The study of human communicative behavior while solving a tangram puzzle for its subsequent transfer to a robot

Anna Zinina[1] [0000-0001-9575-1875], Liudmila Zaidelman[1] [0000-0002-2941-144X], Nikita Arinkin[1] [0000-0003-2303-2817], Artemiy Kotov[1] [0000-0003-3353-5549]

[1] National Research Center "Kurchatov Institute", Moscow, Russia
Zinina_aa@nrcki.ru

Abstract. In a more general project, the present study is a part of, we apply natural communicative behavior to a robot in a situation where it acts as an assistant for a child while the child is solving a spatial problem – in our case, a Tangram puzzle. In order to find the key features of such tutoring behavior, we have arranged a natural situation with two people, helping each other to solve a puzzle. To date, 10 adult (5 pairs) men and women have been recorded. We analyzed the communication strategies that respondents use to effectively help in solving the task. We pay special attention to the emotional dynamics of the participants in each of the two dialogue positions: the assistant and the Tangram solver. We use the obtained data to develop deep and differentiated emotional model for the robot that is applicable to the situation of free assembly of the Tangram puzzle.

Keywords: Multimodal communication, robot-to-human interaction, affective robot tutors.

1 Introduction

What makes a robot attractive? In natural communication, a person adjusts his behavior to the situation and actions of other people. Therefore, for the effective interaction with humans, robots also need to perform coordinated and timely actions based on the analysis of their social environment. A key feature of social robot behavior is the ability to adapt to the changing needs of the user [1]. Perceived adaptability affects the perceived utility of the robot, increases user's satisfaction from the interaction as well as the intention to use it in the future [2]. Robot's ability to respond to changes in the surrounding situation, to adapt its behavior and emotional expression to the users is an important factor to create a positive impression of interaction with a companion robot. It is important for robots to demonstrate emotional dynamics and expression depending on incoming events, such as reactions to successful or incorrect user's actions, user's questions, and user's gestures directed to the robot (for example, touching). The emotional model of companion robots might provide flexible and diverse behavior that underlies social interaction with humans.

The robot's ability to simulate 'feelings' and express the variety of emotional reactions using expressive means (e. g. movements with eyes, head, and hands) is highly appreciated in learning [3-6]. Robots are used for teaching natural sciences [7], mathematics [8], music [9], and foreign language [10]. The use of robots might be also effective for developing children's cognitive skills [11-12]. Social robots involve children in learning, increase their motivation and curiosity, as well as the number of emotional responses [13-15]. Robots can demonstrate various means of communication: using gestures, boy postures, and facial expressions, that also helps to increase the interest and motivation of users to learning. Thus, modeling complex emotional behavior is one of the key characteristics for educational robots.

So many researchers pay special attention to the development of an emotional robot model [16] in learning. For example, in [17] iCat robot plays chess with a child. The robot's emotional state and expression is affected by every move of the child. Children may interpret the robot's affective behavior and by that acquire additional information to better understand the game. The robot has empathic abilities, that also contributes to improving children's chess skills.

In our lab, we are developing a robot that acts as a child's assistant in solving puzzles. In this way, the robot controls the solution of the task: it introduces the puzzle to a child, gives instructions, and monitors the progress of the task.

2 F-2 robot platform for experiments

2.1 Modeling of multimodal robot behavior

We are developing the F-2 robot, which can be used as an experimental platform for the development of interaction models between humans and robots (Fig. 1). Robot's movements are modeled based on The Russian Emotional Corpus (REC) [18]. In this way, we model a complex robot behavior that is as close as possible to natural communication behavior. This behavior allows the robot to interact with people naturally and intuitively.

Numerous experimental studies have shown that complex nonverbal behavior of a robot has an effect on the attractiveness of the robot to the user. For example, in [19], we investigated the effect of complex robot eye movements on users. The experiment [20] evaluated the contribution of various means of communication (eye movements, facial expressions, gestures, speech) to the positive impression of the robot. It was shown that emotional gestures of the F-2 robot increase its attractiveness to the user, more than head movements and facial expression. Another experiment [21] investigated the effect of oriented robot gestures on users in spatial game situations. It was found that subjects implicitly prefer the robot that uses pointing gestures in its instructions. It is also shown that some participants in the experiment follow the robot's pointing gestures, without realizing it.

Fig. 1. The robot F-2.

In a recent study [22], we test the effect of robot's emotional gestures and speech on participants in a game situation. In the experiment two identical robots helped children to complete the tangram puzzle. In the experiment two independent variables were varied, each of which had two levels: (a) robots demonstrated expressive (emotional) or neutral gestures, (b) robots could react with emotional or neutral statements. It was found that emotional gestures are the key factor that influences the attractiveness of the robot for the child. In addition, children noted that the robot with emotional speech and gestures is more kind, empathic, it "has interesting words". In our experiment it was found that the robot F-2 successfully acts as a teacher, children like his assistance in solving spatial puzzles.

Previous simulations of game assistance were organized in the Wizard of Oz paradigm: the moves by the player were evaluated as successful or not by a remote human operator. Robot has been suggesting to complete the puzzle in a fixed order. As we develop an automatic system computer vision recognition system for Tangram puzzle, our attention is focused on the development of an extended emotional model of the robot with more complex system of robot responses with the optimal frequency of suggestions. The emotionality of gestures and speech should be more differentiated. On the one hand, when developing a model, it is necessary to focus on well-known classifications of emotions. For example, the robot must be able to look surprised, sad, happy, angry, frightened or fell shame. On the other hand, a qualitative analysis of the behavior of real people in identical situations is necessary. For example, in [23] the analyses of human-human interaction (HHI) has been suggested as the basis of multidisciplinary approach to the development of empathic robotic tutor. In [24] the Inter-ACT (INTEracting with Robots–Affect Context Task) corpus was presented, an affective and contextually rich multimodal video corpus containing affective expressions of children playing chess with the iCat robot.

2.2 Video corpus

To develop a deeper and more differentiated emotional model, we need a qualitative analysis of behavior of real people in identical situations – in situations of assembling a puzzle. In our work, we create a corpus with recordings of pairs of people helping each other to solve a tangram puzzle. To date, 10 adults (5 pairs, average age 34.5 years) have been recorded. We analyze the communication strategies that respondents use to effectively assist in solving the problem and pay special attention to the emotional dynamics of the research participants in each of the dialogue positions (assistant and puzzle solver).

The participants helped each other solve a tangram puzzle by suggesting the right moves. Each subject had to solve and to explain 4 figures: two from a single set of Tangrams (7 elements), two from a double set (complex figures, made of 14 elements). In total, solutions to 40 Tangram tasks were recorded. Before the beginning of each session, the solver received the outer contour of the figure to the solved, and the assistant received the detailed composition of the figure – one of the possible solutions. The assistants were not limited to follow a specific strategy – it was important to create the situation of free assembly to select the appropriate strategies for the robot. The respondents could use pointing gestures, but they were not allowed to touch the figures and the target place for the game element. The experiment was recorded on a video camera (Fig. 2).

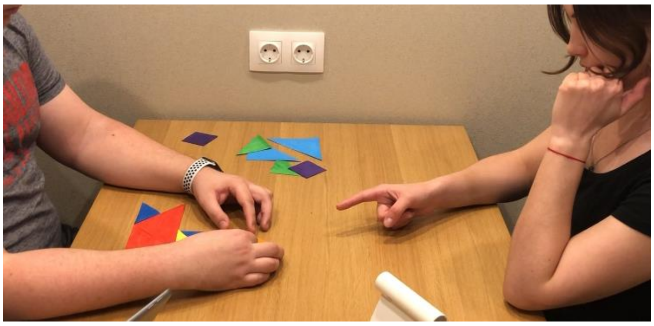

Fig. 2. The experimental situation.

Following the analysis of video recordings, we have identified the main assisting strategies. The strategy in which the assistant focuses on the actions of the solver was called "**Helping**". It is characterized by the increased time for requesting a hint, adjustment of hints to the current arrangement of elements on the field, a general orientation towards the behavior of the puzzle solver. The strategy, in which assistants insistently try to impose their strategy on the solver, was called "**Dominant**". Study participants used both strategies depending on the complexity of the task and the previous interaction experience. The average waiting time for a hint or comment for the *Helping* strategy is about 10 seconds, for the *Dominant* – 3 seconds. Consequently, while assisting a human, a robot must wait for a hesitation pause of variable duration to give an advice, so as not to look intrusive and not deprive the subjects of the opportunity to solve complex problems without assistance.

The analysis of video recordings revealed the types of instruction used by the study participants:

1. Instruction about specific operations with a single game element.
2. Instruction about the decision procedure.
3. Instruction about the general structure of the figure or a general instruction.

These types of instructions can be initiated by the assistant before the puzzle solver begins the corresponding actions – this is typical for the *dominant* strategy. Instructions can also be offered when the puzzle solver meets the difficulties, in response to the solver's request, which is usually typical of a *helping* strategy. When assembling a tangram puzzle, respondents balance between different types of prompts, develop the most optimal strategy for interacting with each other.

After analyzing the corpus, it was also found that participants in the experiment demonstrate complex behavioral patterns for emotions of different "depths". The emotional expressions can be divided into *push emotions* (internal or experienced) and *pull emotions* (external or expressed).

According to the data obtained, half of the subjects experienced great difficulties not in assembling itself, but rather in assisting. The respondents began to get nervous, when the solver misinterpreted the advices, tried to hide their irritation, got upset because of the lack of mutual understanding in the pair, etc. Informants often intentionally used exaggerated expressions of fatigue or surprise to indicate the opponent's wrong action – e. g. behavioral patterns corresponding to the statements: *I can't stand your mistakes anymore!* or *Why is it so hard?* Such emotions were demonstrated to make the assessment of the interlocutor's actions more explicit. In other words, if a person wants to describe the interlocutor's action as incorrect, he can imitate emotion and broadcast a message about the interlocutor's incorrect actions not only through a direct statement (*Wrong*), but also through an emotional pattern.

The identified patterns can be interpreted based on K. Scherer's concept of *push* and *pull* emotions (experienced and expressed emotions) [25]. Based on experimental and corpus studies, Scherer showed that *push-emotions* are experienced internally by a person, while their external expression is suppressed as much as possible. This is typical for those emotions that are not approved in society (aggression, disgust, gloating). At the same time, *pull-emotions* can be experienced by a person relatively poorly, but their external expression is significantly exaggerated, for example, this is typical for empathy and guilt – for emotions in which the expression is associated with social approval.

Implementing such emotional dynamics on a robot will allow us to design complex emotional responses. The obtained data is considered when developing the emotional model.

3 Formal model of emotional dynamics

To develop a formal emotional model, we proceed from the following requirements:

1. The robot should process the majority of incoming stimuli, giving preference to the correct and incorrect movements of the user.
2. The robot can take the initiative and give advice to the user.

3. The choice of communication strategy (as in 1 and 2) depends on (a) the significance of the stimulus, (b) the robot's simulated personal characteristics – emotional profile, and (c) the overall simulated emotional state of the robot.

Система управления роботом позволяет нам комбинировать мимические и жестовые реакции робота из различных поведенческих пакетов (в формате BML). Например, робот может в один момент выражать отрицание или недовольство в движениях головы, и при этом компенсировать их движениями рук (например, почесыванием или автоманипулированием). Архитектура позволяет моделировать выражение complex behavioral patterns for emotions of different "depths".

Мы разрабатываем эмоциональную модель, которая в рамках игровой ситуации имитирует как краткосрочную эмоциональную динамику (в рамках одной коммуникативной реакции) с помощью инвентаря выражаемых эмоций (pull-эмоций), так и более долгосрочную (на интервалах в несколько действий) коммуникативную динамику – в функции переживаемых push-эмоций.

Для комбинаций эмоций используется следующий шаблон реагирования:

вход от пользователя: <действие пользователя>

выход от робота: <паттерны коммуникативных функций выражаемых эмоций> <высказывание робота> <паттерны коммуникативных функций переживаемых эмоций>

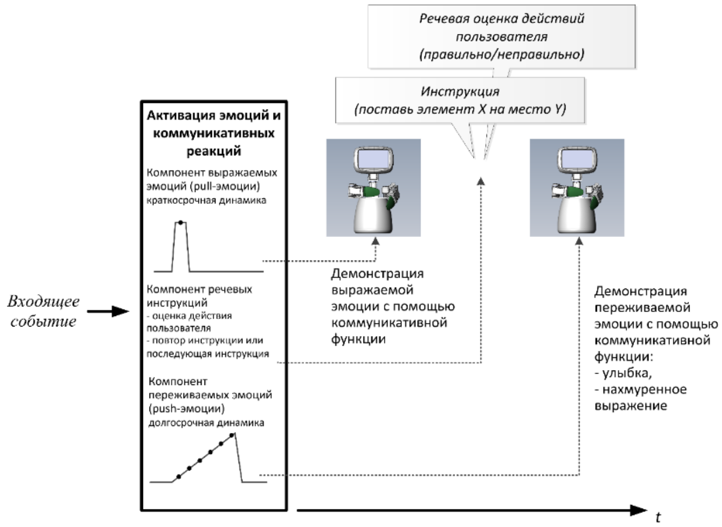

Fig. 3. ___ Схема эмоциональной модели

При поступлении на вход правильного или неправильного действия пользователя вычисляется близость события со стимульными шаблонами выражаемых и переживаемых эмоций. Выполняется расчет и обновление уровня активизации эмоций. После чего, в соответствии с приведенным шаблоном, формируется исходящее сообщение, в начале которого выполняются паттерны выражаемых эмоций. Это, прежде всего, эмоциональные знаки, ориентированные на обратную связь – пользователь должен понять, оценивает ли робот его последнее действие позитивно или негативно. Входящее событие также меняет активацию общих переживаемых эмоций, которые демонстрируются с помощью достаточно слабых внешних проявлений.

4 Conclusions

Our goal was to develop the communicative behavior of the robot for a situation in which it acts as a tutor in solving a tangram puzzle. For this, the corpus was collected and analyzed, which includes video recordings of the puzzle assembly during the interaction of two adults. Analysis of video recordings allowed us to identify key features that need to be implemented in robot's communicative behavior: types of assistance advices, the strategies of their requests and offers. In addition, we found that participants in the experiment demonstrated complex behavioral patterns for emotions of different "depths". Their emotions can be divided into internal (experienced or *push-emotions*) and external (expressed or *pull-emotions*). Based on the data obtained, we develop a flexible emotional robot model that adapts to a variety of communication situations.

Acknowledgments. The present study has been supported by the Russian Science Foundation, project No № 19-78-00113.

References

1. Heerink M., Krose B., Evers V., Wielinga B.: Assessing the acceptance of assistive social agent technology by older users: the Almere model. International Journal of Social Robotics. **2**(4), 361–375 (2010).
2. Broadbent E., Stafford R., MacDonald B.: Acceptance of healthcare robots for the older population: review and future directions. International Journal of Social Robotics. **1**, 319-330 (2009).
3. Kanda, T., Hirano, T., Eaton, D., Ishiguro, H.: Interactive robots as social partners and peer tutors for children: A field trial. Human-Computer Interaction. **19**(1-2), 61-84 (2004).
4. Alemi, M., Meghdari, A., Ghazisaedy, M.: Employing humanoid robots for teaching English language in Iranian junior high-schools. International Journal of Humanoid Robotics. **11**(03), 1450022 (2014).
5. Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., Tanaka, F.: Social robots for education: A review. Science robotics. **3**(21), eaat5954 (2018).
6. Dautenhahn, K.: Socially intelligent robots: dimensions of human–robot interaction. Philosophical transactions of the royal society B: Biological sciences. **362**(1480), 679-704 (2007).
7. Shiomi, M., Kanda, T., Howley, I., Hayashi, K., Hagita, N.: Can a social robot stimulate science curiosity in classrooms?. International Journal of Social Robotics. **7**(5), 641-652 (2015).

8. Brown, L. N., Howard, A. M.: The positive effects of verbal encouragement in mathematics education using a social robot. In: IEEE integrated STEM education conference, pp. 1-5 (2009).
9. Han, J. H., Kim, D. H., Kim, J. W.: Physical learning activities with a teaching assistant robot in elementary school music class. In: 2009 Fifth International Joint Conference on INC, IMS and IDC, pp. 1406-1410 (2009).
10. van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., van der Ven, S., Leseman, P.: Social robots for language learning: A review. Review of Educational Research. **89**(2), 259-295 (2019).
11. Park, H. W., Rosenberg-Kima, R., Rosenberg, M., Gordon, G., Breazeal, C.: Growing growth mindset with a social robot peer.: In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, pp. 137-145 (2017).
12. Cook, A., Encarnação, P., Adams, K.: Robots: Assistive technologies for play, learning and cognitive development. Technology and Disability. **22**(3), 127-145 (2010).
13. Kanda, T., Hirano, T., Eaton, D., Ishiguro, H.: Interactive robots as social partners and peer tutors for children: A field trial. Human-Computer Interaction. **19**(1-2), 61-84 (2004).
14. Gordon, G., Breazeal, C., Engel, S.: Can children catch curiosity from a social robot?: In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, pp. 91-98 (2015).
15. Wainer, J., Feil-Seifer, D. J., Shell, D. A., Mataric, M. J.: The role of physical embodiment in human-robot interaction. In: ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication, pp. 117-122 (2006).
16. Cuadrado, L. E. I., Riesco, Á. M., de la Paz López, F.: FER in Primary School Children for Affective Robot Tutors. In: International Work-Conference on the Interplay Between Natural and Artificial Computation, pp. 461-471. Springer, Cham (2019).
17. Leite, I., Pereira, A. Icat, the affective chess player. In: Proceedings of the systems demonstrations. Second international conference on Affective Computing and Intelligent Interaction (ACII 2007), pp. 29-33 (2007).
18. Kotov, A., Zinina, A.: Functional analysis of non-verbal communicative behavior. In: Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2015". **14**(21), 308-322. (2015).
19. Zinina, A., Arinkin, N., Zaydelman, L., Kotov, A.: Development of communicative behavior model for f-2 robot basing on «REC» multimodal corpora [Razrabotka modeli kommunikativnogo povedeniya robota f-2 na osnove mul'timodal'nogo korpusa «REC»]. In: Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2018". **17**(24), 831-844 (2018).
20. Zinina, A., Zaydelman, L., Arinkin, N., Kotov, A.: Non-verbal behavior of the robot companion: a contribution to the likeability. In: Procedia Computer Science. Postproceedings of the 10th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2019 (Tenth Annual Meeting of the BICA Society). pp 800-806. Elsevier (2019). DOI: 10.1016/j.procs.2020.02.160
21. Zinina, A., Arinkin, N., Zaydelman, L., Kotov, A.: The role of oriented gestures during robot's communication to a human. In: Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2019". **18**(25), 800-808 (2019)
22. Zinina, A., Zaydelman, L., Arinkin, N., Kotov, A.: Emotional behavior of robots during of the spatial puzzle solving with children. In: Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2020". **19**(26), 811-826 (2020).

23. Jones, A., Küster, D., Basedow, C. A., Alves-Oliveira, P., Serholt, S., Hastie, H., Castellano, G.: Empathic robotic tutors for personalised learning: A multidisciplinary approach. In: International conference on social robotics. pp. 285-295. Springer, Cham (2015).
24. Castellano, G., Leite, I., Pereira, A., Martinho, C., Paiva, A., Mcoan, P. W.: Multimodal affect modeling and recognition for empathic robot companions. International Journal of Humanoid Robotics. **10**(01), 1350010 (2013).
25. Scherer, U., Helfrich, H., Scherer, K. R.: Internal push or external pull? Determinants of paralinguistic behavior. In: Giles, H., Robinson, P., Smith, P. (eds.) Language: Social Psychological Perspectives. pp. 279-282. Oxford: Pergamon. (1980).