Human-Robot Plan Communication Strategies

Rahil Rajaa, Layla Alarifi, Walter Bond, Marius C. Silaghi
(Florida Institute of Technology, USA)

***Abstract*—In this research, the problem we address concerns finding the better mechanisms to communicate to a robot the directions for navigation in indoor environments. We identify which out of a set of combinations of speech, gestures, and drawing mechanisms are the most comfortable, easier to learn, and least error-prone for human users. Three different methods: a Speaking method, a Gesturing and Speaking method, and a Drawing method were tested for guiding the NAO robot to the desired place, assuming that the robot has no prior map of the building and even no information about the involved human faces. Our experiment consists of having participants ask the robot to accomplish a complex indoor navigation task. The task communication is attempted with each of the above methods. In one experiment, no additional details were provided about the proposed methods to help the participants, except for a video showing the problem in neuter human terms. In subsequent tests, the participants were also given restrictions concerning the robot communication capabilities: only to use speech, gestures, or vision. In the third set of tests, the participants were also given training using examples of efficient communication with corresponding methods. The task was decomposed, with points assigned for each component. The methods have been investigated and evaluated based on the average task success in simulated plan execution. In our results the Drawing method leads to the highest level of task accomplishment, and it is also the fastest communication method. Further, according to the final short questionnaire, most participants feel that the Drawing method is more comfortable and they can use it immediately without thinking, as compared to methods based on speech.**

I. Introduction

In human-robot teamwork, a human can communicate his intent to the robot via explicit communication such as speech, gesture, and haptics. Speech provides the information in the form of utterances of words and sentences, while gestures can provide the information through the head, arm, and body gestures. As well, head gestures represent facial expressions and eye gaze direction [1]. Robots, especially service robots, may encounter problems during the navigation process. The common challenges are [2]:

1) High localization uncertainty.
2) Unreachable goal positions.
3) Collisions with obstacles.

Many researchers have dealt with these issues and provided various solutions such as involving humans to solve uncertainty and to optimize the robot efficiency [3], [4], [5], [6].

With the increasing use of autonomous robots in the human environment, the capabilities of robots have been enhanced, and the navigation capability has particularly increased. So, many mobile robots use robust navigation strategies to move around in an indoor and outdoor environment without collisions. These robots, such as tour guides [7], [8], [9], interact with the human in their surroundings, assist and guide them, and provide them with the required information. In contrast, when the robot has no previous spatial details to do a specific navigation task, it needs human support to obtain information for moving from one place to another without colliding with obstacles and walls, or going into a wrong direction.

In a human-robot interaction scenario, we study three ways to communicate with the robot for navigation in indoor environments: speech, gestures, and drawings. Our experiment lets the participants guide the robot to a target with various methods and degrees of training. The goal of this work is to identify which human-robot communication mechanism is more comfortable and easy for users, and lets them route the robot to the desired place with fewer errors.

The structure of the rest of this paper is as follows: After brief details about related work in section II, Section III explains our experiment. Section IV presents result and discussion. Section V concludes our work.

II. Related Work

A robot appearance can affect human-robot interactions. For example human-like robots can motivate humans to interact with them as if they interacted with humans, enabling them to understand the robots gestures. This supports human-robot communication such as route-guidance [10]. Also, Cassell et al. [11] explained that the anthropomorphic expressions of embodied robots are improving human-robot communication. Weiss et al. [12] developed an autonomous mobile robot without GPS support. The robot navigates outdoors and asks people for directions. Peoples gestures (pointing) helped the robot to find the target. In [13], the researchers produced an office robot Jijo-2 that can ask people for help via the spoken-dialogue system to learn office locations. Kanda et al. [14] conducted a study to prove that the robot can learn his route by listening to the human teacher. The human can use gesturing and speech to provide the information to the robot. Also, he can communicate with the robot by using spatial datum to navigate the robot. The spacial information can improve the robots awareness of the surrounding environment [15], [16]. To sketch the path of the robot in the robot guidance, Chronis et al. [17] proposed a method that allowed users to draw a path for the robot and all objects and obstacles around this path. Users used an interface for sketching on a PDA, which in turn extracts the prominent landmarks for the robot route. Boniardi and his colleagues in [18] used a sketch interface to enable users to draw an abstract map with the important landmarks, and the robot path.

III. THE EXPERIMENT

The experiment consists of three parts. In each part, the participants should ask the robot to accomplish the corresponding task. Since the robot had no prior knowledge of the building map, participants were asked to give directions to the robot. As well, we gave each participant the plan that included explanations and instructions that they were to perform.

In general, our study aims to find a way to compare methods to help the robot reach intended places. To achieve this goal, we have introduced three methods:

1) Speaking: The participants involved with this method communicate with the robot through voiced utterances, where they can give directions to the robot to reach the target.
2) Gesturing and Speaking: Here, two methods have been integrated to increase the attention of the robot, where the participants can tell the robot about the desired direction by using their voice simultaneously with making gestures with their arms, hands, and fingers. The combination of the two methods is replicating what humans do when speaking spontaneously.
3) Drawing: In this method, the participants can draw the robot route on a map with predefined symbols for assistance. We suggest symbols that can help the robot understand operations that can occur when following directions towards a target. This method may be appropriate when there are a lot of details in the scene, or the human has difficulty in speaking the language of the robot. Participants were provided with the layout map, in Part 3, to depict all details in the scene, and to save the time for participants.

A. Task environment

The experiment scenario is indoors, being conducted in the Department of Computer Engineering and Sciences at Florida Institute of Technology (FIT), where corridors are used as a path for the robot to reach the target. Also, in one of the proposed methods we used a layout map of the second floor of the Harris Center for Science and Engineering building at FIT, where the experiment was conducted. The total distance of the robot path from the starting point to the goal is approximately 3418 cm.

B. The participants

In this study, we recruited 16 participants to apply our experiment (9 men, 7 women), and their ages range in the interval (23 - 39). Most of the participants are graduate participants in the computer engineering and sciences department, and they have different experience with robots. Six out of sixteen participants have a robot experience. Before starting the experiment, we informed each participant that we would record their work and keep it private, only using it for statistical studies, discarding any identifying information. The entire sessions with the participants were conducted in a private and quiet room, and each participant was engaged individually. Each session lasted around 20 minutes.

C. The experiment scenario

At the beginning of an experiment, each participant signed the consent form and then read the descriptions of what he should do (in 5 minutes), as we allowed participants to ask us if they need any additional explanations. Then, they filled a demographic information form. Moreover, the participant was shown a video explaining the intended route for the robot, and he was told that the robot was not familiar with the building and did not know the face of his friend.

In each part participants must give orders to the robot, to guide it. The task is described with the question:

"How can you give an order to the robot to bring the paper from your friend (Who's seat is described in the associated video) in the Ph.D. Lab 211?"

1) Part 1: This part intends to identify the most expected way for participants to communicate with the robot. So, no details are provided about the proposed methods and no example is given to help them. We asked them to give orders to the robot to based on the aforementioned question.

2) Part 2: In the second part, participants use each proposed method freely, without seeing examples, to give the orders to the robot, to accomplish the task in the given question. The methods tested in a randomized order are:

1) Speaking: To start this method, participants have to tell the robot, in English, any instruction they want. They are informed that the only possible input is their voice.
2) Gesturing and Speaking: In this method, participants can give commands to the robot using voice and gestures (i.e., gestures with arms, hands, and fingers).
3) Drawing: participants are asked to sketch the robot's path on a blank piece of paper, such that the robot can follow it.

We presented a short questionnaire to participants, after Part 2, which allows them to evaluate each method from the user perspective, choosing one option from Table I and writing the reasons behind their selection.

TABLE I: Evaluation options presented in the questionnaire after the second and third parts.

Evaluation options
1- I was able to do it immediately without thinking.
2- I need to think several seconds.
3- I need to think over a minute.
4- I had to make several trials.
5- I'll never manage to do it right.

3) Part 3: The same previous scenario is applied in this section, but after providing participants with recorded examples and a set of instructions for routing the robot. For the studied methods tested in a randomized order:

1) Speaking: After presenting the video, the participants were given a list of instructions they might prefer to use while telling the robot the directions such as:
   - Go to the (right/left).
   - Turn (right / left).
   - Pass through the door.

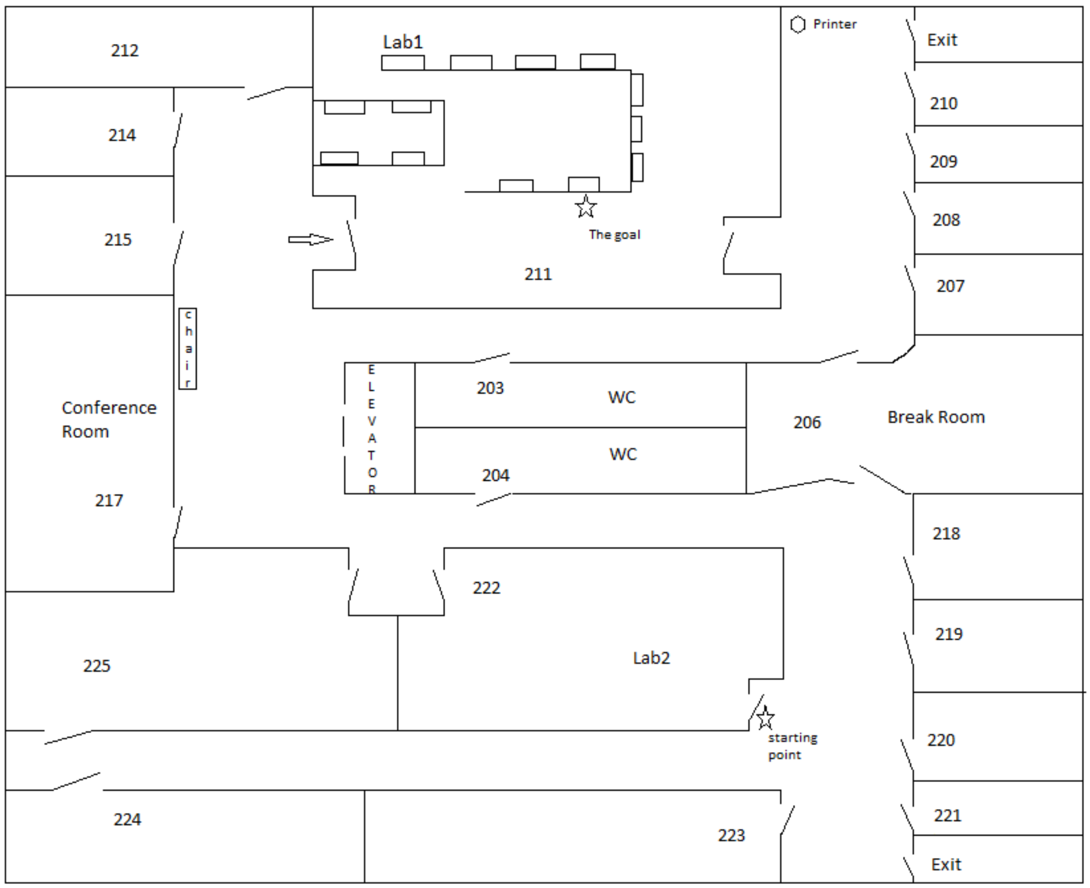

Fig. 1: Layout map of the second department floor.

- Pass by the door.
- Turn left / right [at the corner].
- Start with the back at the [door].
- Go straight until you arrive at [the wall/ the corner].
- Go straight along the wall on the (left/ right) (passing by Y)* until you
- arrive at X.
- Your destination is X.
- Return on the same path.
- If X then Y.

Moreover, participants can use their own words when talking to the robot.

2) Gesturing and Speaking: Participants used the same instructions in the Speaking method, to which gestures are added.
3) Drawing method: Initially, participants were provided with an illustrated map example. Then they penciled the robot path on a prepared layout map of the second department floor as shown in Figure 1, and used the symbols in Figure 2 to explain the direction.

Finally, the participants completed the short questionnaire to select the easiest method after displaying the recorded example; also they comment on why they would choose a method.

IV. Results and Discussions

Our trial consists of three parts, where a total of 16 participants were asked to guide a humanoid NAO robot v5 and help it to figure out the right directions to reach a target, by using the proposed methods, as we described in the experiment section. After Part 2 and Part 3 we asked everyone to assess these methods, by choosing one option from a set of ratings ranging from 1 to 5 (see Table I), through a short questionnaire. The total duration time of the experiment is about 20 minutes.

To evaluate our experiment, we imitated the robot route and tried to see how well the robot can successfully execute a person's guidance, and from the given results, we determined which method is best. The total distance that the robot is supposed to walk to the target if it goes on the safe path without errors is 3418 cm.

We suggested five types of errors that the robot might encounter during his task:

1) Hitting a wall.
2) Going in the wrong direction.
3) Hitting an obstacle.
4) Missing landmarks such as door, corridor, obstacles.
5) Not following the optimal path.

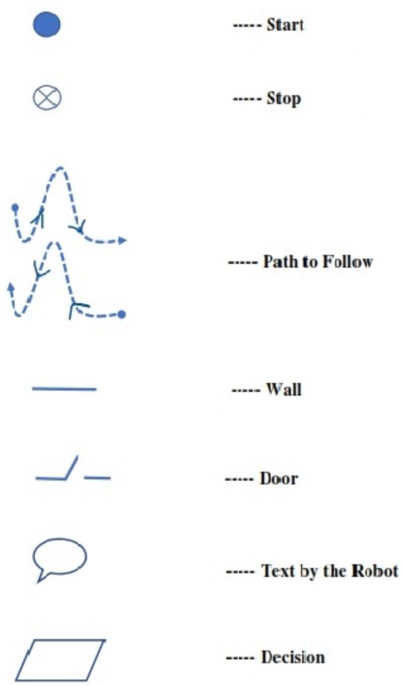

Fig. 2: The symbols to help participants to draw directions on a map.

Furthermore, we allocated a "No Recovery probability" to each type of error based on the importance of the impact of the error on the robot route (Table II).

TABLE II: The percentage of each error, the probability of No Recovery and the probability of success after each error type.

Error Type	Percentage	Pro. No Recovery	Pro. Success
Wall	50%	0.5	0.5
Direction	60%	0.6	0.4
Obstacle	60%	0.6	0.4
Miss	30%	0.3	0.7
Sub optimal	20%	0.2	0.8

Accordingly, we used the following formula to calculate the distance that the robot could take to reach the target with possible errors:

$Totaldistance = Distance + Psuccess * (Distance + Psuccess * (Distance + Psuccess * (...)))$.

To examine whether there was a difference between the distances traveled by the robot in the three methods (i.e., the performance of the participants before and after guidelines), we conducted ANOVA: Two-Factor Without Replication test on our data for Part 2 and Part 3.

The null hypothesis is:

H_0 : *There is no significant difference between the method types in terms of route length achieved.*

The alternative hypothesis:

H_1 : *There is a significant difference between at least two method types in terms of route length achieved.*

In both the first and the second parts we chose the default alpha value of 0.05 in the ANOVA test.

With the results for Part 2, since the $Fvalue$ exceeds the $Fcritical$ ($Fvalue = 55.56 > Fcrit = 3.32$) and $p - value = 8.18791E - 11 < \alpha = 0.05$, we reject the null hypothesis stating that there are no differences between the three methods.

Similarly in Part 3, since the $Fvalue$ exceeds the $Fcritical$ ($Fvalue = 103.12 > Fcrit = 3.32$) and $p - value = 3.60269E - 14 < \alpha = 0.05$, we reject the null hypothesis.

The results, in Part 2 and Part 3, showed a significant difference between the methods; also we can see that the results in Part3 are much better than the results in Part 2.

For further explanation, we plot a graphic shape to display the average value of each method in Part 2 and Part 3 from the ANOVA outputs, as shown in Figure 3. The figure shows the differences between the introduced methods. There is an insignificant difference between the M1 and M2 in both Part 2 and Part 3. However, we can see a significant difference between M1 and M2, and the M3. Also there is an important difference in the same third method in Part 2 and Part 3, as a result of participants' performance improvement after providing participants with the guidelines.

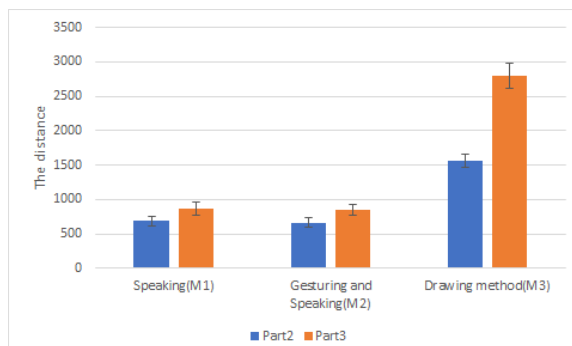

Fig. 3: The average distance for all methods in part2 (without seeing examples) and part3(with seeing examples).

A. Questionnaires

A short questionnaire was presented after Part 2 and Part 3, to find out the participants' opinions on the proposed methods, both before and after studying examples. The questionnaire contains five levels of evaluation to assess these methods (see Table I). A review of the questionnaire revealed that there was a divergence in the participants' views on these methods and that the opinions of some participants were not compatible with their corresponding performance.

Further, the questionnaire shows that, in Part 2, about 37.5% of participants preferred the Drawing method (level 1), and 18.5% of them preferred the two Speaking-based methods. While about 62.5% of participants chose level 2 for Speaking methods, 68.75% of them selected the same level for the Gesturing and Speaking method; the reason is that most participants considered that Speaking and Gesturing are natural for humans. On the other hand, while about 50% of participants choose the Drawing method as a somewhat difficult, most of them (about 68.75%) changed their minds in Part 3, where they saw the prepared map and the set of symbols.

V. CONCLUSION

In this paper, we aim to establish methodologies for measuring which communication mechanisms are better to guide the Nao robot to the desired place in an indoor environment, to do a specific task. Three different methods: Speaking, Gesturing and Speaking, and Drawing were introduced for a navigation task. It was assumed that the robot has no spatial information about its building and did not know the human faces. We recruited approximately 16 participants to join our three-part experiment. In the first part, we asked the participants to direct the robot to the specified place based on their preferred method. In Part 2 we limited the participants to each of the three proposed methods to guide the robot without seeing an example. Ultimately, we allowed the participants to use the methods in the navigation task after training with sample instructions and videos of related solutions.

We investigated and evaluated our experience, and the results show that participants achieved better results when they used the Drawing method to guide the robot. Also, our findings show that the difference is indistinguishable between the performance of the participants with M1 and M2. However, there is a considerable difference between M3 and the rest of the methods, in both Part 2 and Part 3. In addition, the participants' performance was improved after they were trained with an example. The short questionnaire results reveal that most participants, especially in Part 3, decided that the Drawing method is a more comfortable way, as compared to methods based on speech.

REFERENCES

1 Bauer, A., Wollherr, D., and Buss, M., "Human–robot collaboration: a survey," *International Journal of Humanoid Robotics*, vol. 5, no. 01, pp. 47–66, 2008.

2 Savkin, A. V., Matveev, A. S., Hoy, M., and Wang, C., *Safe robot navigation among moving and steady obstacles*. Butterworth-Heinemann, 2015.

3 Fong, T., Thorpe, C., and Baur, C., "Robot, asker of questions," *Robotics and Autonomous systems*, vol. 42, no. 3-4, pp. 235–243, 2003.

4 Bajones, M., Weiss, A., and Vincze, M., "Help, anyone? a user study for modeling robotic behavior to mitigate malfunctions with the help of the user," *arXiv preprint arXiv:1606.02547*, 2016.

5 Heger, F. W., Hiatt, L. M., Sellner, B., Simmons, R., and Singh, S., "Results in sliding autonomy for multi-robot spatial assembly," 2005.

6 Argall, B. D., Chernova, S., Veloso, M., and Browning, B., "A survey of robot learning from demonstration," *Robotics and autonomous systems*, vol. 57, no. 5, pp. 469–483, 2009.

7 Burgard, W., Cremers, A. B., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., Steiner, W., and Thrun, S., "Experiences with an interactive museum tour-guide robot," *Artificial intelligence*, vol. 114, no. 1-2, pp. 3–55, 1999.

8 Thrun, S., Beetz, M., Bennewitz, M., Burgard, W., Cremers, A. B., Dellaert, F., Fox, D., Haehnel, D., Rosenberg, C., Roy, N. *et al.*, "Probabilistic algorithms and the interactive museum tour-guide robot minerva," *The International Journal of Robotics Research*, vol. 19, no. 11, pp. 972–999, 2000.

9 Gross, H.-M., Boehme, H.-J., Schröter, C., Müller, S., König, A., Martin, C., Merten, M., and Bley, A., "Shopbot: Progress in developing an interactive mobile shopping assistant for everyday use," in *2008 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, 2008, pp. 3471–3478.

10 Ono, T., Imai, M., and Ishiguro, H., "A model of embodied communications with gestures between human and robots," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 23, no. 23, 2001.

11 Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H., and Yan, H., "Embodiment in conversational interfaces: Rea," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 1999, pp. 520–527.

12 Weiss, A., Igelsböck, J., Tscheligi, M., Bauer, A., Kühnlenz, K., Wollherr, D., and Buss, M., "Robots asking for directions: the willingness of passers-by to support robots," in *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 2010, pp. 23–30.

13 Asoh, H., Motomura, Y., Asano, F., Hara, I., Hayamizu, S., Itou, K., Kurita, T., Matsui, T., Vlassis, N., Bunschoten, R. *et al.*, "Jijo-2: An office robot that communicates and learns," *IEEE Intelligent Systems*, vol. 16, no. 5, pp. 46–55, 2001.

14 Kanda, T., Kamasima, M., Imai, M., Ono, T., Sakamoto, D., Ishiguro, H., and Anzai, Y., "A humanoid robot that pretends to listen to route guidance from a human," *Autonomous Robots*, vol. 22, no. 1, p. 87, 2007.

15 Zender, H., Mozos, O. M., Jensfelt, P., Kruijff, G.-J., and Burgard, W., "Conceptual spatial representations for indoor mobile robots," *Robotics and Autonomous Systems*, vol. 56, no. 6, pp. 493–502, 2008.

16 Hemachandra, S., Walter, M. R., Tellex, S., and Teller, S., "Learning spatial-semantic representations from natural language descriptions and scene classifications," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 2623–2630.

17 Chronis, G. and Skubic, M., "Robot navigation using qualitative landmark states from sketched route maps," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, vol. 2. IEEE, 2004, pp. 1530–1535.

18 Boniardi, F., Valada, A., Burgard, W., and Tipaldi, G. D., "Autonomous indoor robot navigation using a sketch interface for drawing maps and routes," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 2896–2901.