Overall: 88/100
Excellent Job!

Group 1

Presentation Mark: 39.5/45

Comments: Great presentation! Please specify PEAS (-1) for your task environment based on class discussions. Please use FSA diagrams, stimulus-response diagrams, function notation for at least one example of the robot's behaviors and show their coordination (-1). Please specify all the roles of the human (supervisor, bystander, mechanic, etc.) for your robot (-0.5). Please include detailed discussions on the design of the body and face of your robot i.e., discuss the Uncanny Valley, McCloud's design space as presented in the lectures as they apply to your robot (-1). Good overview of cybersecurity and patient safety. Please include a discussion of the ethical laws and principles discussed in lectures that apply here (-2).

Report Mark: 48.5/55

Comments: Excellent report! Please include a clear societal justification for your robot by providing, for example, statistics to emphasize its need and/or appropriate references (-0.5). You do mention the pandemic and mental health in general, which is good. Please include HRI taxonomy classification and metrics/benchmarks used to evaluate your system (-2). Good job on your discussion regarding privacy. Please also include discussions of the specific ethical laws and principles discussed in lectures that can directly be applied here (-4). Please note HRI as you have discussed here is a different topic than roboethics.

TASO: Caregiver Robot Under COVID-19 Pandemic

Wenyue Feng, Haocong Song, Zian Zhuang

Mechanical and Industrial Engineering, University of Toronto

MIE1070: Intelligent Robots for Society

Dr. Goldie Nejat

2022 Aug 23

I. Introduction

TASO is a semi-humanoid personal robot, designed to serve in hospitals and nursing homes. The robot is optimized to assist the patient, the elderly, and the health worker to adapt to the changes and overcome the obstacles brought by the COVID-19 pandemic. The most common preventive measures for COVID-19 includes disinfection, quarantining, reducing contacts, or carrying out mask regulations and visitor restrictions. While strictly following such guidelines, the hospitals and the nursing houses can efficiently protect their patient and workers from getting infected, but more challenges are exposured in the actual operation. The connection of the patient and the outside world is weakened, and more extra daily routine tasks are added into workers' schedule, such as checking people's body temperature regularly, reminding patients to wear masks in public areas, etc. Considering the possible problems, TASO is built as the best solution to enhance patient's mental wellness and complete the daily routine tasks for health workers, in the meanwhile reducing the physical contacts of people. The robot performs six major tasks including remote communciation, companion, health diagnosis, contactless delivery, navigation and wondering, and disinfection. To satisfy the complex environment and application, our design adopt the hybrid system to allow the existence of both discrete and continuous mode in the robot motion.

The report consists of five parts. The description of robot appearance and hardware used is in section two. In section three, the six major functionalities are introduced in detail, including the sensors used, behaviours, and the finite state acceptor. The hybrid system design is in section 4. Also, the robot design related ethical concerns and solutions are addressed in section five. Last but not the least, the conclusion and further thoughts are stated in the last section.

II. Robot Appearance & Hardware

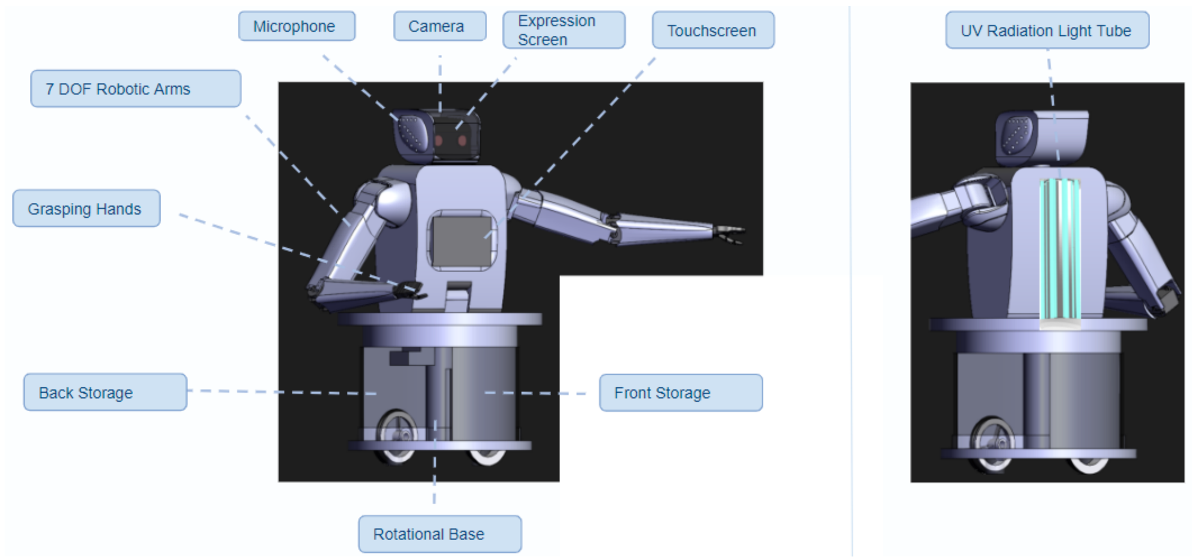


Figure 1: The appearance of TASO

TASO is a semi-humanoid robot, designed to simulate the human body, expression, dynamic movement, and interaction. The actuators act like joints and muscles to support the completion of the goal. Also, the proprioceptive is integrated to sense the orientation, position, the contact in the environment. The components integrate with each other to support the completion of tasks and consist of the lightweight multi-task robot server. To better understand the robot structure, we introduce the robot by the main body component, head, body, and base.

A. Head

Four components are attached to the robot head - camera, microphone, speaker, and expression screen. The speaker is used to display the system voice, for example in the mode of remote chatting; The microphone is used to collect the user's voice; The camera is to record the videos of the users in the video chat mode. The expression screen is used to display the cartoonish expression similar to Zenbo [1]. The expression screen has two modes, one is to reflect the emotion of the person who is chatting with the user, and another one is to display the pre-programmed expression to boost the user's mood if depression symptons of the user is detected. In addition, a motion sensor and light sensor are used to control the on and off of the expression screen. The camera can detect if the user is in the detectable area, if not, the expression screen is off automatically to save energy. Once the user is back to the detectable area, the expression screen is on again. Also, the light sensor can detect the darkness inside of the room. At night, if all the lights are off, the expression screen will also set off automatically.

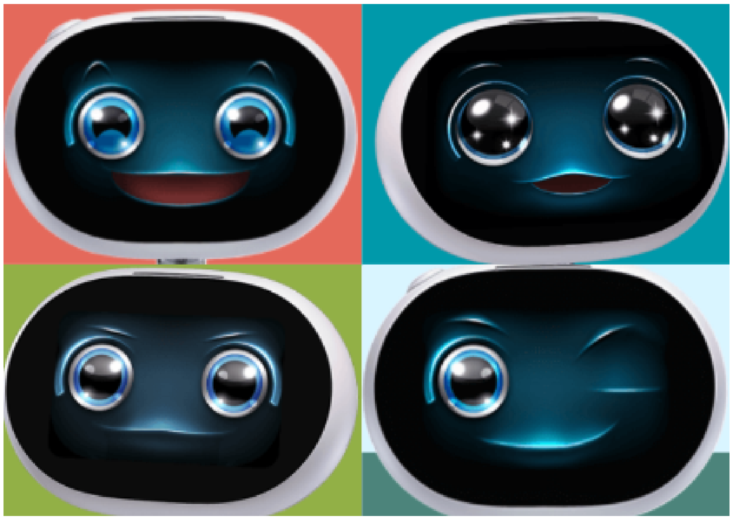


Figure 2: The cartoonish expression of Zenbo

B. Body

A touchscreen is attached to the front of the body. In the remote communication task, the touchscreen is for users to select the options from voice call or video call. In the companion task, the touchscreen displays the music recommendation and food order recommendation for the user, then the user can use the touchscreen to select the desired options. The robot also has two bio-inspired robotic arms and grasping hands. The robotic arms have 7 degree-of-freedom to mimic the actual rotation of human arms. The grasping hands are a 3-finger adaptive soft robot gripper to ensure a solid grip with enough flexibility and versatility.

C. Base

The base contains two major parts, the storage base, and the rotational base. The back storage is storing the devices supporting the operation of the robot, such as field programmable gate arrays, batteries, and wires. The front storage is storing the deliverable items, like healthcare equipment, food boxes, masks, sanitizers, etc. A stationary robot column has located in the middle of the storage base, used to support the height adjustment of the robot to fit the height of different users. The lowest height of the robot is 120 cm, and the highest height can be adjusted to 165cm. At the lowest stage, the robot arms can reach the front storage to grab the deliverable items and pass them to the users. The rotational base supports the 360-degree rotation of the robot, so the robot can always face the users without any manual adjustment.

III. Functionality

The six major tasks are designed to reduce the workload for health workers, and provide a range of services for users' mental and physical wellness while living in an isolated environment. In the remote communication task, the robot plays as a medium encouraging human-human communication. While the robot acts as a companion, it aims to eliminate the depression symptoms by doing activities like chatting, and dancing with the users. It also provides convenient options for food ordering, emergency call making, etc. The health diagnosis task can greatly reduce the human cost under mask regulations - the robot is able to assist the health workers to do mask checking and body temperature measurement, etc. For contactless delivery, the robot helps distribute masks and sanitizers. It also assists with medical equipment delivery, and food order delivery. The gripper can successfully hold and release the item to be delivered. In the navigation and wondering task, the SLAM technology is applied to support TASO easily adapt to the new environment. While TASO is moving around in the public or private area, it performs monitoring and screening at the same time to ensure fast and collision-free movement. The last task is disinfection. TASO disinfects both public and private areas. In addition, to ensure public health regulation, the robot can provide a safety check for the disinfection area before human plans to enter the disinfected area.

A. Remote Communication

Remote communication Supports the remote communication between patients and doctors, or the elderly and family members by touchscreen tablet. While users are chatting with others, the robot can imitate the gesture and the expression of the person in the actual communication that the user is facing. The touch screen is used to receive the video or voice call. Also, the user can use the touch screen to view the real-time image of other people in chat. While the user is in the chatting mode, the user does not have to always sit in front of the robot. The user can continue with other activities, or walk around the room. The robot can detect the movement of the user with the camera, and then use the bumper switch to act as the obstacle detection sensor to automatically avoid hard collisions in the environment with obstacles, such as chairs, beds, tables, etc. The rotation wheels also support the 360-degree rotation of the robot, so the robot can always face the users. The robot arms are used to mimic human gestures in communication. If a gesture can be detected in the video, the movement of the arms will reflect the actual gesture acted by the person that the user is chatting with. Otherwise, the simulation gesture will be generated based on the content of the chat.

In the deliberative system, the plan sequencer includes imitation learning and pose estimation. Imitation learning is applied in the case that the human pose is detectable in the video. To map human pose in video chat to robot pose by artificial neural network mapping model [2]. The pose estimation is applied in the case that the human pose is not detectable in the video. The robot detects the speech content and uses NLP to analyze the emotion of the people who are chatting with the user, and then generates the pose and gesture based on the voice content by GAN [3].

The following is the primary low-level behaviour analysis for the robots in this task. In the perceptual schema, the robot detects the user's location using the camera, uses the bumper switch to detect and track obstacles in the environment, uses the camera to process the live video and detect human arm movement and poses, and then detects the speech content in the live call to understand human emotion and simulate the human poses. In the motor schema, the robot rotates towards the user by generating a vector toward the target user detected by the user location detection perceptual schema. TASO also generates a vector away from any detected obstacles to avoid static obstacles. The arms are moved by the generated vector toward the mapped position or the estimated position processed by the pose estimation.

The detailed finite state acceptor is described below,

1. The robot begins in the start state. No motor schemas are active. The robot is waiting for a signal to begin execution.
2. The signal is the start of the video/voice call.
3. When the video/voice call starts, the robot starts detecting the human pose in the video.
4. If the human pose is detected, the robot starts mapping the pose in the video to the robot's arm movements. If the human pose is not detected, the robot analyzes the speech content and simulates the human pose with the robot arms. If it is in the voice call mode, the robot also simulates human emotion with the cartoonish expression shown on the tablet.

B. Companion

TASO also works as a companion robot to provide companionship for isolated patients or elderlies under the quarantine period or in other isolated scenarios. Compared to the remote communication task, the companion tasks have more interaction between robot and human. The following table includes the five subtasks that the robot can perform, including chatting, music display, food order, emergency call making, and dance coaching.

Task	Sensor	Plan	Actuator
Chatting	Voice Detection Camera (Gesture recognition)	Text/Mood Recognition (NLP, CNN) Text Generation (NLP) Search Engine Memory	Speaker Expression Display
Music Display	Voice Detection Camera (Gesture recognition)	Text/Mood Recognition (NLP, CNN) Text Generation (NLP) Search Engine Recommender System	Speaker Touchscreen
Food Order	Voice Detection Touchscreen Camera (Gesture recognition)	Text/Mood Recognition (NLP, CNN) Text Generation (NLP) Search Engine Recommender System	(Through App) Touchscreen
Emergency Call (Families, Doctors, Police)	Voice Detection Camera (Gesture recognition)	Text Recognition/Generation (NLP) Phone Call Matching	(Through API)
Dance Coaching	Camera Obstacle Detection Touch Sensor	Text Recognition/Generation (NLP) Pose Simulation	Robot arms Twistable base Wheels AR (Augmented Reality)

For all the five subtasks, the common hardware and sensors are microphone, speaker, voice detection and gesture recognition. The microphone is used to accept voice input. Users can speak out about their needs and trigger the voice detection feature of the robot. The speaker is to display the audible voice content for users. Voice detection is to detect the voice input of the user. Considering the possible disability user group, the camera is used to recognize the gestures, so the users do not have to communicate with the robot by voice. After the robot receives the inputs, TASO can analyze the speech content using NLP, and then trigger the rest of the tasks such as replying to the user, displaying music, or order food, etc.

The plan sequencer included in the chatting task involves human intention analysis, text generation, search engine, and short-term memory. NLP is applied to analyze the voice content received from the users to support the intention analysis. Then, the search engine is triggered to browse the related information online based on the user's requirement. Based on the language analysis and search result, the text generation model is started to generate the proper response for the users. Most importantly, short-term memory is also required to allow the robot to memorize the chat content within a chat session, so the follow up questions can be properly responded.

In the music display and food ordering subtasks, the touchscreen is used to display the recommendation options. So, the users can use the touch screen to choose the desired dishes or music. In settings, the recommendation results can also be set to display from the speaker. Then, the users can also use voice to choose the options. In the planning phase, the recommendation system is used to provide the best fit options for the users if the user provides a category of options instead of a specific command. The recommendation system works based on the emotion analysis result, and the personal preference of the user. The emergency call task has the most simple design, as the system will contact the target and then automatically display the pre-coded content for the user, like “successfully contacted the doctor.”

In the dance coaching mode, users can hold the robot’s hands to dance with the robot, following the pre-coded dancing poses. The start of the dance coaching mode is triggered by the touch sensor. If the touch sensor detects the touch, which indicates the start of the dance. The camera is used to detect the position of the user and calculate the body information, for example, the length of the arms. Combining with the obstacle detection sensor, the robot can detect the obstacles in the environment and prevent the user from hitting the obstacles. Then, based on the obstacle detection and the user’s body information, the pose simulation plans the best movement for the user to prevent hurting the user from inappropriate stretching or obstacle collision. Moreover, the user can wear an associated VR glass to select the virtual reality surrounding scene to simulate the environment outside of the private room environment.

C. Health Diagnosis

The COVID pandemic has caused increasing challenges for health care. The increasing number of patients and shortage of health care human resources can lead to delayed care and cost pressure on healthcare systems. Our design tends to provide health diagnosis services for patients in order to meet the health care demands and make healthcare more efficient. The main objective of the health diagnostic system is to make diagnosis results as accurate as possible while ensuring the patient’s safety. The environment for this task is the registered patient, and the sensor information is the patient’s tailored questionnaire and clinical data. The environment type is partially observable because the patient's health information cannot be fully determined with common sensors, stochastic and dynamic because the human health condition is constantly changing and cannot be fully determined by treatments, sequential and continuous because the current diagnosis is based on the past medical records and could affect future diagnosis, and single-agent as the health diagnosis is a one-on-one examination. A utility-based agent with classification models can be built according to the structure shown in Figure 3.

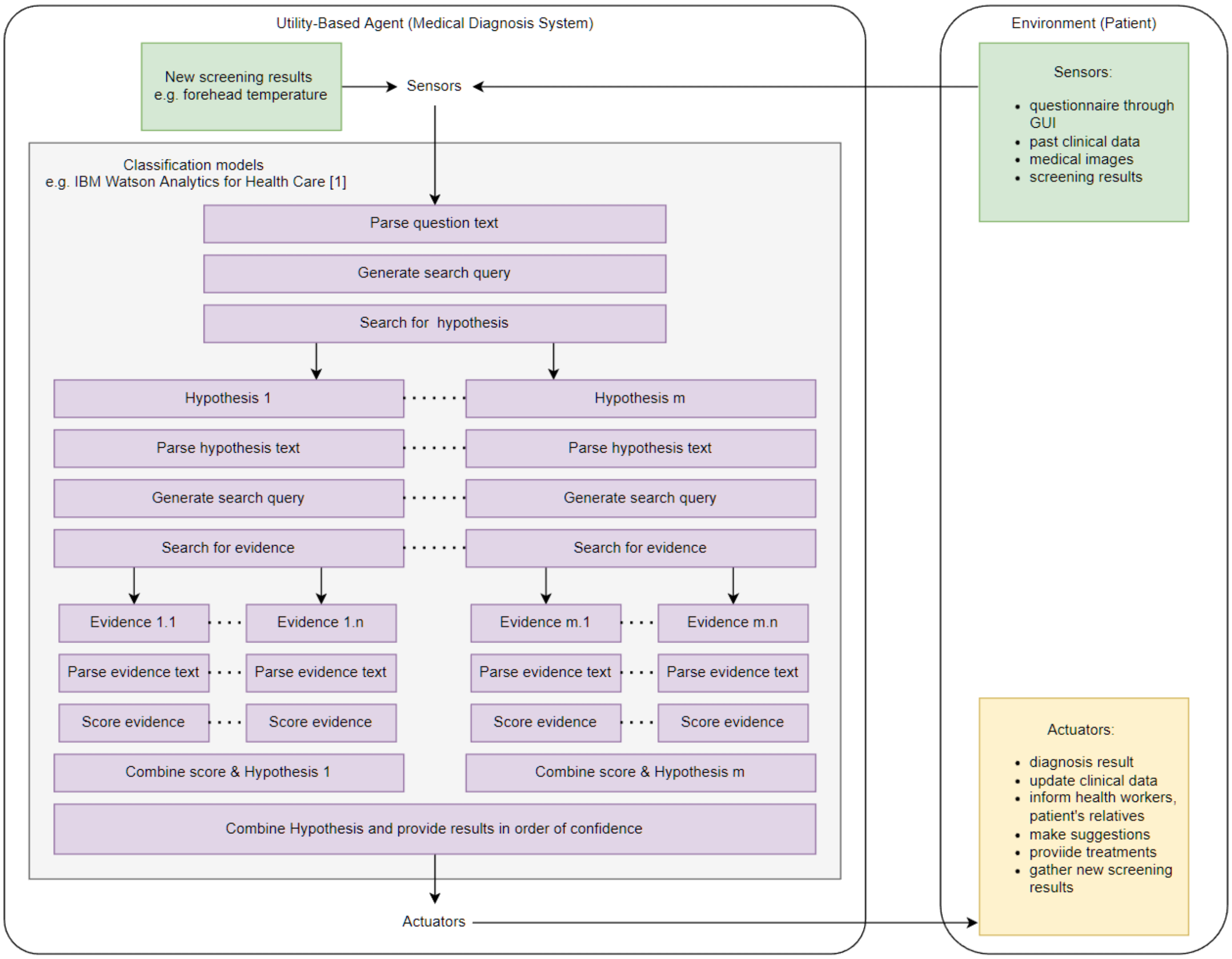


Figure 3. Health diagnosis agent structure

The robot can gather patient information from the patient’s clinical records and by asking a set of health-related questions. These questionnaires can include patient daily behaviour and description of syndromes. Besides that, the recently collected screening results, such as the detected forehead temperature, may be used to mimic how the person’s physiological condition evolves. Following the data preparation, a utility function, such as IBM Watson Analytics [4], can be implemented to rank hypotheses in order of confidence to ensure the accuracy of the diagnosis results. These machine learning classification models can be updated with robot maintenance care. Based on the diagnosis results, several actuator actions can be made by the robot. The robot can inform the patient of the results via a GUI display, update the patient's medical record, and gather more results by conducting the advanced screening. It is worth mentioning that the patient's health record needs to be saved locally and cannot be transmitted over the internet to prevent information hacking. An example of advanced screening can be measuring the patient's blood pressure, so the task scheduler will shift to the delivery functionality for the robot to provide the patient with the blood pressure measuring kit. Moreover, the robot can provide medical treatments and suggestions. To meet the three laws of robotics, these treatments must be verified by authoritative healthcare personnel before putting into action to ensure the treatments are not harmful to humans. To improve human-robot interaction, the agent may provide more persuasive suggestions in a peer-to-peer manner through daily companion tasks instead of representing an authority figure [5].

D. Contactless Delivery

One of the measures to reduce the risk of exposure to COVID-19 in the workplace is contactless delivery. In the delivery task, the goal is to find the right item, pick it up, and then carry it to the target user. The task environment consists of lots of different items, obstacles, and target users. To accomplish this goal, a plan sequence is established in FSA below, as shown in Figure 4.

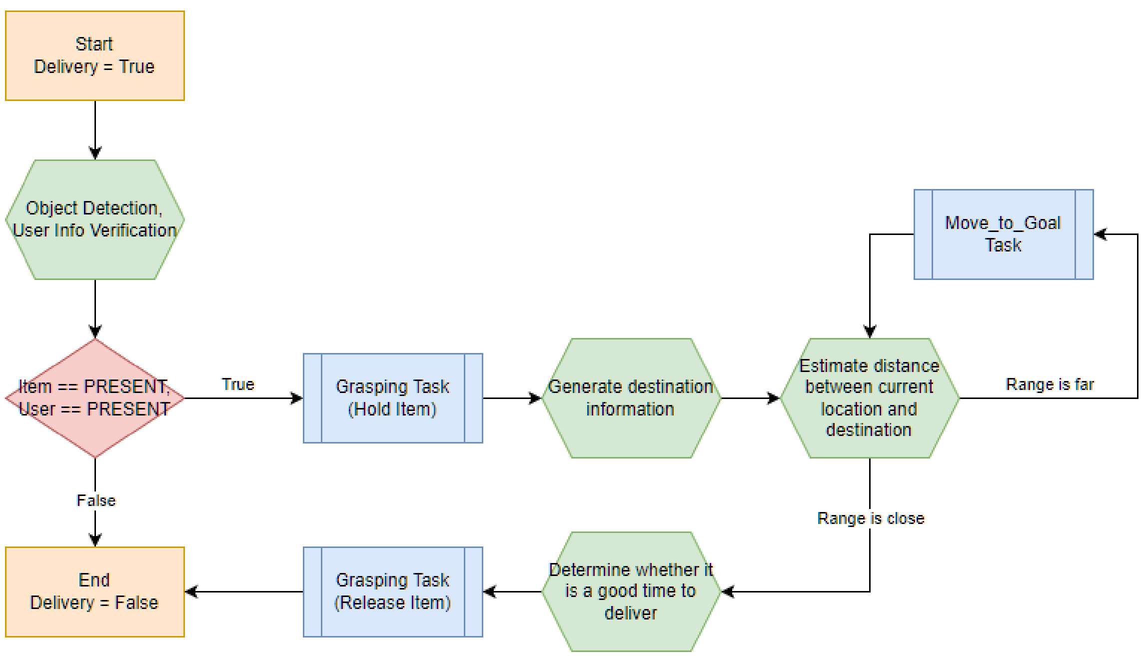


Figure 4. FSA for delivery

The robot first needs to distinguish the items in the environment and determine whether the target user is in the reachable workspace. An object detection model such as CNN can be used to detect object location in the camera input feed. Also, the robot can check the target user's location based on its built world map. For example, if the target user is a patient, the robot can target the patient's residence room as the target location. If both stimuli are present, the releaser initiates the low-level behaviour, grasping mechanism, to pick up the object. After the item is picked up, the target user's coordinates are set as the destination, and the motor schema "move_to_goal" is triggered. The move_to_goal schema, as part of the navigation task, can contain the shortest path finding algorithm to ensure the item can be delivered in an efficient manner. To improve human-robot interaction, the robot can undergo a perceptual schema to decide whether or not to complete delivery based on target user behaviours once the robot reaches the destination. This perceptual schema can be built with a human pose estimation model and speech recognition model for the robot to understand the user's current status. For example, if the person is on the phone, the robot can recognize the person is busy at the moment, and therefore postpone the releasing behaviour.

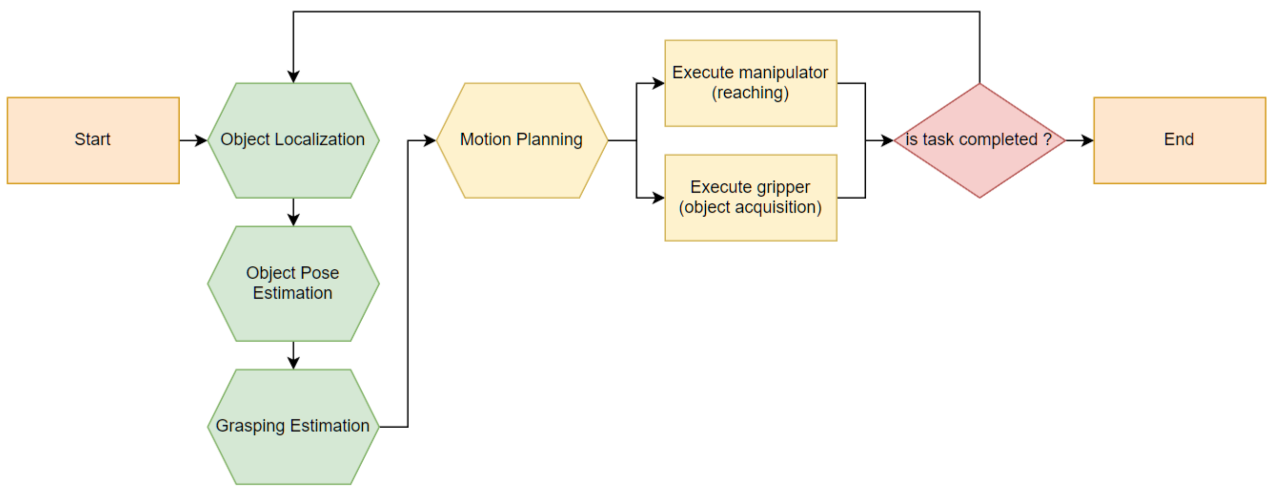


Figure 5. FSA for grasping mechanism

The low-level behaviour, grasping mechanism, can be implemented following the above FSA, in Figure 5. The robot first perceives the object's location and grasping points through pose estimation. The sensors for this perceptual schema can be the color video camera, and infrared sensors. With just RGB image feed, some object pose estimation models, such as DOPE [6], can be utilized to determine the 6-DoF object grasping points. After the grasping points are determined, a motion planning schema can control how the motor schema functions, then execute a manipulator to reach near the object location and execute a gripper to acquire the object.

When designing these actuators, the design team proposes to build two 7-DoF robotic arms combined with 3 jointed cable-driven soft grippers. The 7-DoF robotic arms are inspired by real human arms as the human arms have 7 degrees of freedom [7]. Also, the gripper can be built according to this article to mimic human finger motions and ensure adaptable and effective grasping [8]. This bio-inspired design can also have the capacity for the robot to learn from how humans handle objects and improve its grasping mechanism via imitation learning. The purpose of having two sets of manipulators is to improve delivery efficiency and human-robot interaction. For example, the robot may use the right arm to bring the user the item while holding another item in the left arm.

E. Navigation & Wandering

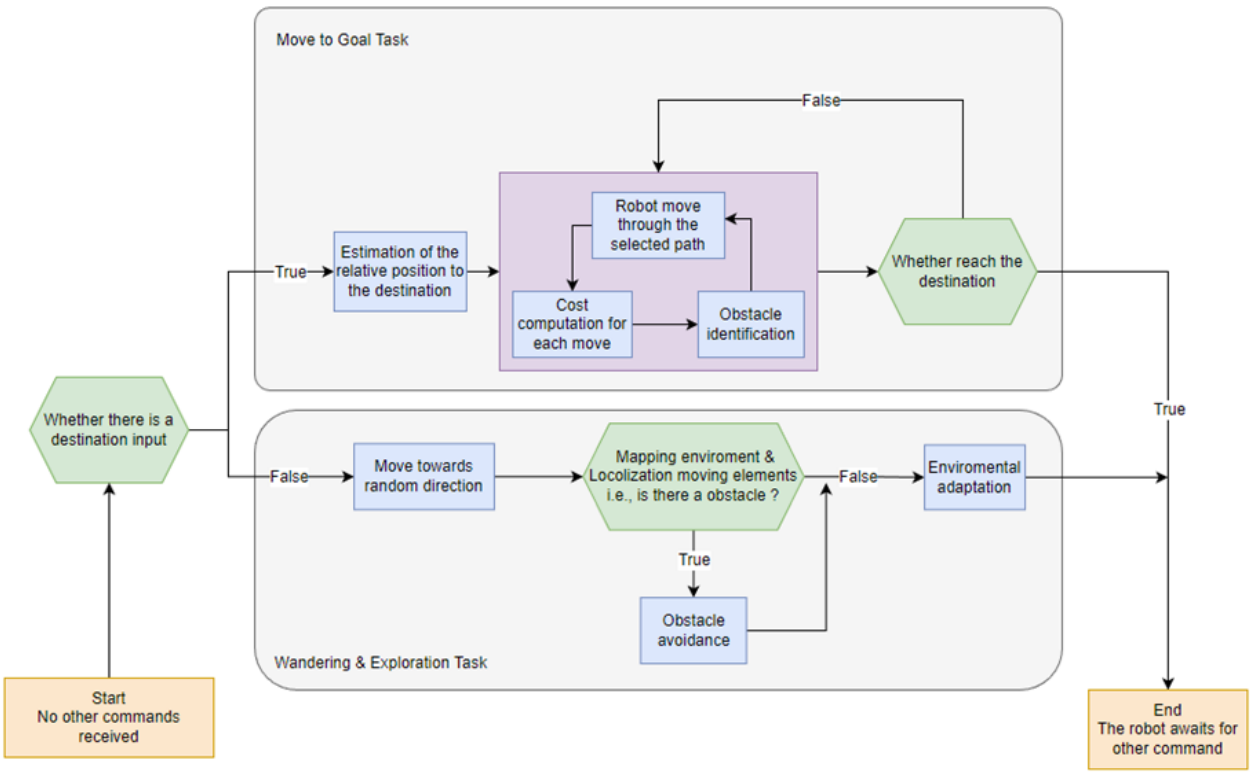


Figure 6. FSA for Navigation

In this part, our robot focuses on its task of navigation through the complex environment under various deployment scenarios, i.e., hospitals, clinics, or public service centers. Navigation & wandering is a fundamental task of our robot as it falls in the wandering state while there are no higher priority commands from users. As is shown in Figure 6. , this task is divided into two subtasks, move to goal and wondering. The plan begins with a perceptual schema to determine whether there is a destination input. If there is a destination input, the robot will first estimate its relative position to the destination and then begin a recurrent state of cost computation, obstacle identification, path generation, and step movement. The state ends when the destination is reached given by the mapping system. The plan for the wondering task is to map the surrounding environment to better adapt to other tasks. For example, when the robot is in a wandering state, the lidar system will construct a 3D representation of the surrounding environments into the memory so that the robot will not have to rebuild the map when navigating through the same place. During the mapping process, obstacle avoidance is achieved when obstacles are detected by the sensors.

The sensor system which supports this task mainly consists of Lidar and cameras. The lidar system serves as the main sensing method to construct the 3D representation of the surrounding environment. The camera system serves as the auxiliary sensing method to detect human poses in the surrounding environment. When sudden movements of humans which can cause collisions are detected, the bumper switch is actuated to avoid collisions. The camera is also integrated with an OCR system to read the text information of the surrounding environment such as room numbers and sector IDs.

The primary low-level behaviour analysis of navigation & wandering tasks is presented below:

- Perceptual schema
    1. Detect surrounding environment: Use Lidar to build a 3D representation of the environment.
    2. Detect obstacles: Use the camera to perform real-time object detection to detect static obstacles.
    3. Detect human pose: Use the camera to perform real-time detection of human body movement and possible incoming collision.
    4. Detect text information: Use the OCR system to detect the text information in the surrounding environment.
- Motor schema
    1. Collision avoidance: Generates a vector against the incoming collision.
    2. Avoid static obstacles: Generates a vector away from the static obstacles.
    3. Move towards the destination: Generates a vector towards the destination.

F. Disinfection

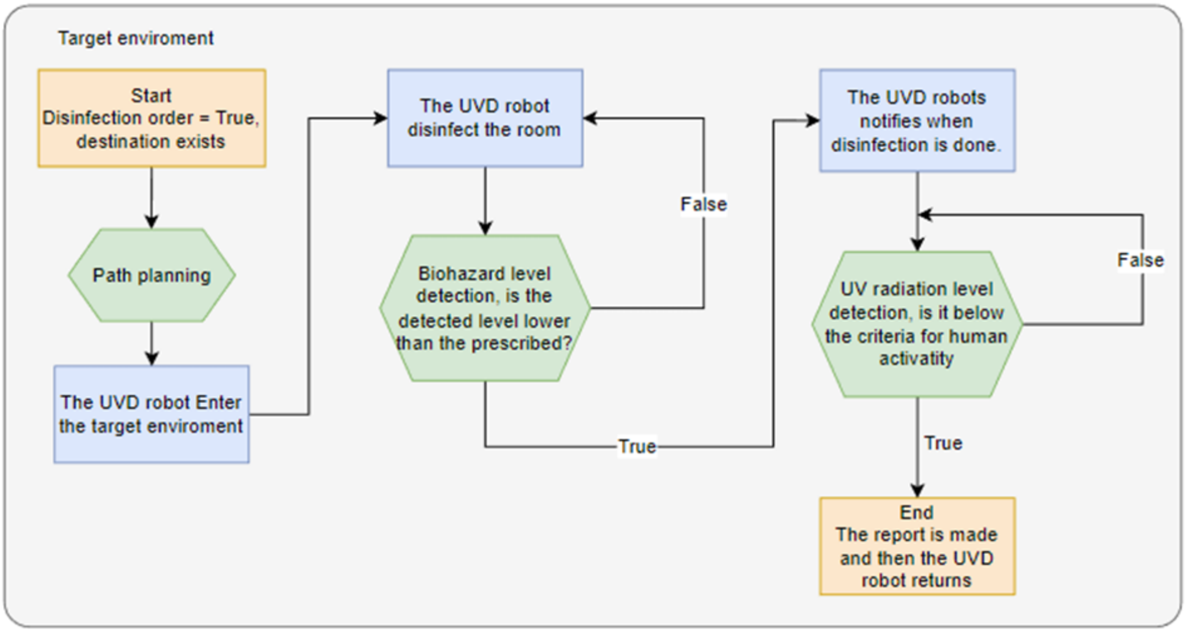


Figure 7. FSA for Disinfection

On top of our navigation system, we provide environmental disinfection capability of our robot called UVD (ultraviolet disinfection). In order to ensure the safety of users to our best, the two major perceptual schemas are biohazard level detection to monitor the disinfection quality and UV radiation detection to monitor the remaining radiation level after the disinfection. As shown in Figure 7 the detailed finite state acceptor of the UVD robot begins with the cleaning staff ordering UVD Robot via the communication system (app). The UVD robot then makes path planning via the navigation system, i.e., opening the doors, taking the elevators, and entering the rooms. After the cleaning staff goes through the security checklist, the UVD robot disinfects the room. The biohazard level detector will evaluate the disinfection quality at the end of each cleaning round. After the UV radiation level is below the criteria given by the UV radiation probe, the UVD robots notify when disinfection is done. The final report is made and then the UVD robot returns.

The primary low-level behaviour analysis of navigation & wandering tasks are presented below:

- Perceptual schema
    1. Detect obstacles during navigation: Use sensors form the navigation system to detect obstacles during the navigation process.

2. Detect the target environment for disinfection: Use a camera and OCR system to identify the target room that awaits disinfection.
3. Detect bio-hazard level according to the disinfection criteria: Use a bio-hazard level detector to determine whether the disinfection is qualified.
4. Detect UV radiation level: Use a UV radiation probe to detect whether the environment after disinfection is available for human activity.

- Motor schema
    1. Move towards the destination: Generates a vector towards the destination, including necessary actions such as taking the elevator, opening the door, and entering the room.
    2. UV disinfection: Generates ultraviolet radiation to disinfect the target environment.

IV. System Design

Since the task environment for each function is different and the world model keeps changing while the robot is moving, a hybrid system is required to ensure the robot can perform well in dynamic settings. The hierarchical integration model is designed following AuRA (Autonomous Robot Architecture) as the principle interface strategy. The navigation functionality allows the robot to create, store and update the world representation while moving in the environment. The cartographer maintains this mapping information in the long-term memory and provides two sets of planner logic for the robot to follow: public space and private space. The public space planner logic is suitable for areas such as the hospital corridor, lobby, and main entrance, where there is massive human traffic and screening and health monitoring is necessary. While operating in private spaces such as waiting rooms, isolation rooms, and recovery rooms, where access is limited to patients, visitors, and health workers, the other planner logic can ensure the robot provides a better user-targeted health care service. By setting the Start_State as wandering, the robot may decide which planner logic to follow depending on its current position via localization. Regardless of the environment in which the robot is operating, the mission planner always schedules the user-defined tasks as the highest priority for the robot to process. At the current stage of the design, the user-defined functionalities include disinfection and equipment delivery, and the system can add more in the future as needed. The robot can execute these common tasks by interpreting human commands via speech recognition and GUI (graphical user interface). These common tasks can override other tasks which the robot has at hand and cannot be interrupted until they are completed or abandoned.

Planner for public space

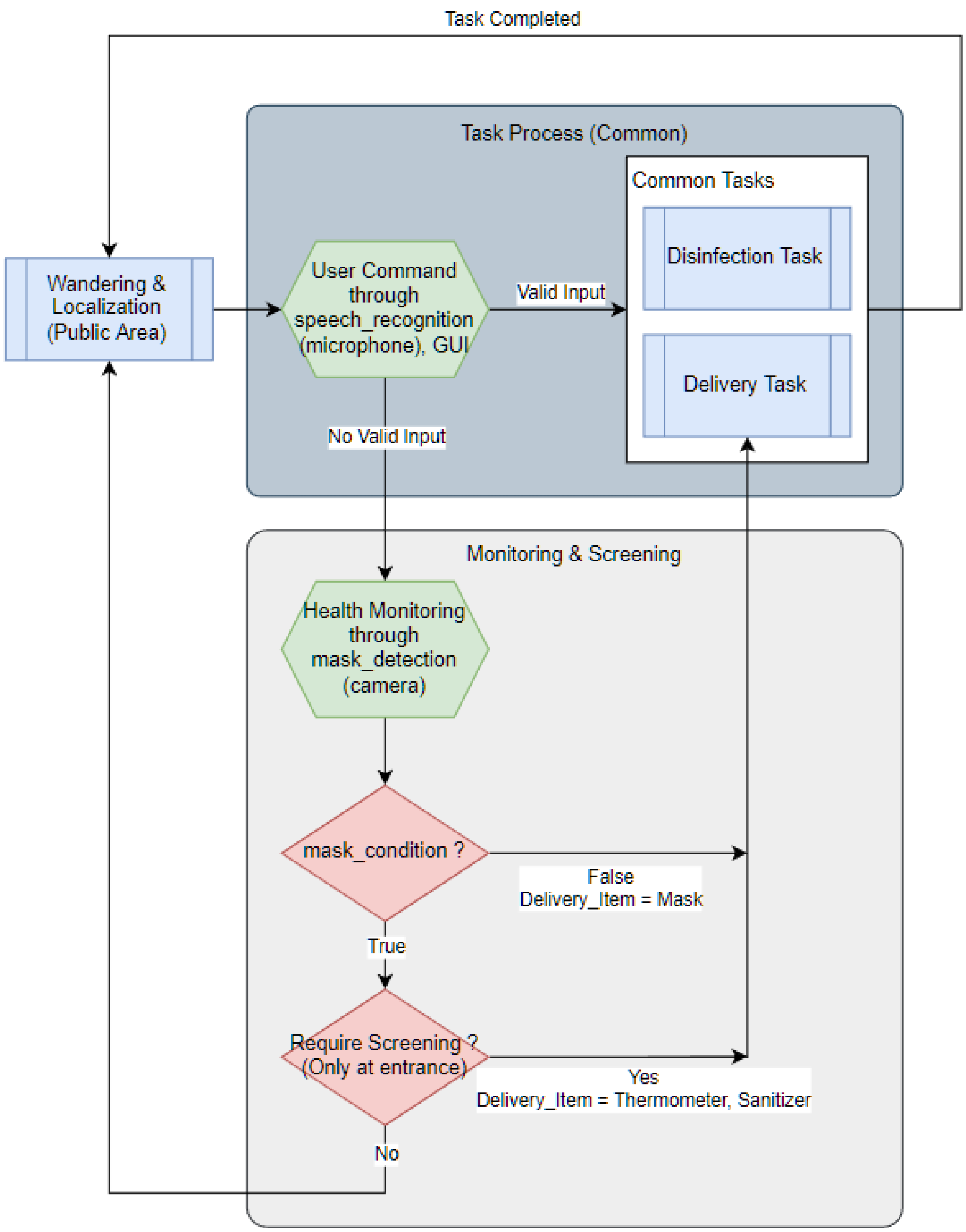


Figure 8. FSA for public space planner

While the robot is functioning in public space and there's no user-defined task at hand, the robot is designed to perform health monitoring and screening duties while moving around. To accomplish this, a plan sequencer, as shown in Figure 8, can be implemented. The logic first checks whether there is a valid user command input, if there is no such command, the robot will continue wandering around while detecting the surrounding environment. If the robot senses there is a person nearby, it will apply a mask detection algorithm to determine if the person is wearing a mask. This mask detection algorithm can be a convolutional neural network (CNN), which can be trained with tons of masked and unmasked images to classify the problem. This machine learning model can be updated through robot software maintenance in the future. If the result of the model, mask_condition, is false, meaning the person doesn't have a mask equipped, a releaser will trigger the robot to perform the delivery functionality as a pilot subtask. In this case, the delivery_item would be a mask and the target_user would be the person who needs the mask. In certain areas where further screening is necessary, such as at the building's entrance, the robot can perform other similar tasks, such as distributing sanitizer or checking forehead temperature.These certain areas can be marked in the world map in the robot's long-term memory. While the robot is located in these regions, the robot would perform the additional tasks. Besides checking the mask condition of the nearby humans, the robot can check whether they have a fever by moving the digital thermometer towards their forehead. Additionally, the robot can recognize people's voice commands via a microphone and supply their sanitizers if they ask for them. If there are no people nearby, the robot will keep wandering around while getting ready to handle new tasks.

Planner for private space

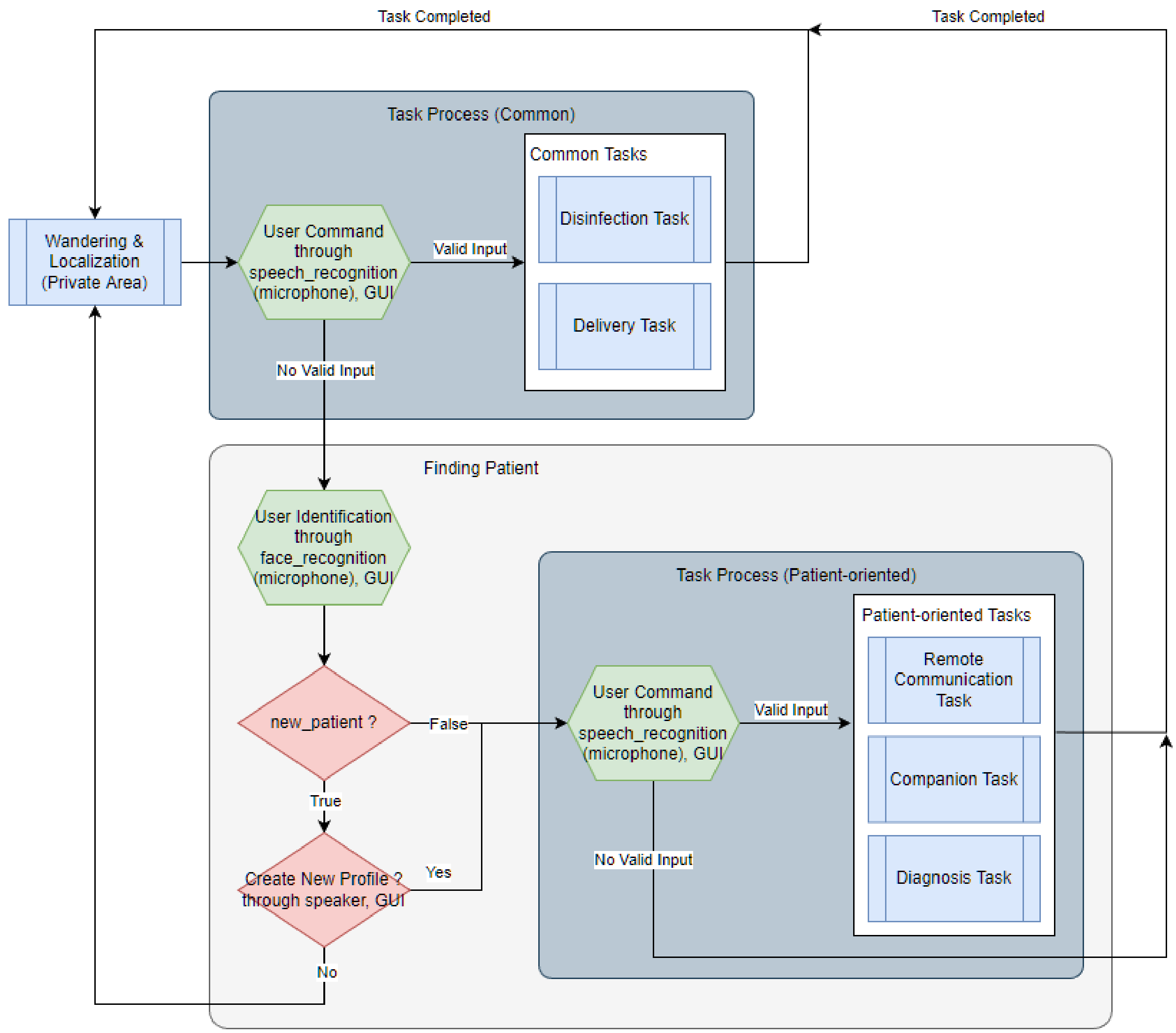


Figure 9. FSA for private space planners

Unlike in public places, masks are not always mandatory in private places. Patients in the recovery room usually do not have to put the masks on at all times unless they are meeting with health care staff or visitors. While operating in these areas, the robots should be able to identify human looks and provide services specific to each individual. A plan sequencer for private space, as shown in Figure 9, is created to achieve this goal. Similar to the public space planner logic, the logic first executes user-defined tasks when there is a valid user command input. If there is no valid command, the robot will move around while identifying each person nearby with face recognition technology. An example of face recognition technology can be a combination of face detection and feature-based face analysis. First, a face detection model is applied to extract face regions in the camera image input to detect human beings. Then, the face analysis algorithm transforms 2D analog face information into a numerical code called faceprint based on the person's facial features. Since we want one robot to serve as many people as possible, the robot can save this face data in its local database. Depending on whether the detected faceprint matches with the robot database, it can construct new patient profiles or load past patient profiles as needed. After the patient profile has been loaded, the robot can perform patient-based tasks similar to the main mission planner based on user input commands. At the current stage in the design process, the user-defined health care service includes remote communication, companionship, and health diagnostics, and the system has the ability to add more in the future. If the person detected does not need such service, the robot will leave the person and provide assistance to other patients.

V. Roboethics

To address the robot ethic concerns that the application scenarios could bring, we propose the following five strategies with respect to data security and human-robot interaction.

1. As patient information and clinical data are confidential and sensitive, our diagnosis system will store them locally to ensure cybersecurity.
2. Treatments such as prescriptions generated from the diagnosis system will be verified by certified health workers before being recommended to patients.
3. For the human-robot interaction, we choose cartoonish expressions rather than those of real human faces in our communication interface to prevent the uncanny valley effect.
4. Only Pre-programmed poses are adopted by the robot to prevent harmful or offensive actions.
5. In order to support minority groups with possible disabilities, we include voice and gesture recognition in our communication system.

VI. Conclusion

In this project we present TASE, a semi-humanoid personal robot designed to serve in hospitals and nursing houses. Designed with six major functions including remote communication, companionship, health diagnosis, contactless delivery, navigation, wondering, and disinfection, TASO can well assist the patient, the elderly, and the health worker to adapt to the changes and overcome the obstacles brought by the COVID-19 pandemic. The general appearance of the robot is first introduced to offer an overall look. Demonstrations of each function are discretely presented with flow charts, schema lists, and descriptions of the finite state acceptor. Hybrid systems for both public and private spaces are designed for our robot to satisfy the complex environments and different application scenarios, allowing the existence of both discrete and continuous mode in the robot motion. Considering about ethic problem, our design allow accepting information by voice and gesture to support various disability groups. However, the current design has not consider about the convinience of the group with weak vision ability. How to improve the robot acceptance rate of the group with weak vision ability is the next step that need to be carefully desiged. Moreover, in future work, VR/AR technology can be integrated with the robot companion tasks, so the patient and the elderly can have more chance to interact with their familly members, or friends.

VII. References

[1] "Zenbo smart home", Zenbo Global, 2022. [Online]. Available: https://zenbo.asus.com/whatsnew/features/smart-home/.

[2] Z. Al-Qurashi, "Robot-Human Mapping Model Learning for Robotic Imitation using Deep Learning and Virtual Reality", figshare, 2022. [Online]. Available: https://indigo.uic.edu/articles/thesis/Robot-Human_Mapping_Model_Learning_for_Robotic_Imitation_using_Deep_Learning_and_Virtual_Reality/13475985/1.

[3] Chuang Yu, Adriana Tapus. SRG3: Speech-driven Robot Gesture Generation with GAN. 16th International Conference on Control, Automation, Robotics and Vision, Dec 2020, Shenzhen (virtual), China.

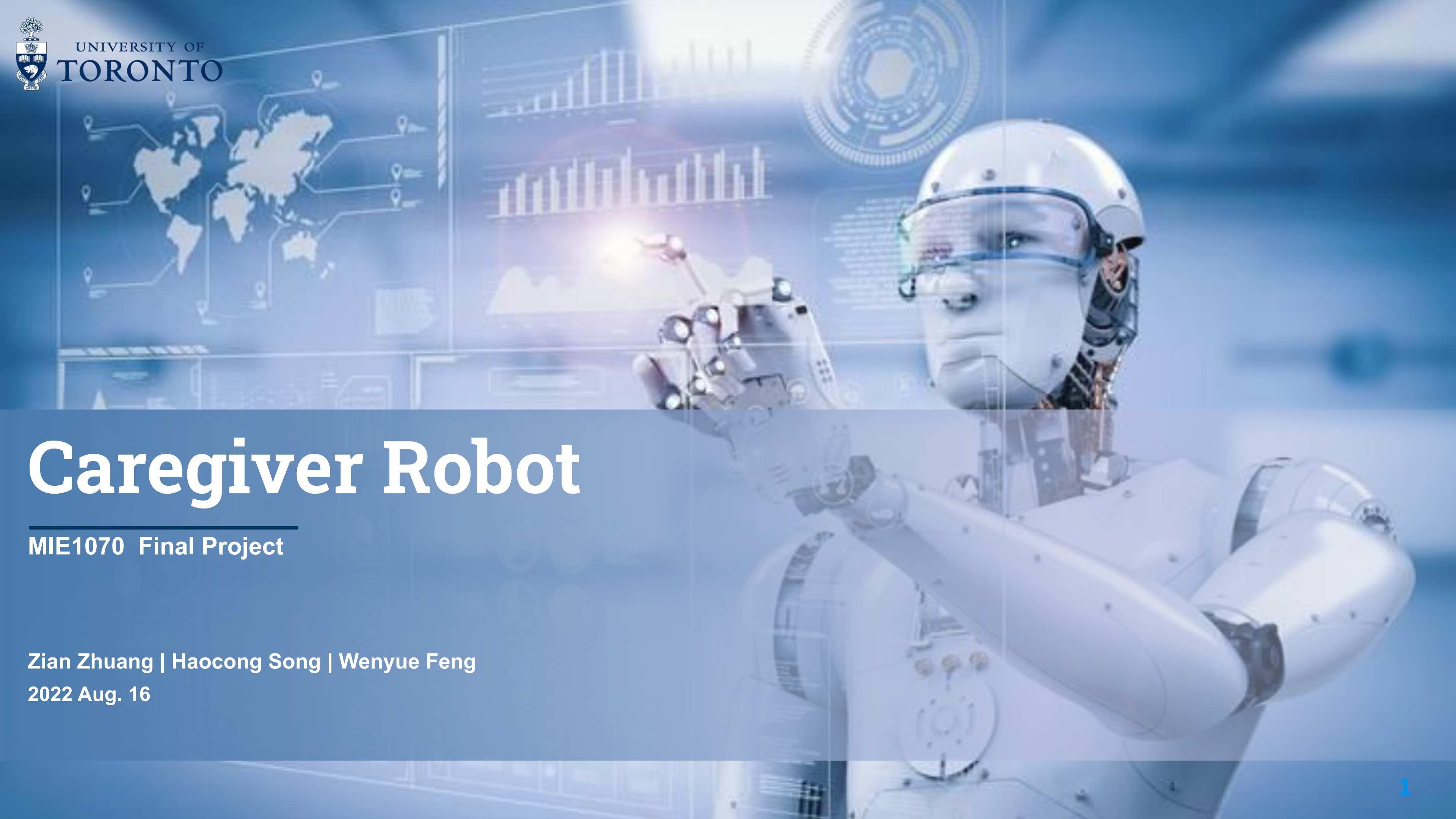
[4] Aggarwal, Mayank & Madhukar, Mani. IBM's Watson Analytics for Health Care:. 2017.

[5] S. P. Saunderson and G. Nejat, “Persuasive robots should avoid authority: The effects of formal and real authority on persuasion in human-robot interaction,” Science Robotics, vol. 6, no. 58, 2021.

[6] J. Tremblay, T. To, B. Sundaralingam, Y. Xiang, D. Fox and S. Birchfield, "Deep Object Pose Estimation for Semantic Robotic Grasping of Household Objects", arXiv.org, 2022. [Online]. Available: https://arxiv.org/abs/1809.10790.

[7] Hyunchul Kim, L. M. Miller, A. Al-Refai, M. Brand, and J. Rosen, “Redundancy resolution of a human arm for controlling a seven DOF wearable robotic system,” 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011.

[8] Manti, Mariangela & Hassan, Taimoor & Passetti, Giovanni & NicolòD’Elia, & Laschi, Cecilia & Cianchetti, Matteo. . A Bioinspired Soft Robotic Gripper for Adaptable and Effective Grasping. Soft Robotics. 2015.



UNIVERSITY OF TORONTO

Caregiver Robot

MIE1070 Final Project

Zian Zhuang | Haocong Song | Wenyue Feng

2022 Aug. 16

Contents

Introduction

Objective

- Support multiple one robot-one human and one robot-multiple human interaction tasks in COVID settings

Environment

- In hospital or nursing house, including both public area and private room area with obstacles such as chairs, tables, bed, etc.

Target Users

- The patient living in hospital | The elderly living in nursing house | Health workers

Design Idea

- Hybrid system

6 Major Tasks

Task 1

Remote Communication

- Encourage the communication between patient and doctors, or isolated elderlies and family members.

Task 2

Companion

- Eliminate depression symptoms by activities like chatting, dancing with users. Provide convenient options for food ordering, or emergency call making.

Task 3

Health Diagnosis

- Reduce human cost on mask regulations and body temperature measurement, etc.

Task 4

Contactless Delivery

- Distribute mask, sanitizer
- Assist on medical equipment delivery, food order delivery, etc.
- Hold and release the item to be delivered.

Task 5

Navigation & Wondering

- Adapt to new environment (through SLAM)
- Monitoring and screening while moving.
- Fast and collision-free movement

Task 6

Disinfection

- Disinfect both public and private areas.
- Provide safety check for the disinfection area before human entering

Appearance Overview

Microphone

Camera

Expression Screen

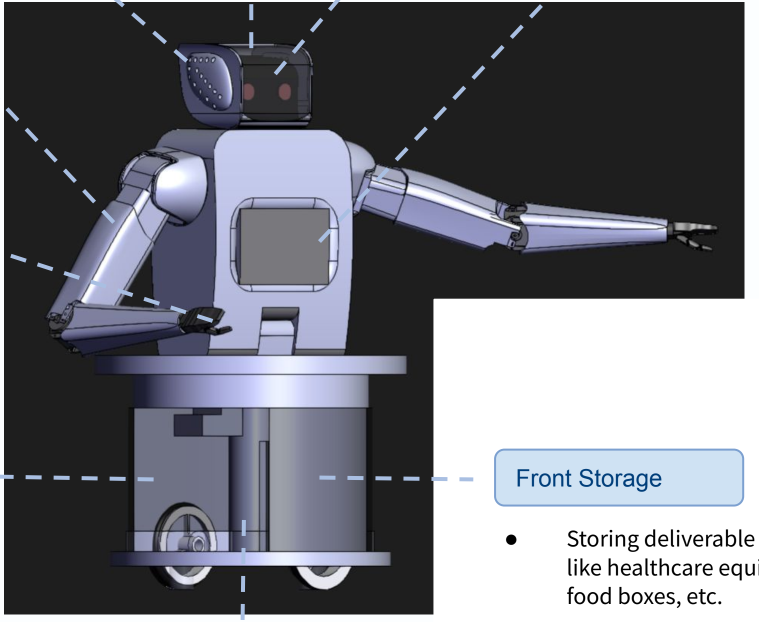
Touchscreen

7 DOF Robotic Arms

- Bio-inspired from human arm

Grasping Hands

- 3-finger adaptive soft robot gripper
- Bio-inspired from human hands



Back Storage

- Storing FPGA (Field Programmable Gate Arrays), batteries, wires

Front Storage

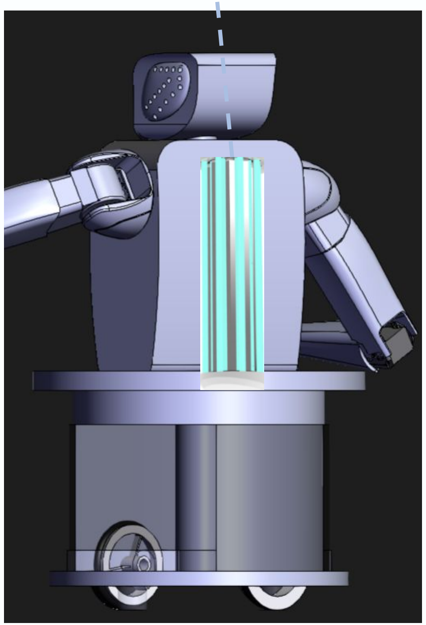
- Storing deliverable items, like healthcare equipments, food boxes, etc.

Rotational Base

- Support 360-degree rotation

Front View

UV Radiation Light Tube



Back View

Contents

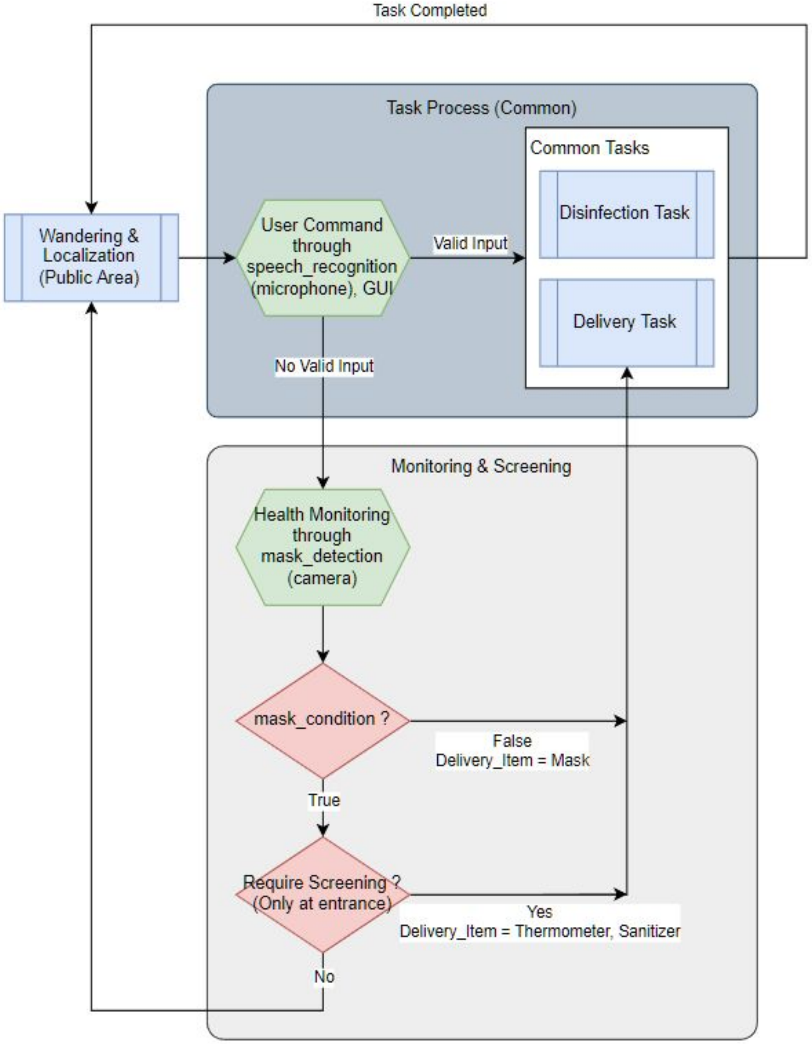
1st Deliberative Layer

Environment: Public space (e.g. corridor, lobby)

Mission planner: Common Task Execution
- Navigation & Wondering
- Disinfection
- Equipment Delivery

Plan Sequencer: Monitoring & Screening
- Supply mask
- Distribue sanitizer
- Check forehead temperature

HRI awareness:
- interpret user commands
- know human locations
- understand human needs (e.g. mask)



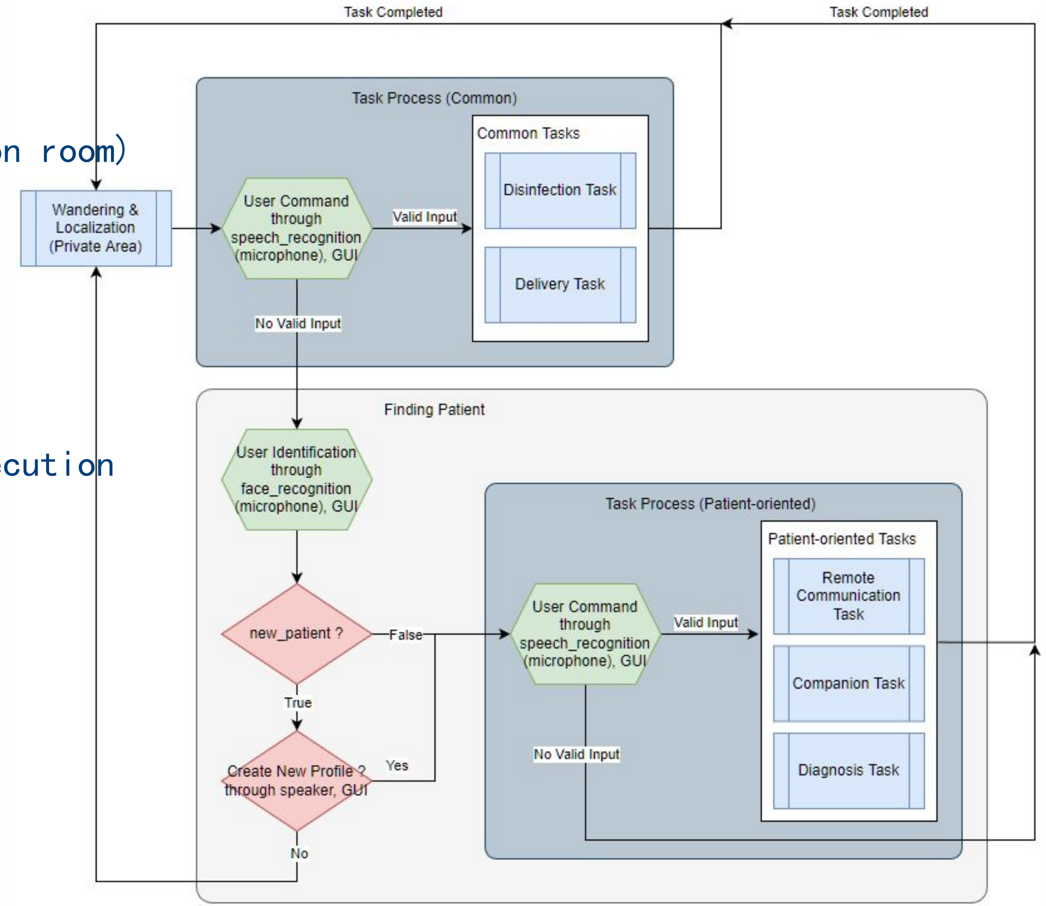
2nd Deliberative Layer

Environment: Private space (e.g. isolation room)

Mission planner: Common Task Execution
- Navigation & Wondering
- Disinfection
- Equipment Delivery

Plan Sequencer: Patient-oriented Task Execution
- Patient Identification
- Remote communication
- Companion
- Health Diagnosis

HRI awareness:
- interpret user commands
- know human locations
- identify human appearance



Contents

Task 1: Remote Communication

Objective

Supporting video and voice call between patients and doctors, or elderlies and family members by touchscreen tablet. Imitating the human gesture in actual communication.

Environment

Involves obstacles such as table, bed, chairs, wastebasket, etc.

Target Users

Patient or isolated elderly

Hardware & Sensing

Touchscreen: to receive video or voice calls, and view the real-time image of other people in chat.
Bumper Switch: to autonomously avoid collisions, like hitting a table.
Camera: to record videos of the users in the video call mode.
Robot Arms: to map the human pose in video, or mimic human pose based on speech content.
Rotational Wheels: supports the 360-degree rotation of the robot, so the robot can always face the users.

Task 1: Remote Communication

Cartoonish Expression & Pre-programmed Poses

Expression screen

Sensors: Motion sensor, light sensor

Plan:

Mood Recognition detects emotion with CNN, displays the pre-programmed expression to boost user's mood if depression symptoms are detected.

How it works? Detect user location. When user is away, the screen will be automatically off to save energy.

Robot Arms

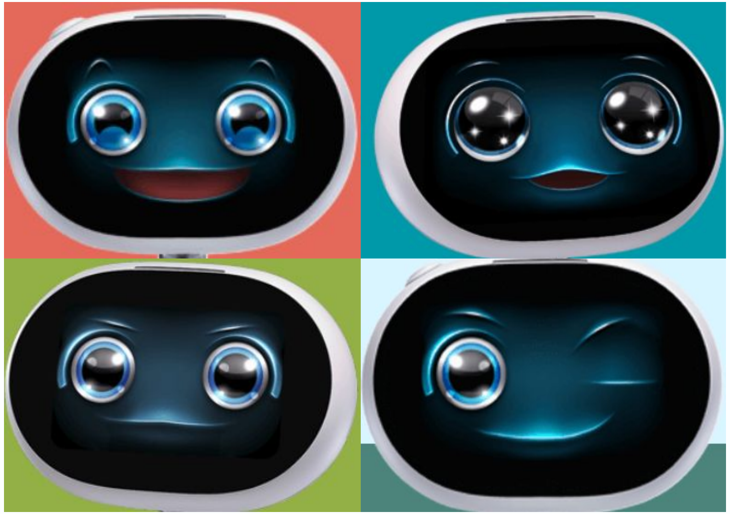
Sensors: Obstacle detection

Plan:

Imitation learning: To map human pose to robot pose by artificial neural network mapping model

Pose estimation: Analyze speaker's emotion from the speech content and act the pre-programmed poses with robot arms

How it works? Detect content and emotion -> Map to the pre-programmed poses -> Generate arm movements



Zenbo: https://zenbo.asus.com/whatsnew/features/smart-home/

Task 2: Companion Robot

Task	Sensor	Plan	Actuator
Chatting	Voice Detection Camera (Gesture recognition)	Text/Mood Recognition (NLP, CNN) Text Generation (NLP) Search Engine Memory	Speaker Expression Display
Music Display	Voice Detection Camera (Gesture recognition)	Text/Mood Recognition (NLP, CNN) Text Generation (NLP) Search Engine Recommender System	Speaker Touchscreen
Food Order	Voice Detection Touchscreen Camera (Gesture recognition)	Text/Mood Recognition (NLP, CNN) Text Generation (NLP) Search Engine Recommender System	(Through App) Touchscreen
Emergency Call (Families, Doctors, Police)	Voice Detection Camera (Gesture recognition)	Text Recognition/Generation (NLP) Phone Call Matching	(Through API)
Dance Coaching	Camera Obstacle Detection Touch Sensor	Text Recognition/Generation (NLP) Pose Simulation	Robot arms Twistable base Wheels AR (Augmented Reality)

Task 3: Health Diagnosis

Performance Measure:

Healthy patient, accurate diagnosis result, lawsuits, minimize diagnosis cost

Environment:

Partially Observable, Stochastic, Sequential, dynamic, Continuous, Single-Agent

Sensors:

Questionnaire, Clinical Data, Screening Results

Actuators:

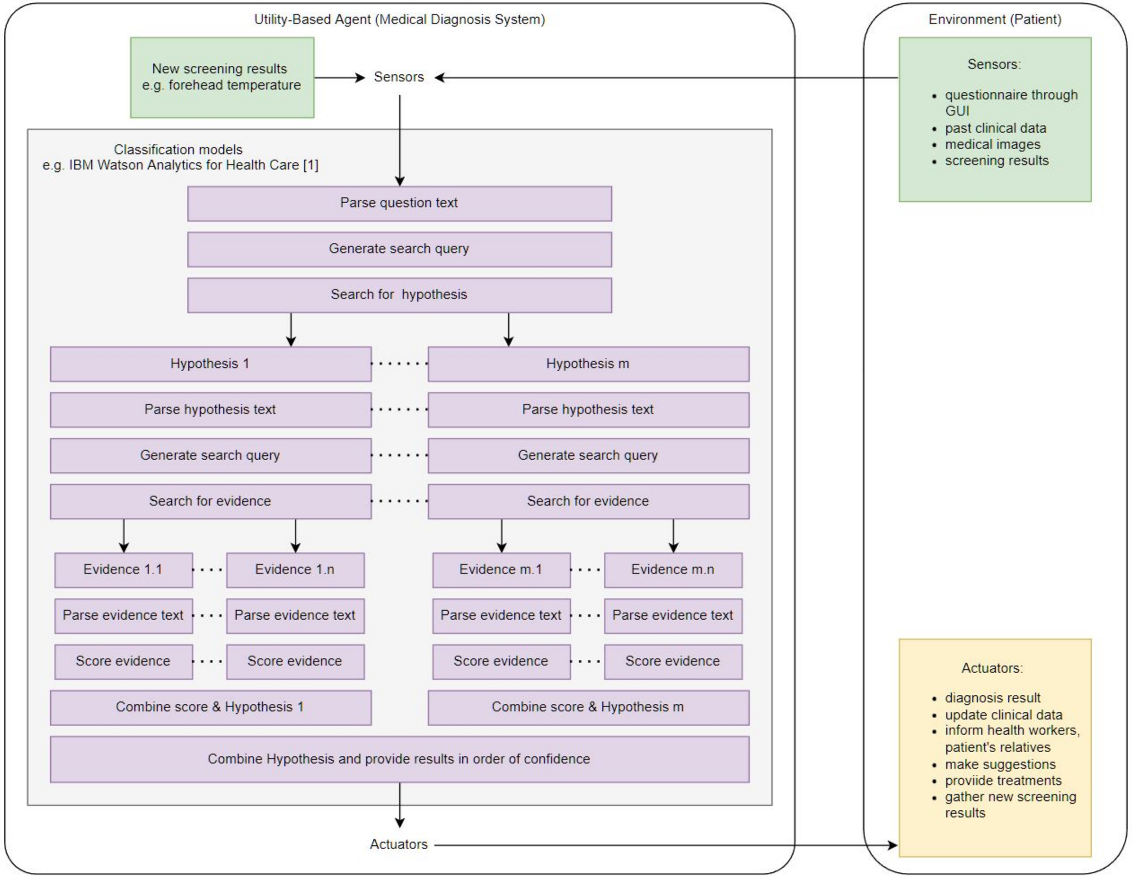
Diagnosis Results, Provide Treatments, etc

Type of Agent:

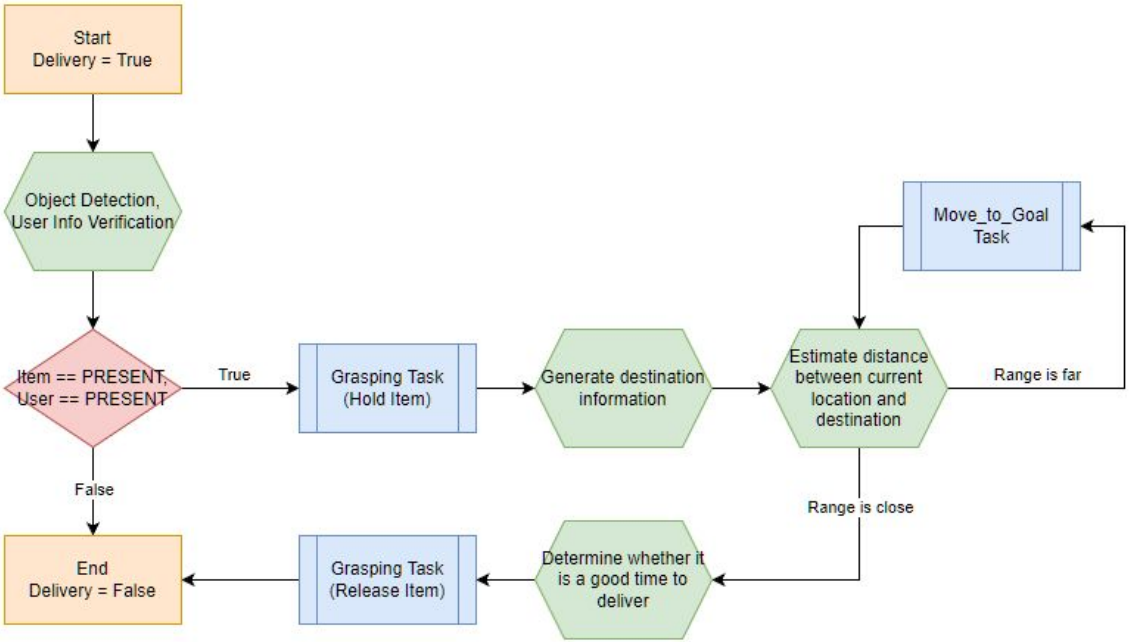
Utility-based agent with classification models

HRI/Roboethics:

Store patient information locally
Provide harmless treatment
Make persuasive suggestions as a peer role



Task 4-1: Delivery



Objective:

Search for correct item, pick up, carry to user

Environment:

Consisting of lots of different items,
obstacles, target users

Hardware & Sensing:

Color Video Camera, Infrared Sensors,
Gripper & Manipulator

Low-level Behaviours:

Perceptual schema:

- Object Detection, Object Localization, Object Pose Estimation, Grasping Estimation

Motor schema:

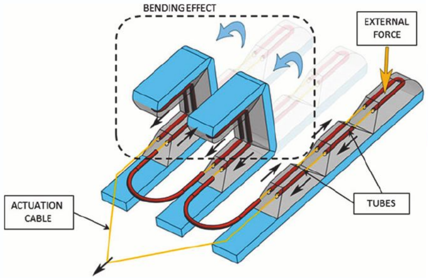
- Motion Planning, Execute Manipulator & Gripper, Move to Goal

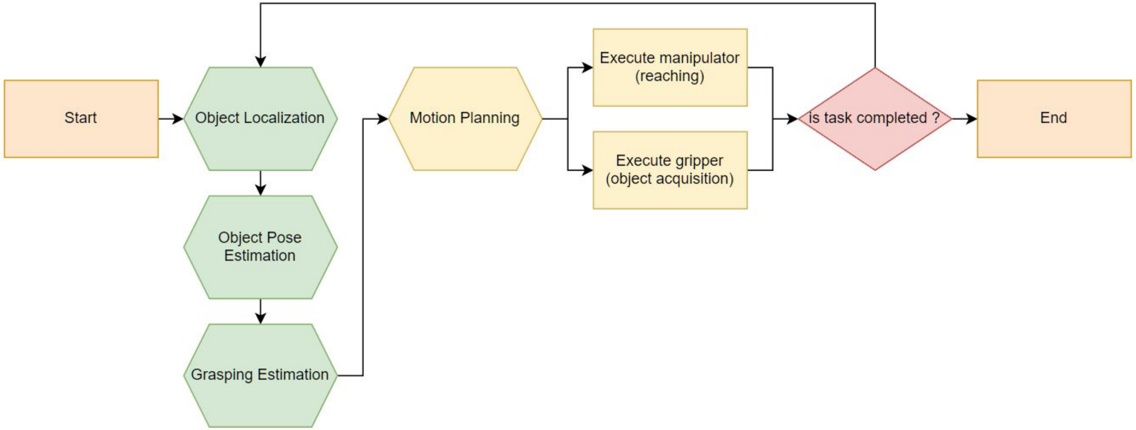
HRI Awareness:

Deliver the item to the user at the appropriate time
Understand user behaviour through pose, body language and voice

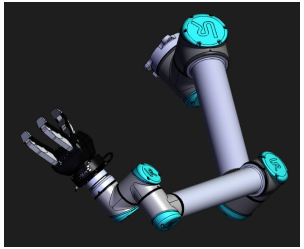
Task 4-2: Grasping Mechanism (Low-Level Behaviours)

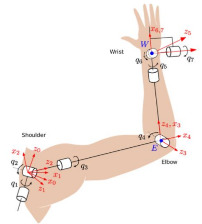
Bio-inspired 3 jointed cable cable-driven soft gripper





Bio-inspired 7 DOF Robotic Manipulator





Low-level Behaviours:

Perceptual schema:

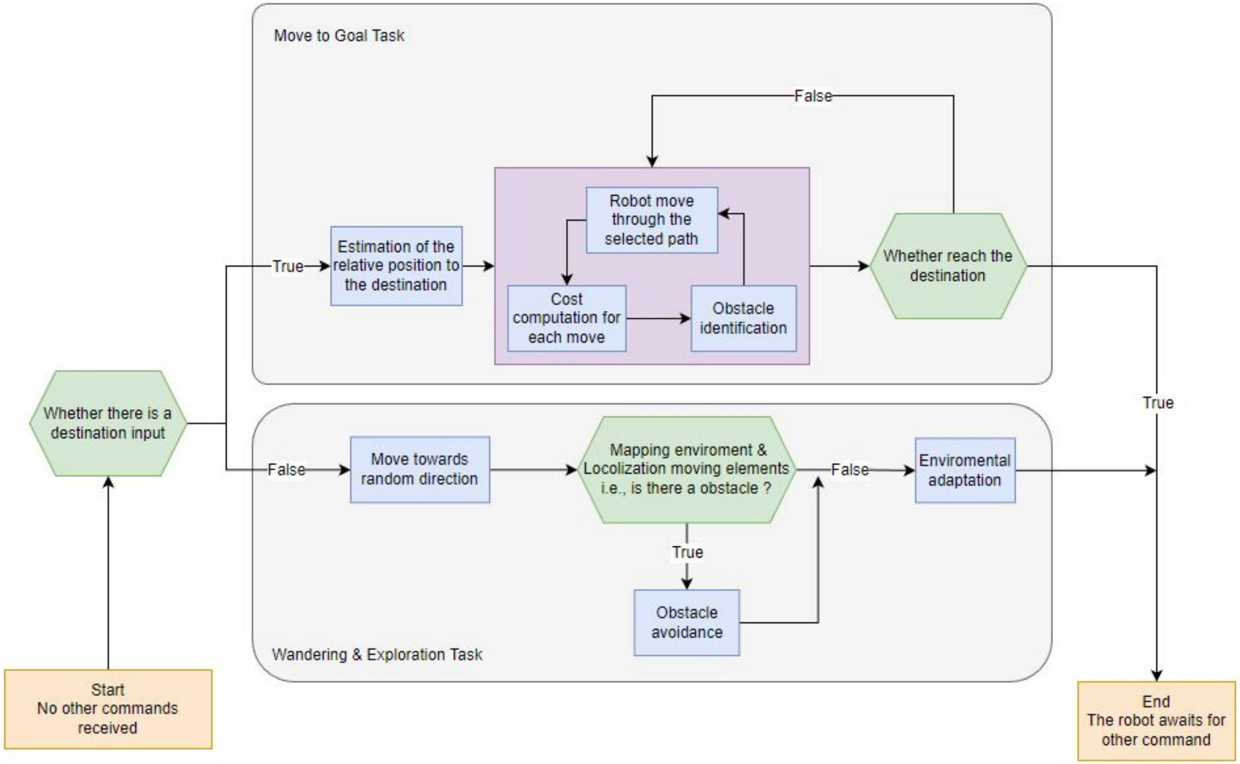
- Object Detection, Object Localization, Object Pose Estimation, Grasping Estimation

Motor schema:

- Motion Planning, Execute Manipulator & Gripper, Move to Goal

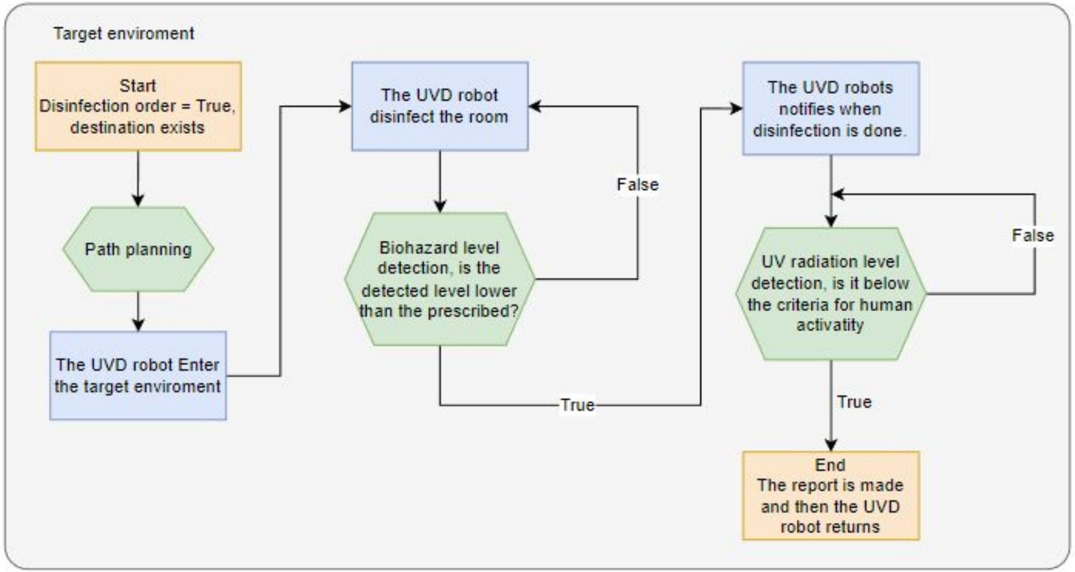
Task 5: Navigation & Wondering

- **Objective:**
  Move to goal;
- **Schema:**
  position estimation, path planning;
- **Objective:**
  Wandering;
- **Schema:**
  Representation construction of the environment, obstacle avoidance;



Task 6: Disinfection

- **Objective:**
    - Disinfection with UV;
    - Monitoring the disinfection quality and the remaining radiation level;
- **Perceptual schema**
    - Biohazard level detection;
    - UV radiation detection;
- **Motor schema**
    - UV power light;



Contents

Roboethic Concern & Design

1. Diagnosis and clinical data is stored locally to ensure cybersecurity.
2. Treatments generated from the diagnosis system must be verified by certified health workers before being recommended to patients to ensure patient safety.
3. Cartoonish expression design is to prevent the uncanny valley effect.
4. Pre-programmed robot poses is to prevent the harmful actions.
5. Include voice recognition and gesture recognition to support various disability groups.


THANK YOU !