Homework 01

Klara Golubovic, K.M. Moshiur Rahman Songlap, Akhilan Ashokan, Daniele Affinita

Ungraded Task:

Task 1: Human-Centered Interaction in Robotics (21%)

1.1. Give an example of a human-robot interaction scenario. (5%)

One example for a human-robot interaction scenario can be a person throwing a ball at a robot to catch it. In this scenario the robot can catch the ball, but it does not throw the ball back to the person. Through a sound signal as for example calling the robot's name, the robot can be prepared for the upcoming ball to catch it. This is one of the examples we looked at in our partner work during the lecture.

1.2. Discuss any four human-centered aspects of interaction between humans and robots in the above scenario. (16%)

1) Verbal and Nonverbal Interaction: This would be the first aspect and concerning the mentioned example it would express nonverbal interaction. The robot does not speak in this scenario. Instead, by catching the ball it involves body movement. Additionally, the communication is one-way as the robot is trained to catch the ball and not to throw it. Because of that, the robot never initiates the game, it only reacts by catch the ball. When the robot catches the ball, the communication is over and the person must retrieve the ball to throw it again.
2) Understanding Human Behavior: In this example, the robot must understand the *multimodal expression* of the person. The robot should know when the person starts to throw the ball and needs to recognize *context-based* when a person is ready to start the game. To catch the ball, the robot must focus on the ball and how it has been thrown. Depending on the configuration of the robot, there is a specific sound signal, such as a call, to inform the robot that the ball is about to be thrown at it. In such case, the robot also needs to understand the person concerning *interpersonal differences*. Different people have different voices and may throw the ball in a different way. This aspect can be extended to the fact that the same person can call or throw differently each time. This aspect is a part of the *intrapersonal* difference in expression.
3) Learning from humans: The aspect of demonstration can help to develop a strategy for the robot to catch the ball in the best way possible. By analyzing how humans catch a ball from different heights and distances, the robot can learn through imitation. Moreover, the robot might be able to learn through evaluative feedback. If the robot can recognize when it has caught the ball, feedback might help it to learn and to improve. Possibly with the help of AI, the robot can change its behavior through these experiences and train itself to improve the catch.
4) Understanding Robot's Capabilities and Intentions: To clarify, in this scenario, the robot can catch the ball and hold it. It cannot throw the ball back to the person nor drop it after catching it. But the person can take the ball from the robot's "hands" after it was caught. The robot will catch the ball after the person gives a verbal sign, and it will recognize the ball flying towards it. The robot moves its arm and fingers to catch the ball. Therefore, behavior is on-request and nonverbal.

Task 2: Robot Autonomy v/s Human Control (24%)

Read the following paper carefully:

Ben Shneiderman (2020) Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy, International Journal of Human–Computer Interaction, 36:6, 495-504, DOI: 10.1080/10447318.2020.1741118

Imagine that you are designing an autonomous drone for supporting the fire- fighters in search and rescue missions.

Consider the 2D framework of robot-autonomy and human-control presented in the above-mentioned paper and answer the following questions:

2.1. Give an example of a human-drone interaction design for each of the following categories (8%):

a) low robot autonomy + low human control

b) low robot autonomy + high human control

c) high robot autonomy + low human control

d) high robot autonomy + high human control

a) **Low robot autonomy + low human control (Remote Controlled Drone)**

Example 1: In this case, it is like a toy drone, where the humans can control it (which is low autonomy) and cannot efficiently get the output from the available resources. i.e. having a proper control over the maneuver (low human control).

Example 2: Pizza delivery drone. The drone receives orders from an online application and delivers food to customers. In doing so, the humans don't control the drone at all (low human control) as the drone follows a fixed path (just a straight line), whereas the drone cannot plan its goal (the customer position) so there isn't any robot autonomy.

b) **low robot autonomy + high human control (Drone used for surveillance and filming)**

It is like a surveillance drone, where humans can maneuver the drone for surveillance, yet it will have an inbuilt hover and soft-landing mechanism (low robot autonomy) and humans can have a high control to maneuver the drone without any difficulties. (High human control)

c) **high robot autonomy + low human control (fertility spraying/programable drone)**

It is like using a drone which is used for fertilizer spraying, where the drone autonomously identifies the field area and sprays the fertilizers to the field. (high robot autonomy). Humans merely turn on the drone and just do the surveillance in emergency situations. (Low human control)

d) **high robot autonomy + high human control (swarm robots)**

Use of swarm robots has high robot autonomy as it can communicate within the robots to avoid collision (high robot autonomy), where humans must have greater control to make the swarm robots to do certain activities.

2.2. Explain the pros and cons of each design you mentioned above (16%).

1. Low robot autonomy + low human control (Remote Controlled Drone):

- Pros Example 1: Humans can have fun maneuvering a drone and having control to actuate it according to their wishes. In addition, the human can also modify the goal on the fly.
- Pros Example 2: No need for employees to deliver the food. Easy technical implementation as there are neither planning to implement nor complex decisions.
- Cons Example 1: It has no advanced feature like hovering the drone, so it will be difficult to maneuver. The operator must focus on the drone control to avoid potential crashes.
- Cons Example 2: No reliability nor robustness. The drone cannot face unexpected situations.

2. Low robot autonomy + high human control (Drone used for surveillance and filming)

- Pros: Humans have high control to maneuver, collect data and make predictions according to it. The human can handle trajectories also in an unknown environment.
- Cons: Data filtering, and object detection all those things has to be completed manually

3. High robot autonomy + low human control (fertility spraying/programable drone)

- Pros: Robot have high knowledge of actuating according to the environment as they are preprogrammed
- Cons: When there is a chance of failure in detection of the field there could be a chance of wasting the resource. It's impossible to detect and correct false positives so the detection must be really accurate.

4. High robot autonomy + high human control (swarm robots)

- Pros: Better performance rate as it as better communication between the drones. This approach is a good tradeoff between autonomy and controllability.
- Cons: More data has to be monitored regularly. Robot self-coordination requires a communication channel (network).

Task 3: Reading – Human-Robot Interaction (HRI) (10%)

Read the following paper and write a scientific summary in 300 - 500 words.

P. Vogt et al., "Second Language Tutoring Using Social Robots: A Large- Scale Study," 2019 14th ACM/IEEE International Conference on Human- Robot Interaction (HRI), 2019, pp. 497-505, doi: 10.1109/HRI.2019.8673077.

Please note that this task would help you in completing Task 4

Hint: You can check the following articles for tips on how to write scientific summaries:

https://www.lib.sfu.ca/about/branches-depts/slc/writing/sources/summarizing

https://writingcenter.uconn.edu/wp-content/uploads/sites/593/2014/06/How_to_Summarize_a_Research_Article1.pdf

This paper discusses the effectiveness of social robots as teaching tools that help young preschool children learn second languages. It highlights that because social robots may display socially supportive behavior, they can deliver efficient one-on-one tutoring that can result in greater learning gains than group-based education. In short-term studies, the effectiveness of social robots in education may be better which is why long-term research is needed to evaluate properly. Research on human-robot interaction can be time-consuming and expensive, especially in the educational sector. As for educational robots, long-term studies are crucial because of skills needed for topics like a second language that need for repetition and patience. They also address some conflicting findings from earlier research on the efficiency of robots in many contacts for language learning.

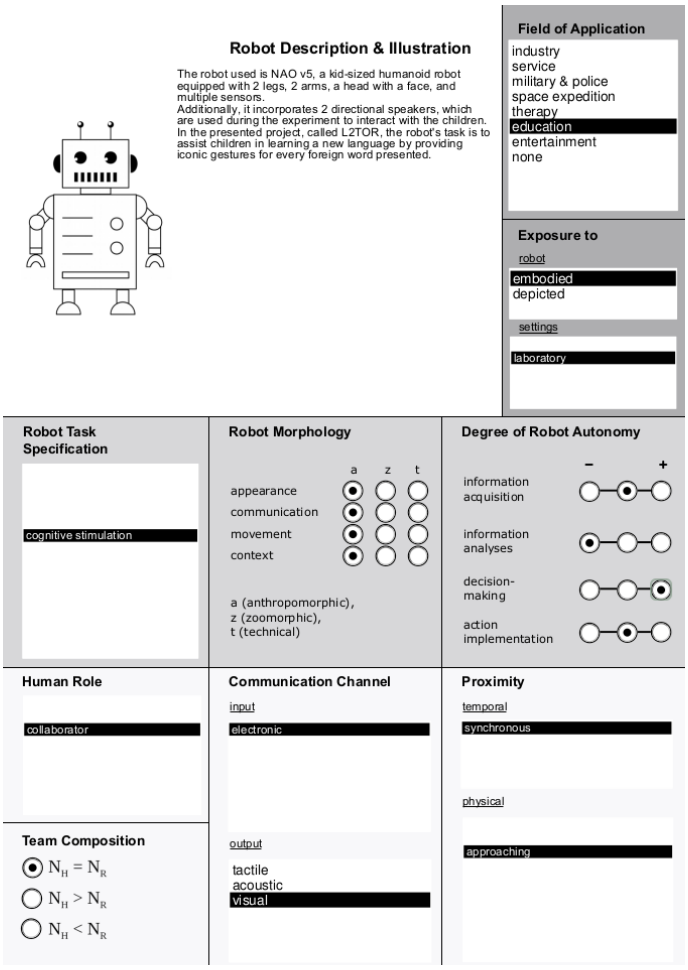
The study also presents that face-to-face interaction with robots provides a better learning foundation than systems with screen applications. It has some significance of both deictic and iconic gestures in the acquisition of new languages as well and robots use these gestures efficiently. For learning vocabulary in a foreign language, robots use both with and without gesture systems which affects young children. The study has four different conditions and a well-powered sample size, one of which involves a control group that dances with the robot without any instructional information. Several theories can be evaluated based on the efficiency of robots and tablets in terms of learning a language.

In this study, to determine the effects of various circumstances on children's language learning capacities researchers perform some examinations. Four criteria and several assessments of word knowledge are used for their study. According to the findings of MANOVA and chi-square tests, they divided children into four groups with similar age, gender, Dutch vocabulary, phonological memory, selective attention, and prior knowledge of the target words. Table III shows the key results from word knowledge tests. The result shows kids considerably outperformed zero on two post-tests and a pre-test that involved translating from English to Dutch. For all conditions of the control group and translations scores, in the immediate post-test were higher than in the pre-test. As well as post-test scores is greater than immediate post-test score with some delay. That shows the primary impact of time while indicating word consolidation with time. The comparison task results performed well compared to translation and chance level. Children under experimental circumstances performed poorly in the control group in all tasks, based on a repeated measures analysis, demonstrating a substantial main impact of the condition. There are no discernible variations existed between the various testing settings. In addition, with, delayed post-test scores are greater than immediate post-test scores which indicate word consolidation with time. Children with bigger vocabularies learned more English words, according to a second study that used Dutch receptive vocabulary and phonological memory as control variables. However, there are no discernible effects of phonological memory or interaction effects are discovered. These findings show that environmental factors greatly influenced word learning and that children with larger Dutch receptive vocabularies learned English words more successfully.

Task 4: Classify HRI Study (25%)

Apply the taxonomy developed by Onnasch and Roesler to describe the condition “1) Robot with iconic gestures + tablet” in Section II C of the paper that you read in Task 3 (10%) and explain your choice of categories (15%).

Note: Please use the HRI canvas provided in supplementary materials. Fill this canvas and submit it along with your solution PDF. If you have trouble in editing this file, then you can provide your answer directly in the solution PDF. However, explanations for your choice should be written in the solution PDF.



The field of application is clearly Education. Indeed, the experiment's main goal was to teach children new foreign words, supporting the learning process with robot gestures. The robot is embodied because it has a physical body (hardware). The setting can be considered a laboratory since it was designed specifically for conducting the experiment.

The robot's task can be described as “cognitive stimulation”, indeed its primary goal is to assist the children during the learning process. Due to the humanoid nature of the robot, its morphology can be considered humanoid mimicking human communication and movements.

In general, the robot is not fully autonomous in this context, the information is retrieved directly from the tablet and the children's speech was provided by a human operator.

“We had to rely on Wizard of Oz (WoZ) to indicate whether the child had said something, because neither automatic speech recognition nor automatic voice activity detection worked sufficiently reliably.” (Second Language Tutoring using Social Robots: A Large-Scale Study, page 499 section B)

So, we can say that the information is not processed at all. On the other hand, the robot can make decisions by reacting to the information provided. For instance, it can introduce the children to the lesson, correct the child if they make a mistake more than twice, and use iconic gestures to help the child understand the meaning of the word. Therefore, the robot is capable of differentiating its actions based on the situation. Finally, the robot can implement the actions (I.e. gestures) effectively and autonomously.

In this experiment, there is a one to one interaction between the child and the robot. The child can be considered a collaborator, as both the human and the robot share the same goal, which is to help the child learn new foreign words. They work together to achieve this goal, with the robot reacting to the child's choices, and the child observing the robot's gestures.

The robot communicates with the child through gestures and speakers while receiving inputs from the tablet and a human operator (WoZ). The interaction occurs synchronously in the same physical space, but there is no physical contact between them.

Task 5: Realize Behaviors on Pepper (15%)

In this task, you will use the qibullet simulation tool to complete the following subtasks:

1. Implement a behaviour that enables Pepper robot to speak a given text. (5%)

2. Implement a waving gesture on Pepper robot. (5%)
3. Add comments in your code to explain your logic. (5%)

Submission:

- Along with your solution PDF, please upload the source code that you have written or used for solving the task. You do not need to embed code snippets in the solution PDF. But you should comment your source code appropriately.

- Please also upload the screen recording demonstrating the above behaviours on Pepper.

Instructions to install and use qiBullet and Pepper robot simulation:

- To install and use qiBullet, please follow the instructions and examples provided in this github repository: https://github.com/softbankrobotics- research/qibullet.

- See also the tutorial at https://github.com/softbankrobotics-research/qibullet/wiki/Tutorials:- Virtual-Robot

- For the names of Pepper’s joints, see here: https://github.com/softbankrobotics-research/qibullet/blob/75c3669a03316 38bf195a50d42edb51f9da2edd2/qibullet/robot_posture.py

- A skeleton code has been uploaded in LEA to help you with loading and connecting to Pepper in qibullet.

Feedback:

Please answer the following:

1. How much time did you spend on doing this sheet per person? Anonymize your answer! 1 hour per person
2. Was this sheet too easy / easy / ok / hard / too hard? Ok
3. What additional resources (blogs, papers, books, tutorials, etc.) did you use? Please provide links or references.

   QiBullet Documentation

   ChatGPT just for spelling and grammar corrections.

4. Did you face any issues while solving this sheet? Nothing in particular

Submission Procedure:

Upload the PDF of your solutions and the relevant source code files in LEA as a single Zip archive. For the naming convention for your submission, please follow the instructions under Course Rules in LEA.