Empathy in Humanoid Robots

Marina Sardà Gou[1], Vasiliki Vouloutsi[1], Klaudia Grechuta[1], Stéphane Lallée[1], and Paul F.M.J. Verschure[1,2]

[1] Laboratory of Synthetic, Perceptive, Emotive and Cognitive Systems – SPECS, Universitat Pompeu Fabra, Roc Boronat 138, 08018, Barcelona, Spain
[2] ICREA Institució Catalana de Recercai Estudis Avançats, Passeig Lluís Companys 23, E-08010 Barcelona, Spain

Abstract. Humanoid robots should be able to interact with humans in a familiar way since they are going to play a significant role in the future. Thus, it is necessary that Human-Robot Interaction (HRI) is designed in such a way that allows humans to communicate with robots effortlessly and naturally. Emotions play an important role in this interaction since humans feel more predisposed to interact with robots if they are able to create an affective bond with them. In this study, we want to know whether humans are able to empathize with a humanoid robot. Therefore, in the present research, we are going to recreate a Milgram experiment in which we expect participants to empathize with the robot while playing a matching game. Like in Milgram's experiment, they will have to give fake electrical shocks to the robot thinking that they are punishing it. In that way, an empathic state, which we expect to see in our results, may be induced.

Keywords: Milgram experiment, empathy, humanoid robot, Human-Robot Interaction (HRI).

In the future, humanoid robots will undoubtedly play an important role in our society since they can perform a variety of practical activities in peoples' everyday life. The robots which are meant to interact with people should be therefore designed to provide a smooth and comfortable Human-Robot Interaction (HRI). Studying and incorporating emotions is a crucial element of HRI research which aims at achieving that goal since robots, which are able to display emotions through facial expression and/or gestures are perceived as more familiar and warm. In that way, humans feel more likely to interact with them. (Fong, 2003) This may also result in the increased level of empathy towards the robot, and consequently in an affective bond from the human to the robot. The purpose of this study is to investigate whether a person can feel empathy towards a humanoid robot, and if so, to which extent.

In order to answer our question, first, it is necessary to define and understand what "empathy" means and how we can measure it. According to Hoffman, empathy is "an affective response more appropriate to another's situation than to one's own." (Hoffman 2001). We can feel empathy in two different ways; one being cognitive and another being affective. Cognitive empathy is the understanding of another's emotions,

A. Duff et al. (eds.): Living Machines 2014, LNAI 8608, pp. 423–426, 2014.
© Springer International Publishing Switzerland 2014

while affective empathy is the possession of that emotion. (D'Ambrosio, 2009) It is "an important contributor to successful social interaction, allowing us to predict and understand others' behaviour and react accordingly." (Engen et al., 2012) However, it is not something that happens every time we perceive emotions from another person. Empathy depends on several factors such as the social context, culture or the characteristics of both the empathizer and the object of empathy (Engen et al., 2012). Therefore, in order to feel empathy, one must feel, in some way, identified with the person having an emotional response.

Humans do not only feel empathy for human beings. Indeed, we can also empathize with animals as well as physical objects once we perceive that they are expressing emotions. For instance, in fiction, there are quite a few films about inanimate objects which show feelings through facial expressions or gestures and, consequently, people empathize with them. Since we are aware that objects have no emotions at all, it is our imagination that plays an important role in the process. Humans can empathize with an inanimate object when, through their imagination, they can perceive that a given object has/displays an emotional response, and this perception causes humans to feel empathy for this particular object. (Misselhorn, 2009) That is to say, with our imagination, we can attribute affective behaviour to some inanimate objects and, as a consequence, we may feel empathy towards them.

In order to test whether humans can feel empathy for a humanoid robot, we will run an adaptation of Milgram experiment (Milgram, 1963). Milgram did his experiments in the 1960's in which he proved that ordinary people are able to cause a huge amount of pain to another person if there is an authority figure that forces them to do so. However, this does not mean that they do not feel empathy for the victim. In fact, as Milgram claimed, the participants were under extreme stress during the experimental procedure. "An unanticipated effect was the extraordinary tension generated by the procedures. […] In a large number of cases the degree of tension reached extremes that are rarely seen in socio-psychological laboratory studies." (Milgram 1963) Therefore, the participants did suffer because of the victim's pain. Two of the variations Milgram did of his own experiment were bringing the victim closer and removing physically the authority figure from the laboratory. There, he could see that the obedience of the participants dropped drastically. (Milgram, 1974) This also supports the idea of participants being empathic towards the victim. In that way, Milgram's experiment can be used in order induce this emotional state to participants.

In fact, there are already some studies in which researchers performed Milgram's experiment with robots and virtual humans. Slater et al. found that, using an avatar as the victim, participants tend to respond, physiologically and behaviourally, in a similar way that they do with a human victim. Some of them hesitated to continue the experiment or, sometimes, did not want to proceed. In a different way, using less anthropomorphic robots than a virtual human, participants felt compassion for the robot, but the urges of the authority figure were enough to make them proceed. (Bartneck et al., 2005)

In the present study we will apply a similar method. The robot we are going to use is the humanoid robot, iCub. Our hypothesis is that when a robot uses eye contact and shows emotions, humans feel more compassion towards it. As a consequence, they will administer less amount of shock to the robot. Like in Milgram experiment, participants won't know the actual purpose of the experiment. They will be able to interact with the robot through a “matching game” where the goal is to teach the robot the colours in Spanish. The participant and the iCub will be playing the game at the Reactable.

In the game, first, the participant says one colour to the robot and, after that, the robot touches the correspondent colour in the Reactable. There will be 40 trials and, like in Milgram's experiment, three wrong answers to one correct answer. Every time the matching is incorrect, the human needs to punish the robot by administering simulated electrical shocks. The shocks are not real but we want the participant to think that they are real because, in that way, the subject might feel that he is causing pain to the robot and may empathize with it. The shock generator is a crucial issue that needs to be explained since we are not going to use the one that Milgram used. Instead, in the present scenario there is a regulator button on which the user can choose what amount of pain s/he wants to give to the robot. However, we needed to constrain the participant’s choice to prevent from choosing the same shock amount. We will therefore force him to increase the shocks as the experiment proceeds. Since we allow the participant to stop the experiment at any time we need an authority figure. It will not be a human but a set of pre-recorded sentences that will instruct the participant. If the participant does not want to proceed anyway, the experiment will end. Since we want to know how the robot’s behaviour can affect participants' empathy, we will have four conditions in which we will change the iCub responses. Two independent variables will be the iCub’s eye contact and its emotional responses expressed through speech and facial expressions. In the first condition, which is the control group, the robot will not have any of them; in the second condition, it will use the eye contact; in the third condition, it will show emotional responses, and in the fourth condition it will have the both of them.

In every condition, we will have several dependent variables. We will measure participants' heart rate and since it has been proved that, when humans feel empathy, the heart rate increases (Miu 2012, Silva 2011). We will also record participants’ performance and a naïve judge will further analyze the eye contact, facial expressions and the speech of all the subjects. We will store the Reaction Times (RTs) between the robot making a mistake and the participant giving the shock, and measure the amount of shock. If the participant hesitates to continue, the system will measure the number of audio pre-recorded sentences s/he needs to hear in order to proceed. Finally, every participant will answer a questionnaire about empathy in which, among other issues, they are going to be asked if they think that the robot can feel pain. If that is the case, this would help us to understand any empathic response participants may have.

When the robot shows emotional responses and eye contact, and as the amount of shock increases, we expect to see the following participants' reactions: an increasing heart rate, less eye contact with the iCub (Milgram showed that participants refused to look at the victim when giving the shocks), more reaction time, less amount of shock administered, the participants' necessity to hear more audio instructions to continue, and finally, the answers in the questionnaire, which show that the participant feels empathy for the robot.

The goal of the present study is to investigate whether humans are able to empathize with a humanoid robot and, if so, what are the behaviours, which cause empathy in humans. It is also necessary to say that the experimental procedure described above should be tested in a pilot experiment that we are going to perform in the time coming.

References

Bartneck, C., et al.: Robot abuse—a limitation of the media equation. In: Proceedings of the Interact 2005 Workshop on Agent Abuse (2005)

D'Ambrosio, F., Olivier, M., Didon, D., Besche, C.: The basic empathy scale: a French validation of a measure of empathy in youth. Personality and Individual Differences 46, 160–165 (2009)

Engen, H.G., Singer, T.: "Empathy circuits.". Current Opinion in Neurobiology 23, 275–282 (2012)

Fong, T., Nourbakhsh, I., Dautenhahn, K.: "A survey of socially interactive robots". Robotics and Autonomous Systems 42, 143–166 (2003)

Milgram, S.: Behavioral study of obedience. Journal of Abnormal and Social Psychology 67, 371–378 (1963)

Milgram, S.: Obedience to Authority. Harper & Row, New York (1974)

Catrin, M.: Empathy with Inanimate Objects and the Uncanny Valley. Minds and Machines 19, 345–359 (2009)

Miu, A.C., Balteş, F.R.: Empathy Manipulation Impacts Music-Induced Emotions: A Psychophysiological Study on Opera. PloS ONE 7(1) (2012)

Silva, P.O., Gonçalves, Ó.F.: Responding empathically: a question of heart, not a question of skin. Appl Psychophysiol Biofeedback 36(3), 201–207 (2011)

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., et al.: A virtual reprise of the Stanley Milgram obedience experiments. PLoS ONE 1(1) (2006)