HKH:Helen Keller Heuristic,a Common Ground Scenario for Human-Robot Interaction

Mohammed A. El-Dosuky
Computer sciences Department, Faculty of Computers and Information
P.D.Box 35516, Mansoura University, Egypt

ABSTRACT

Lacking a common ground in Human-Robot Interaction is the main source of errors and miscommunications especially for remote exploration robotics. Advances in language and social psychology studies are the rescue, particularly the common ground theory. This paper reviews common ground theory, then it introduces a framework inspired by the miraculous history of Helen Keller. Then a case study is presented to investigate the proposed framework. Finally, the paper explores some potential future issues in the application of the proposed common ground approach in Human-Robot Interaction.

General Terms:

Human Robot Interaction, Common Ground

Keywords:

Human Robot Interaction, Helen Keller Heuristic, Common Ground

1. INTRODUCTION

Lacking a common ground in Human-Robot Interaction (HRI) is the main source of errors and miscommunications especially for remote exploration robotics [19]. Advances in language and social psychology studies are the rescue, particularly the common ground theory [5]. Common ground theory [4] has been applied to human robot interaction ([18],[10],[15]). Recently, many heuristics are proposed for robotics ([25],[26],[27]).
This paper reviews common ground theory, then it introduces a framework inspired by the miraculous history of Helen Keller. Then a case study is presented to investigate the proposed framework. Finally, the paper concludes with a discussion of future works.

2. COMMON GROUND THEORY

Some of the concepts and ideas studied within the framework of the common ground theory are listed below:

Speaking while monitoring addressees for understanding [5] : How people monitor their addressees, including how they watch their partners and the areas around their partners as well as how they use gestural acts .

Referring as a collaborative process [3] : Participants repair, expand on, or replace a noun phrase until they reach a version they mutually accept. Excellent table on mutual acceptance as a recursive process .

References in conversation between experts and novices [8] : As novices and experts talk with each other while completing a task, they assess each other's level of expertise and adjust their descriptions accordingly.

Language efficiency and visual technology [7] : Minimizing Collaborative Effort with Visual Information. How visual information can make collaboration more efficient. Participants complete a puzzle without a shared visual space, shared space with immediately updated visual information, shared space with delayed visual updating.

Mutual knowledge problem in dispersed collaborations [6] : A Study of six-person teams distributed across three continents identifies five types of failures of mutual knowledge: failure to communicate and retain contextual information, unevenly distributed information, difficulty communicating and understanding the salience of information, differences in speed of access to information, and difficulty interpreting the meaning of silence.

Visual Information in collaborative physical tasks [12] Bicycle repair task - situational awareness, conversational grounding. Sources of visual information.

Uncertainty, utility, and misunderstanding [14]: A decision-theoretic perspective on grounding in conversational systems. Trying to create a systematic way of providing feedback to a user. Telephone dialing system provides both negative and positive evidence to the user based on the state the system is in. Amount of feedback changes based on adjusting a grounding criterion - initially the system provides a high level of feedback; less feedback will be provided if interactions progress smoothly.

Common ground (GC) is often represented by an external representation, like a chess board or a blackboard and it falls into three main parts [4]

—what the presupposed on entering the activity,
—the current state of the activity, and
—the public events that led up to the current state.

Each of these parts divides further into the information that is officially part of the joint activity and the information that isn't.
A detailed field study is performed to investigate the relation between autonomy and common ground in HRI [18]. This field study

clearly declares that the common ground framework allows the focus on the entire dialogue between a robot and a user rather than only on the users information requirements.
There are four autonomy basic types [23]: information acquisition, information analysis, decision selection, and action implementation. In robotics, these autonomy types are commonly collapsed into three [22]:

autonomous sensing (information acquisition and data transformation): making observations and refining information,

autonomous planning (information interpretation and decision selection): reacting to information or deciding actions and schedule, and

autonomous acting (action implementation): executing a planned task or producing reflexive reactions.

These types decompose information analysis into data transformation during sensing and interpretation during planning. One robotic system can have a different autonomy level for each typethat is, sensing, planning, and acting.

3. HELEN KELLER HEURISTIC

Helen Keller [9] was born with full sight and hearing. At the age of 19 months, she lost her sight and hearing due to a serious illness. The small child terrorized the house with fits of screaming and shouting. After four years of suffering, Anne arrived at the house. Anne immediately started teaching Helen, using finger spelling on Helen's hand. It was difficult for Helen to understand besides, Anne could not control her bad behavior until they moved into the small cottage by the side of the house. A bond is developed between the two. On 5 April 1887 when pumping water onto Helen's hand, Anne spelled out the word 'water' in Helen's other hand. Somehow a trigger was set off in Helen's mind that helped her understand the concept of words and their meanings. Immediately, Helen asked Anne to reveal the name of the pump. Helen learned the name of everything she touched and also asked for Anne's name. Anne spelled the word "Teacher" on her hand. Over the next years, Helen and Anne embarked on a number of lecture tours all over the world which featured Helen talking about her life and beliefs. Anne helped interpreting Helen's words sentence by sentence.

3.1 Proposed Heuristic

For the matter of lab test, let us assume a small world environment that looks like the house of Helen. This environment is divided into rooms named A, B, C, ... Each room contains items. To move between rooms, you pass through doors. If doors are closed, one must possess the corresponding key. Helen and Anne are robots initially setting in room A. Helen robot is interested only in Food and Keys. The behavior of Helen robot is easy to be described. She keeps moving between rooms, taking any reachable objects, only retaining food and keys. For exploring version, Helen can be interested in some task-specific items such as certain minerals instead of food. Based upon these assumptions, let us propose the following algorithm.

Algorithm 1 Helen Operation

▷ Initially, currentRoom can be set randomly
1: $currentRoom \leftarrow RoomA$
▷ Initially, bag is empty
2: $bag \leftarrow \{\}$
3: **while** (TRUE) **do**
4: $objects \leftarrow senseObjects(currentRoom)$
▷ enumerate objects, one by one
5: **for** $i = 0$ to $objects.length - 1$ **do**
6: $currentObject = objects[i]$
▷ Take the current object.
7: $bag \leftarrow take(bag, currentObject)$
8: **if** ($\neg isFoodOrKey(currentObject)$) **then**
9: $bag \leftarrow drop(bag, currentObject)$
10: **end if**
11: **end for**
12: $doors \leftarrow senseDoors(currentRoom)$
13: $index \leftarrow random()$
14: $currentDoor = doors[index]$
▷ Move to another room.
15: $currentRoom \leftarrow move(currentDoor)$
16: **end while**

Anne robot keeps tracking and monitoring Helen, and following her wherever she goes. If Helen touches an object, Anne labels it for her. Anne also keeps listening to any human voice for interacting with Helen. The algorithm describing the control loop of Anne is simple. In the 7[th] line of Helen algorithm, whenever Helen takes the current object, this triggers Anne's attention to spell its name for her. In 15[th] line, whenever Helen moves to another room, this triggers Anne's attention to follow her.

3.2 Implementation

An implementation is coded in Microsoft Visual C#.NET . Earlier versions are coded in Java.
The user interface is initially empty. First, click the generate menu item shown in figure 1 to generate the world. This will load the configuration presented in SmallWorld class.

```
    public class SmallWorld
    {
    public static List<string> foods
        = new List<string>(
            new string[]{
                "an apple",
                "a pear",
               "some tea",
                "some jam",
                "a green pear"
            });
    public List<Room> rooms = new List<Room>(
            new Room[]
    {
    room ("Room A",
"This room contains a box which contains an apple.\n"
    + "There are two doors in this room",
 item_list( item ("a box", contains ("an apple"))),
 door_list( door ("to room B", "Room B"),
door ("to room C", "Room C"))
            ),
      room ("Room B",
```

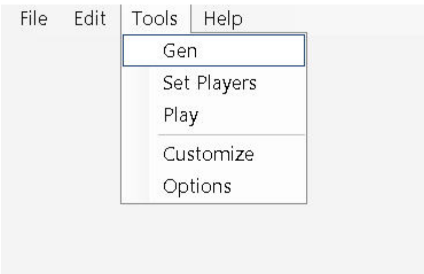

Fig. 1. Generate The World

```
  "This room contains several objects.",
  item_list( item ("a pear"),
item ("a knife"),
      item ("a fork"),
    item ("a blue key")),
 door_list( door ("via a blue door",
                     "Room A","a blue key"),
                     door ("to room C", "Room C"))
           ),
      room ("Room C",
  "This room contains nested objects.",
```

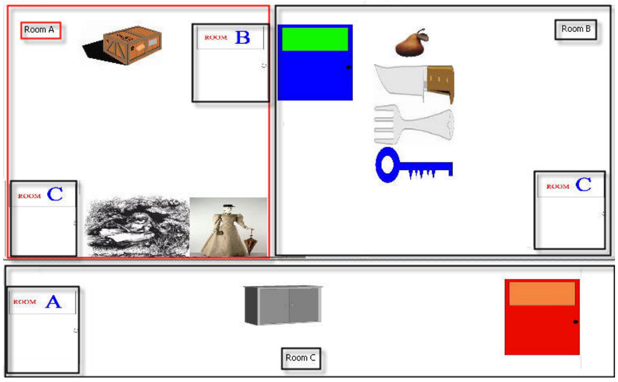

Fig. 2. Small World

```
  item_list( item ("a cupboard",
        contains (  item ("a red key"),
                          item ("a teapot",
                contains ("some tea")),
            item ("a jam jar",
            contains ("some jam"))))),
 door_list( door ("to room A", "Room A"),
                door ("a red door",
                "Room B",
                "a red key"))
            )
        });
}
```

Figure 2 shows the generated environment after clicking the generate and set players menu items.
Figure 3 shows the structure of the proposed framework of communication. Helen has simple sensors to probe the environment to get the next item. Anne keeps an EYE on Helen and an EAR on the Human. If Anne notices Helen grasping an item, she tells her its

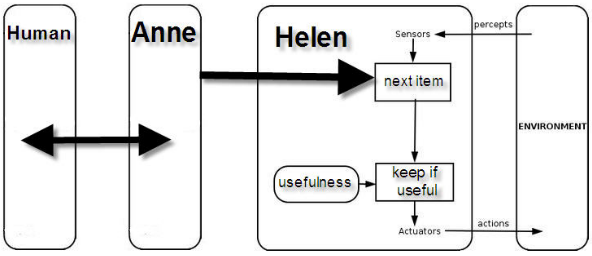

Fig. 3. Proposed Framwork

name and any supporting information. If Anne hears a query from the Human, she reply about the situation.

4. EVALUATION

By clicking the play menu item, Helen and Anne start moving and interacting with the environment and with each other.
Here is a listing of there interaction.

```
-------------------------
Helen is in Room A
-------------------------

Helen grasps a box (ContainerItem)
a box contains :
an apple

Anne checks a box (ContainerItem)
Anne checks a box (ContainerItem)
Helen grasps an apple (Item)
Anne checks an apple (Item)
Anne checks an apple (Item)
Helen keeps an apple
Helen grasps a key to Room B (Door)
Helen delves into Room B

-------------------------
Helen is in Room B
-------------------------

Helen grasps a pear (Item)
Anne checks a pear (Item)
Anne checks a pear (Item)
Helen keeps a pear
Helen grasps a knife (Item)
Anne checks a knife (Item)
Anne checks a knife (Item)
Helen grasps a fork (Item)
Anne checks a fork (Item)
Anne checks a fork (Item)
Helen grasps a blue key (Item)
Anne checks a blue key (Item)
Anne checks a blue key (Item)
Helen keeps a blue key
Helen grasps a key to Room A via blue key (Door)
Helen grasps a key to Room C (Door)
Helen delves to Room C

-------------------------
Helen is in Room C
-------------------------

Helen grasps a cupboard (ContainerItem)
a cupboard contains :
a red key
a teapot
a jam jar

Anne checks a cupboard (ContainerItem)
Anne checks a cupboard (ContainerItem)
Helen grasps a red key (Item)
Anne checks a red key (Item)
Anne checks a red key (Item)
Helen keeps a red key
Helen grasps a teapot (ContainerItem)
a teapot contains :
some tea

Anne checks a teapot (ContainerItem)
Anne checks a teapot (ContainerItem)
Helen grasps some tea (Item)
Anne checks some tea (Item)
Anne checks some tea (Item)
Helen keeps some tea
Helen grasps a jam jar (ContainerItem)
a jam jar contains :
some jam

Anne checks a jam jar (ContainerItem)
Anne checks a jam jar (ContainerItem)
Helen grasps some jam (Item)
Anne checks some jam (Item)
Anne checks some jam (Item)
```

```
Helen keeps some jam
Helen grasps a key to Room A (Door)
Helen delves to Room A
```

5. CONCLUSION

Helen robot designed to be blind and deaf, makes it inexpensive. This allows the construction on many Helen robots with no fear of losing such probing robots in the exploring environment. Keeping an EYE on Helen, Anne robot has a simple control loop, just to track, follow and teach Helen about the item in hand. Anne also plays as a middleware layer between Helen, the exploring robot, and the Human.

6. FUTURE WORK

Future work have many dimension. One dimension is to consider the giving Anne robot the ability to fix and reconfigure the setting on Helen robot. Another dimension is to study the impact of transformable items due to interactions, or something like the following code.

```
        public List<Interaction> interactions
        = new List<Interaction>(new Interaction[] {
    combine ("an apple", "some tea",
            "some apple flavoured tea"),
    combine ("a box", "a pear",
     item ("a new box", contains ("a green pear")))
        });
```

Another dimension is to enhance the accuracy of the framework by incorporating transfer learning [24] , to be able to generalize what learnt from the small world environment to the world. In many machine learning, there is an assumption that the training and future data must be in the same feature space and have the same distribution. However, in many real-world applications, this assumption may not hold. Knowledge transfer, if done successfully, would greatly improve the performance of learning by avoiding much expensive data-labeling efforts.

7. REFERENCES

[1] Leslie Lamport, *A Document Preparation System*. Addison Wesley, Massachusetts, 2nd Edition, 1994.

[2] aSusan E. Brennan and Eric A. Hulteen, *Interaction and feedback in a spoken language system: A theoretical framework*. Knowledge-Based Systems, 1995, volume 8, pages 143-151, issue 2.

[3] H. Clark and D. Wilkes-Gibbs, *Referring as a collaborative process*. Cognition, 1986, volume 22, pages 1-39, number 1.

[4] Herbert H. Clark, *Using Language*, Cambridge University Press, 1996.

[5] Herbert H. Clark and Meredyth A. Krych, *Speaking while monitoring addressees for understanding*, Journal of Memory and Language, 2004, volume 50, pages 62-81.

[6] Catherine Durnell Cramton, *The Mutual Knowledge Problem and Its Consequences for Dispersed Collaboration*, Organization Science, 2001, volume 12, pages 346-371, number 3.

[7] Darren Gergle and Robert E. Kraut and Susan R. Fussell, *Language Efficiency and Visual Technology: Minimizing Collaborative Effort with Visual Information*, Journal of Language and Social Psychology, 2004, volume 23, pages 491-517, number 4.

[8] Ellen A. Issacs and Herbert H. Clark, *References in Conversation Between Experts and Novices*, Journal of Experimental Psychology: General, 1987, volume 116, pages 26-37, number 1.

[9] Helen Keller, *The Story of My Life*, Dover Thrift, 1996.

[10] Sara Kiesler, *Fostering Common Ground in Human-Robot Interaction*, Proceedings of the IEEE International Workshop on Robots and Human Interactive Communication (RO-MAN), 2005, pages 729-734.

[11] R. M. Krauss and Susan R. Fussell, *Social psychological models of interpersonal communication*, Social psychology: Handbook of basic principles, Guilford Press, 1996, editors E. T. Higgins and A. Kruglanski, pages 655-701.

[12] Robert E. Kraut and Susan R. Fussell and Jane Siegel, *Visual Information as a Conversational Resource in Collaborative Physical Tasks*, Human-Computer Interaction, 2003, volume 18, pages 13-49.

[13] Shuyin Li and Britta Wrede and Gerhard Sagerer, *A computational model of multi-modal grounding*, Proceedings of the ACL SIGdial Workshop on Discourse and Dialog, in conjunction with COLING/ACL 2006, year 2006, pages 153-160, ACL Press.

[14] Tim Paek and Eric Horvitz, *Uncertainty, utility, and misunderstanding: A decision-theoretic perspective on grounding in conversational systems*, Psychological models of communication in collaborative systems, Papers from the AAAI Fall Symposium, November 5-7, North Falmouth, Massachusett, 1999, pages 85-92.

[15] Aaron Powers and Adam D. I. Kramer and Shirlene Lim and Jean Kuo and S. Lee and Sara Kiesler, *Eliciting Information from People with a Gendered Humanoid Robot*, Proceedings of the IEEE International Workshop on Robots and Human Interactive Communication (RO-MAN), 2005, pages 158-163.

[16] Abigail J. Sellen, *Speech Patterns in Video-Mediated Conversations*, Proceedings of the 1992 SIGCHI Conference on Human Factors in Computing Systems (CHI 1992), 1992, pages 49-59, ACM.

[17] Kerstin Severinson-Eklundh and Helge Huttenrauch and Anders Green, *Social and collaborative aspects of interaction with a service robot.*, Robotics and Autonomous Systems, Special Issue on Socially Interactive Robots, 2003, volume 42, number 3.

[18] Kristen Stubbs and Pamela Hinds and David Wettergreen, *Autonomy and Common Ground in Human-Robot Interaction: A Field Study*, IEEE Intelligent Systems, Special Issue on Interacting with Autonomy, 2007, volume 22, pages 42-50, number 2.

[19] Kristen Stubbs and Pamela Hinds and David Wettergreen, *Challenges to Grounding in Human-Robot Interaction: Sources of Errors and Miscommunications in Remote Exploration Robotics*, Proceedings of the First International Conference on Human-Robot Interaction, ACM, 2006.

[20] Kristen Stubbs and Pamela Hinds and David Wettergreen, *Challenges to Grounding in Human-Robot Collaboration: Errors and Miscommunications in Remote Exploration Robotics*, Carnegie Mellon University , 2006.

[21] Cristen Torrey and Aaron Powers and Matthew Marge and Susan R. Fussell and Sara Kiesler, *Effects of Adaptive Robot Dialogue on Information Exchange and Social Relations*, Proceedings of the First Annual Conference on Human-Robot Interaction, pages 126-133, ACM, 2006.

[22] Veloso, Manuela M., *Entertainment robotics*, Commun. ACM, volume 45, number 3, 2002, pages 59-63.

[23] Parasuraman, R. and Sheridan, T. B. and Wickens, C. D., *A model for types and levels of human interaction with automation*, Trans. Sys. Man Cyber. Part A, volume 30, number 3, 2000, pages 286-297.

[24] Sinno Jialin Pan and Qiang Yang, *A Survey on Transfer Learning*, Knowledge and Data Engineering, IEEE Transactions on, 2010, volume 22, number 10, pages 1345-1359.

[25] M. A. El-Dosuky, Ahmed El-Bassiouny, Taher Hamza, Magdy Rashad, Simulated Tom Thumb, the Rule Of Thumb for Autonomous Robots, CoRR abs/1210.2421 (2012)

[26] M. A. El-Dosuky, M. Z. Rashad, T. T. Hamza, A. H. El-Bassiouny, Collaborating Robotics Using Nature-Inspired Meta-Heuristics, CoRR abs/1212.5777 (2012)

[27] M. A. El-Dosuky, M. Z. Rashad, T. T. Hamza, A. H. El-Bassiouny, Robopinion Opinion Mining Framework Inspired by Autonomous Robot Navigation, CoRR abs/1209.0249 (2012)