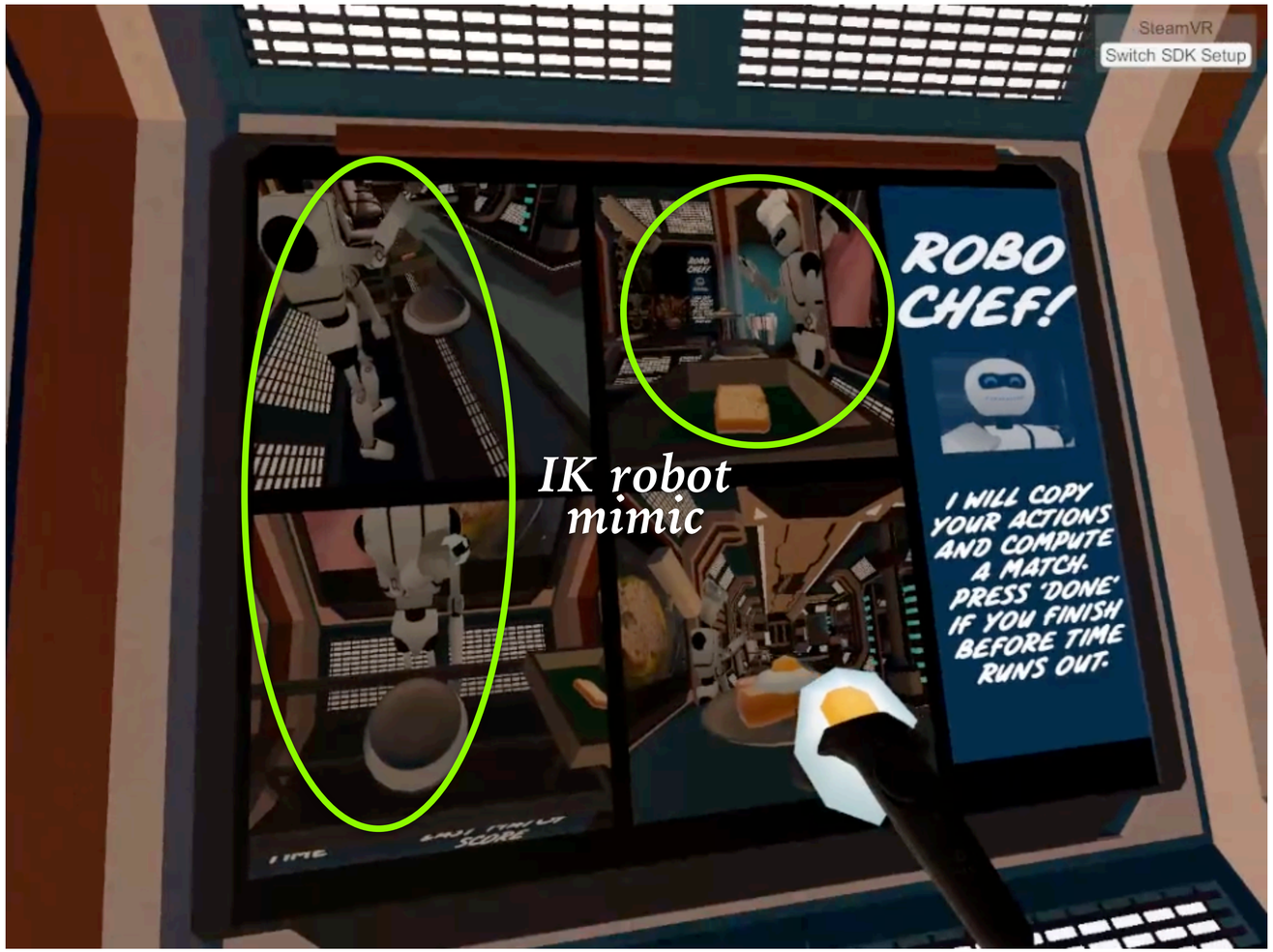
https://youtu.be/YjvpUNM3aP4

- The user must simulate a cooked item by stacking food parts in the same fashion as the exemplar.
- A timer and 3D placement and rotation matching scores are averaged and shown for motivation.
- For flare, cameras hover and the exemplar food rotates.
- The robot mimics then user actions via IK tracking of an object that is moved to a similar position as the user's target object.



DOING/COOKING (2/2)

- Objects are cloned for the robot side and given perfect position symmetry with relative parent position correspondence.
- Using IK targeting actions, the robot appears to mimic the user and is captured via "cameras" as in a TV display.



LISTENING/WRITING (1/2)

https://youtu.be/-QmW1JUYp2c

- In the Listening and writing task, the robot prompts the user with a simple story to complete.
- Instead of a full MadLibs implementation, a small set of words was hand picked for the user to choose from.
- Physics-based push buttons with dynamically populated text are the primary means for interaction.

LISTENING/WRITING (2/2)

- During playback, graphics associated with each user-selected word in the story are displayed.
- In code, stories are easily created and linked to key words via simple templated data structures.
- At time of writing this walkthrough, the stories were not auto-read through a text-to-speech system, but this could be a further refinement.

SEEING/LEARNING (1/1)

https://youtu.be/gugcc3nOjKk

- The seeing task lets the user create basic pictures with a few shape primitives and evaluate them against in-app machine learning.
- Unfortunately, timing prevented this task from being fully completed. But some basic data was collected from a companion app and stored in the repo's /data directory.
- Proximal to the user's interaction was a logic and matching readout to help them understand what was happening in the algorithm itself.

VRND CAPSTONE – ROBO REPLACEMENTS

Eric Zavesky - May-June 2018