METRICS HEART-MET
Physically Assistive Robot Challenge

General Description

In this challenge, the robot is tasked with assisting a person at home by fetching an item for them. The person communicates to the robot which item they would like, and where it is located. The robot should navigate to the given location, find the item, grasp it, and bring it back to the person and safely hand it over to them. The task is based on several functional and task benchmarks from the HEART-MET (Healthcare Robotics Technologies -- Metrified) competition developed in the context of the EU-funded project METRICS (Metrological Evaluation and Testing of Robots in International Competitions), and builds on previous competitions such as RoCKIn and RockEU2.

Main Scientific Challenge

The main scientific challenges addressed in this challenge are socially acceptable interruption, object perception and safe and socially acceptable physical human-robot interaction and collaboration for object handovers. Additional functionalities such as speech recognition, navigation and object manipulation are also required to successfully complete the task.

Assistive robots in human-centred environments may need to interrupt humans who may already be engaged in conversation. The interruption must be performed in a socially acceptable manner; for example by approaching the persons without surprising them and waiting for a pause in their conversation before interrupting.

Additionally, robots which operate in homes or healthcare facilities may interact with persons with cognitive or physical impairments. This could lead to failed verbal or physical interactions with the person which are caused, for example, by the person not responding to the robot, incorrect coordinating during a handover, incorrect modality used for communication, etc.

In unstructured environments such as a home, robots must be able to locate and manipulate objects in cluttered environments. In this episode, the robot must find the target item in a given semantic location, such as the living room, and thus requires the skill to actively *search* for objects, in addition to general object perception capabilities.

Platforms Allowed

A fully autonomous mobile robot with the capability for mobile manipulation, visual perception and speech recognition and generation is required for this episode. The robot must be able to communicate with a person using speech and gestures, and be able to navigate autonomously in a home-like environment. The robot must be able to perceive and manipulate objects at a minimum height of 0.0 m up to a maximum height of 1.2 m. In case of an emergency, the robot must have an easily visible and accessible emergency stop button to shut off all motors of the robot.

Scenario Setup

The arena will represent an apartment, with areas such as a living room, dining room, bedroom, etc. A sample studio apartment at the Cobot Maker Space, University of Nottingham is shown in Fig. 1. The areas may be separated by walls with virtual or real boundaries defined for each area, with the entire apartment covering a minimum area of 64 m^2. The apartment will have furniture such as tables, chairs, sofa, cupboard, etc. Typical household objects, such as kitchen items (plates, cups, bowls, cutlery, food and drinks), personal items (clothes, glasses, shoes), and other miscellaneous items such as books, computers, cellphones, etc., will be present in the apartment.

Fig 1. Sample apartment at the Cobot Maker Space, University of Nottingham

Procedure

There are three main phases for a trial in this episode: 1) communication with the person regarding the item and its location, 2) fetching the item, and 3) handing over the item to the person. The desired outcome for each phase is specified, and is used for scoring the achievements of the robot.

Preparations and start of the trial

- Two volunteers are guided by the referees to sit, stand or lay down in an area of the apartment and engage in a conversation.
- The robot is positioned in the apartment in the same semantic location, **A**, as the persons.
- Several objects are placed by the referees in different locations in the apartment on different surfaces such as tables, shelves, the floor, and chairs. This includes some of the predefined target items (known to the robot beforehand), and other miscellaneous objects.
- Other volunteers may be present in the apartment, and could be moving around.
- The trial begins when a signal is sent to the robot to start.

Phase 1: Communicate with the person

- In this phase, the robot needs to retrieve two pieces of information from a person, namely, the target item to be fetched and the semantic location of the item (e.g. dining table, shelf, coffee table, couch, etc.).
- Once the robot receives a signal to start, it can either a) wait to be summoned by a person in location **A** (e.g. by waving or speech) before approaching them, or b) start to approach the persons in location **A**.
- The robot reaches the persons at **A**, interrupting their conversation in a socially acceptable manner, and asking them whether they would like the robot to bring them something.
- One of the persons responds to the robot, but may choose to react in one of three ways:
    - The person responds by speaking to the robot
        - The person tells the robot which item they would like
        - The robot asks the person where the item is located
        - The person tells the robot the location of the item (for example, in the dining room, or on the kitchen counter)
    - The person does not reply to the robot, but gestures that they would like something (for example, by nodding their head, or showing a thumbs up sign)
        - The robot continues the dialogue via speech, and recognizing gestures performed by the person
        - The person may use gestures such as nodding, shaking their head, thumbs up, thumbs down, stop sign, pointing, etc.
    - The person responds by speaking to the robot, but may also use gestures during the remainder of the interaction
        - The person tells the robot which item they would like
        - The robot asks the person where the item is located
        - The person tells the robot the location of the item (for example, in the dining room, or on the kitchen counter)
        - Gestures may be used to respond to the robot, depending on the question

Desired Outcome: The robot announces that it is going to fetch item **X** from the location **B**

Phase 2: Fetch the item

- In this phase the robot must fetch the requested item from the given location.
- The robot moves autonomously to the location while navigating around persons who may also be moving around in the apartment.
- The robot searches in the location for the target item.
- Once the item is located, the robot approaches it, and grasps it.

Desired Outcome: The robot has possession of the item **X** in location **B**

Phase 3: Hand over the item

- The robot navigates back to location **A**
- The robot approaches the person who requested the item
- In case the person is busy, the robot interrupts them in a socially acceptable manner and announces that it has brought the item and is about to hand it over to the person
- The robot initiates the handover
- The person may respond in different ways
    - Ignores the robot
    - Reaches out but does not grasp the item
    - Reaches out and grasps the item safely
    - Reaches out and grasps the item, but drops it during or immediately after the handover
- In case the robot is still holding the item after the attempted handover, the robot should reinitiate the handover after announcing this to the person
- In case the object has been dropped, the robot must inform the person that the object has been dropped and ask whether they want the robot to recover it, and hand it over again

Desired Outcome: The robot and the person are in location **A**, and the person has possession of item **X**

Timing

The maximum time for a single trial of the episode is 20 minutes. The time starts when the robot is signalled to start the task, and stops when the item is successfully handed over to the person.

Score

The episode is scored based on technical achievements, penalties and a social acceptance score. The technical achievements are based on the desired outcomes of each phase. Penalties and disqualifying behaviours are undesired behaviours of the robot.

The social acceptance score is based on a questionnaire answered by the volunteers who interact with the robot, and additional persons who observed the robot performing the task.

Technical Achievements (TA)

1. Phase 1: The robot confirms it has completed this phase by saying “I am going to **Y** to fetch **X**”
2. Phase 2: The target item is in the robot’s possession (e.g grasped, or on a tray on the robot)
3. Phase 3: The target item is in the person’s possession

Penalties (P)

1. The robot collides with obstacles (such as furniture, walls)
2. The robot picks the wrong item
3. The robot drops the item
4. The robot stops responding
5. The robot stops working

Disqualifying behaviours (D)

1. The robot hits a person
2. The robot collides with and damages obstacles
3. The robot damages the item

The final Technical Score (TS) is computed as **TS = (TA - P)**. A disqualifying behaviour results in a Technical Score of 0.

Social Acceptance Score

The social acceptability of the robot is assessed through a questionnaire answered by the volunteers who interacted with the robot, and/or additional persons (such as experts, spectators, referees) who watched the episode. Some guidelines for socially acceptable behaviour are listed below; the questionnaire, based on works such as [1] and [2], will be formulated primarily to account for these guidelines

1. The robot approaches the two persons so as not to surprise or shock them, and maintains an acceptable distance until it interrupts them.
2. The robot interrupts the persons at an acceptable time, for example, when there is a pause in the conversation, or when they direct their attention to the robot.
3. Once a person responds, the robot addresses its attention to them, for example, by turning towards them.
4. The robot announces its intentions and current status, when appropriate
5. The robot navigates in the environment in a socially aware manner, for example, by avoiding getting in the way of persons and communicating intentions.
6. The handover occurs at a reasonable location, namely, at a reasonable height and distance from the person, who may be sitting, standing or laying down
7. The robot responds to a failed handover by announcing what has happened and asking for further instructions, if applicable.

Overall Score

The teams are ranked separately on their Technical Score and the Social Acceptance Score. The overall ranking is based on a weighted sum of the two rankings. For example, if team A is ranked M for the Technical Score and N for the Social Acceptance Score, an overall score is computed as S_{final} = (M + N) / 2. The teams are then ranked based on S_{final} with the lowest S_{final} being ranked first. In case of ties, more than one team can have the same rank.

References

[1] Kharub I, Lwin M, Khan A, Mubin O. Perceived Service Quality in HRI: Applying the SERVBOT Framework. Front Robot AI. 2021 Dec 13;8:746674. doi: 10.3389/frobt.2021.746674. PMID: 34966790; PMCID: PMC8711722.
[2] Schaefer KE. Measuring trust in human robot interactions: Development of the “trust perception scale-HRI”. In Robust intelligence and trust in autonomous systems 2016 (pp. 191-218). Springer, Boston, MA.

Appendix A

Objects

A subset of the YCB object set is used for this task. Only the following five objects are used in the task directly, meaning that the robot can be asked to fetch one of these objects:

- Windex bottle
- Spatula
- Pringles can
- T-shirt
- Campbell soup

The following objects may also be present in the apartment, in addition to other miscellaneous household objects, but are not expected to be grasped.

- Softball
- Tennis ball
- Spam can
- Plastic jug
- White rope

For all of the above objects, images and models can be found here.

Gestures

The list of possible gestures that can be used by the person interacting with the robot are as follows:

- Nodding
- Stop sign
- Thumbs down

- Waving
- Pointing
- Calling someone
- Thumbs up
- Waving someone away
- Shaking head