BabyRobot - Next Generation Social Robots

Enhancing communication and collaboration development of TD and ASD children by developing and commercially exploiting the next generation of human-robot interaction technologies.

Alexandros Potamianos[1], Costas Tzafestas[1], Elias Iosif [1], Franziska Kirstein[2], Petros Maragos[3], Kerstin Dauthenhahn[4], Joakim Gustafson[5], John Erland Østergaard[1], Stefan Kopp[6], Preben Wik[7], Oliver Pietquin[8], Samer Al Moubayed[7]

[1]Institute of Communication and Computer Systems, National Technical University of Athens, Greece
[2] Blue Ocean Robotics, Odense, Denmark
[3]Athena Research and Innovation Center in Information Communication and Knowledge Technologies, Athens, Greece
[4] School of Computer Science, University of Hertfordshire, Hatfield, UK
[5] Centre for Speech Technology, KTH, Stockholm, Sweden
[6] Computer Science, Faculty of Technology, Universitaet Bielefeld, Bielefeld, Germany
[7]Furhat Robotics AB, Bandhagen, Sweden
[8] CRIStAL (UMR 9189), Univ. Lille, CNRS, Centrale Lille, Inria - Institut Universitaire de France

potam@central.ntua.gr, ktzaf@cs.ntua.gr, iosife@central.ntua.gr, fk@blue-ocean-robotics.com, maragos@cs.ntua.gr, k.dautenhahn@herts.ac.uk, jocke@speech.kth.se, jeo@blue-ocean-robotics.com, skopp@techfak.uni-bielefeld.de, preben@furhatrobotics.com, olivier.pietquin@univ-lille1.fr, samer@furhatrobotics.com

***Abstract*— We present the main ideas of the recently initiated EU-IST H2020 project "BabyRobot". The project is a collaborative, multidisciplinary effort for developing and commercially exploiting the next generation of human-robot interaction technologies in order to promote the adoption of robotic systems in educational settings, consumer applications and beyond.**

Keywords—Child-Robot Communication and Collaboration, Multimodal Interaction, Child-Robot Interaction, Evaluating Child-Robot Interaction, Spoken Dialogue Systems

I. Introduction

Children-robot communication is an emerging, promising field that, at the same time, presents new challenges, e.g., speech recognition for young children. Children are special in that they are fast and adaptive learners, always curious, especially with novel robotic technologies in, e.g., educational and edutainment settings. In addition, children have unique communication skills, meaning that they can easily convey or share complex information and intentions with only some or even no spoken language at all by, for example, efficiently using gestures instead. These, abilities make children a perfect target group for studying the effect of human-robot interaction on the development and enhancement of communicative abilities and shared goal formation. Autistic spectrum children, who often have delayed or not fully developed social, communication and collaboration skills without therapy, can especially profit from the aforementioned interaction

In this paper, we present the main ideas of the EU-project "BabyRobot". BabyRobot is a new European Project within the framework of Horizon 2020 (Information and Communication Technologies, Research & Innovation Action). It is a collaborative effort, based on the common vision of a team of experienced researchers and technologists for developing and commercially exploiting the next generation of human-robot interaction technologies in order to promote the adoption of robotic systems in educational settings, consumer applications and beyond.

The creation of these technologies requires inter- and multi-disciplinary expertise ranging from core robotic technologies to human-robot interaction, multimedia processing, behavioral informatics and machine learning. The consortium is built around research and technological excellence in these areas [1, 2, 3, 4, 5, 6, 7, 8].

II. About BabyRobot

Humans, when communicating, have the unique ability to share intentionality, create and execute on joint plans. Thus, we model human-robot communication as a three-step process: sharing attention, establishing common ground and forming shared goals. Requirements for successful communication are to be able to decode the cognitive state of other humans (intention-reading) and to build trust. In BabyRobot, our main goal is to create robots that analyze and track human behavior over time in the context of their surroundings (situational) using audio-visual monitoring in order to establish common ground and intention-reading capabilities (see Figure 1). In BabyRobot we focus on the typically developing and autistic spectrum children user population. Children have unique communication skills, are quick and adaptive learners, eager to embrace new robotic technologies. This is especially relevant for special education where the development of social skills is delayed or never fully develops without intervention or therapy. Thus, our second goal is to define, implement and evaluate child-robot interaction application scenarios for developing specific socio-affective, communication and collaboration skills in typically developing and autistic spectrum children. We will support not supplant the therapist or educator, working hand-in-hand to create a low risk environment for learning and cognitive development. Breakthroughs in core robotic technologies are needed to support this research mainly in the areas of motion planning and control in constrained spaces, gestural kinematics, sensorimotor learning and adaptation. Our third goal is to push beyond the state-of-the-art in core robotic technologies to support natural human-robot interaction and collaboration for consumer, edutainment and healthcare applications. BabyRobot's ambition is to create robots that can establish communication protocols and form collaboration plans on the fly that will have impact beyond the consumer and healthcare application markets addressed here.

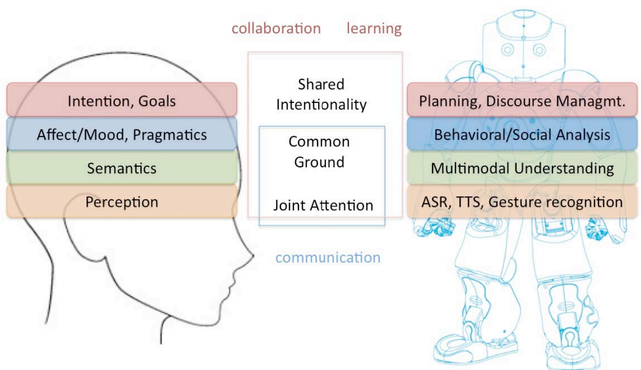

Fig. 1. Main concept of the BabyRobot Project (from [17])

In the BabyRobot project, we will integrate technologies on three different robot platforms: Furhat, Zeno, and KASPAR. Each platform possesses its own capacities and potential for technology implementation and integration. Furhat [8], is a robotic platform very much focused on spoken dialogue and multimodal human-robot interaction. ZENO [6] is a general-purpose commercial platform, which is now widely used in several application domains including child-robot edutainement scenarios. KASPAR [7] is a research platform that possesses some unique features enabling it to be effectively used in social human-robot interaction paradigms

In BabyRobot we will augment the functionality of the robotic platforms with advanced behavior tracking and audio-visual functionality, reinforcement learning and adaptation for long-term interaction, as well as advanced multiparty communication and collaboration capabilities. The development will be both motivated and validated by the three use cases of BabyRobot. These technologies will be designed in a modular and scalable manner in order to be ready for integration also on other robotic platforms.

III. METHOD

A. Use Cases in BabyRobot

Selecting appropriate use cases is essential for algorithmic design, technological development, as well as evaluation of platforms and services. It is especially hard to create HRI scenarios for children users that are at the same time fun, engaging and challenging across age-groups and across users with different level of communication and cognitive skills. Turning to the literature, in the EMOTE project two scenarios are considered, a strongly engaging wack-a-mole style game and a non-engaging touch button task [9]. In LIREC children play chess with the iCat robot and are being taught melodic composition of music by the virtual agent Pequeno [10]. In ALIZ-E three gaming scenarios are used: the "Simon says" game, the dancing lessons game and the quiz game [11]. Finally, in KASPAR a collaborative copycat game is used, in a dyadic (child and robot) or triadic (two children and a robot) setting [12]. The goal is to facilitate social interaction and engagement among autistic children. These skills are developed working with the robot and then used when interacting with humans.

BabyRobot is centered around three use cases that will help identify, ground, develop and evaluate the relevant set of technologies, as well as their application to child-robot interaction scenarios, specifically: 1) natural child-robot interaction scenarios that showcase the joint attention, common ground and shared intentionality modules, 2) child-robot communication and collaboration skill development games, and 3) multi-party collaboration skill development via dyadic and triadic interaction, using the robot as a mediator. The detailed scenarios for each use case will be determined in the first months of the project.

In use case 1, we aim to make use of a setup where a group of children and FurHat robot interact about virtual and physical objects on an interactive table. This is a collaboration exercise that requires joint attention. We will design a game that simulates a situation where the task of the robot is not to assist physically, but to guide a group of children verbally in their problem solving. The game will involve a number of tasks that will require coordinated manipulations of different sets of physical or virtual objects on the smart table. The robot will have access to step-by-step instructions on what and when each of them has to perform certain tasks in order to reach the goal. In the game scenario, the robot will also have access to information unavailable to the child (e.g., about the setting and objects). The task will require the children to perform synchronized actions on several objects in order to solve the task. The role of the robot will be to assist in this coordination, by tracking the visual attention of all humans and provide targeted and timely instructions using a combination of verbal instructions and gaze that accurately picks out both the object and the child that is being addressed. For ASD children, the task can be simplified by involving only one child and less complex objects. Furthermore, the smart table can be used to provide supportive visual feedback on joint attention.

The design, development and evaluation of the communication and learning capabilities will be grounded on child-robot interactive scenarios that constitute the second use case of BabyRobot. Additional algorithmic and software development will be required to add these new capabilities to the robotic platform. Long-term experiments (case study evaluations in schools) will verify the progress in the communicative/collaborative abilities of participants using the data collected. In addition, the communicative capabilities of the robot will be enhanced via reinforcement learning and language learning from a minimal set of examples using shared attention and common grounding modules [13]. Learning scenarios in the literature are based on the "learning from demonstration" principle (LfD) in [14], learning by imitation for kinetic tasks in [15] and learning from partial observations via a reinforcement-learning paradigm in [16]. In use case 2, we will use the humanoid robot Zeno by Robokind that is also able to show facial expressions.

Use case 3 we will study learning and communication of children with ASD in dyadic and triadic interaction games, using verbal and non-verbal modes of interaction that will be integrated in order to provide enjoyable scenarios where children play computer games with KASPAR and/or other people. The robot will use information on the state of the game, the overall objectives of the game, and the behavior of the children in order to motivate, engage and encourage the children to participate in collaborative games. The robot will have dual roles in this case study: as a mediator (mediating child-child interaction) as well as a game partner who is able to autonomously participate in the interaction.

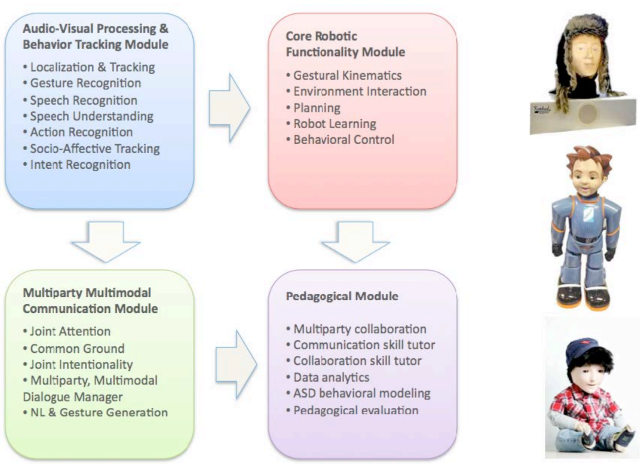

Fig. 2. BabyRobot modular architecture and robotic platforms.

B. *Participants*

Subjects are recruited by the partners via local schools in UK, Denmark, Sweden and Greece that they have been collaborating with in the past. The children to be recruited will be in the age range of 8-12 for ASD children and 6-10 for TD children. Selected ASD children will have good verbal skills. The children will be used as their own controls in case-study evaluations that allows taking an inclusive approach and not prescribe a fixed IQ or verbal age as required in clinical randomized control trials. The participants will therefore include children with a wide range of abilities and skills, which allows identifying the children with a level of functioning who will most benefit from the new technology developed in the project.

C. *Evaluation*

Evaluation metrics will be defined to evaluate all aspects of BabyRobot activities, namely, algorithms, technological claims and robotic platforms. Technical evaluation includes the evaluation of the algorithms proposed for audio-visual processing, behavior tracking, multimodal communication and learning (across use cases and across languages). Robotic platforms will be evaluated both subjectively (in terms of user satisfaction, likability, etc.) and objectively in terms of percent success in communicative and collaborative scenarios, child-robot interaction time, adaptation performance/learning rate of the robotic communication and collaboration capabilities, robustness to environmental changes, etc. Finally, the technological claims for improving the communication and collaboration skills of children via robotic edutainment scenarios will be evaluated using ASD therapists and educators in longitudinal experiments. Specifically, for the evaluation of the audio-visual processing and behavior tracking algorithms we will use objective metrics. Subsets of the collected data (test set) will be annotated, labelled and transcribed and algorithmic results will be evaluated against these labels.

An important part of the evaluation is the learning rate/adaptation of the robot that will be measured both objectively in terms of accuracy of the learned communication codes and collaboration goals compared to manually labelled ones, as well as in terms of improvement of performance of the end-of-end system (see efficiency, satisfaction, engagement metrics above). The capacity of the humanoid robots to learn new skills (gesticulation, environment interaction) will be assessed using both objective and subjective measures, including efficiency and efficacy (accuracy on reproducing a skill, quantitative and/or qualitative effects on human participants, etc.). Last but not least, the communication and collaboration skills of children participating in the use cases, will be evaluated from longitudinal data both manually from educators and ASD therapists using questionnaires and exit interviews, as well as in a machine-assisted framework by analyzing audio-visual and behavioral data automatically.

IV. CHALLENGES, SOLUTION AND IMPACT

BabyRobot addresses an important commercial opportunity in the area of robotics and human-robot interaction. Limited cognitive capabilities, learning/adaptation, communication and collaboration skills are the major bottleneck towards the adoption of robots in new roles as companions, assistants, social mediators, therapy mediators etc. in consumer, healthcare applications and beyond. Our market analysis and results from earlier projects indicate that there is a semantic gap between market/end-user expectations and what robotics are able to provide today. Specifically, there are high expectations regarding cognitive behavior (voice/speech recognition, person/face detection, recognition of objects etc.) and cognitive autonomy such that robots are aware of their environment during an interaction. These two challenges are directly connected to BabyRobot's main goals: creating robots that can establish communication protocols and form collaboration plans on the fly will have impact beyond the consumer and healthcare application scenarios envisioned here. The technologies developed in BabyRobot will act as enablers, lowering the barrier to entry for Europe's SMEs, improving the quality and cost-effectiveness of prototyping novel human-robot communication and collaboration scenarios, as well as facilitating the adoption of robot technologies in everyday life. The exploitation potential of such robots is large.

ACKNOWLEDGMENT

The research leading to this paper has been partially funded by the European Community's H2020 Program (Information and Communication Technologies) under grant agreement no. 687831, Babyrobot.

REFERENCES

[1] S. Narayanan and A. Potamianos, "Creating conversational interfaces for children," IEEE Transactions on Speech and Audio Processing, vol. 10, pp. 65-78, Feb. 2002. [IEEE Signal Processing Society Best Paper Award 2005]

[2] M. Geist & O. Pietquin, "An Algorithmic Survey of Parametric Value Function Approximation", IEEE Transactions on Neural Networks and Learning Systems, 24(6):845 – 867, 2013.

[3] M. Salem, F.A. Eyssel, K. Rohlfing, S. Kopp & F. Joublin, "To Err is Human(-like): Effects of Robot Gesture on Perceived Anthropomorphism and Likability", International Journal of Social Robotics, 5(3):313–323, 2013.

[4] P. Maragos, P. Gros, A. Katsamanis and G. Papandreou, "Cross-Modal Integration for Performance Improving in Multimedia: A Review", in Multimodal Processing and Interaction: Audio, Video, Text, (edited by P. Maragos, A. Potamianos and P. Gros), Springer, 2008.

[5] S. Al Moubayed, G. Skantze, & J. Beskow, "The Furhat Back-Projected Humanoid Head - Lip reading, Gaze and Multiparty Interaction", International Journal of Humanoid Robotics, 10(1), 2013.

[6] F.Kirstein, R. V. Risager (accepted). Social Robots in Educational Institutions. In Proceedings of the 2016 ACM/IEEE international conference on Human-robot interaction ACM.

[7] J. Wainer, B. Robins, F. Amirabdollahian, K. Dautenhahn, "Using the humanoid robot KASPAR to autonomously play triadic games and facilitate collaborative play among children with autism", IEEE TAMD (IEEE Trends in Autonomous Mental Development) 6(3): 183-199, 2014.

[8] Johansson, M., Skantze, G., & Gustafson, J. (2014, November). Comparison of Human-Human and Human-Robot Turn-Taking Behaviour in Multiparty Situated Interaction. In *Proceedings of the 2014 worksho*p on Understanding and Modeling Multiparty, Multimodal Interactions (pp. 21-26). ACM.

[9] Jones, A., Küster, D., Basedow, C. A., Alves-Oliveira, P., Serholt, S., Hastie, H., Lee, C., Barendregt, W., Kappas, A., Paiva, A., & Castellano, G. (2015). Empathic Robotic Tutors for Personalised Learning: A Multidisciplinary Approach. In *Social Robotics* (pp. 285-295). Springer International Publishing.

[10] Koay KL, Syrdal DS, Dautenhahn K, Arent K, Małek Ł, Kreczmer B. 2011. Companion Migration – Initial Participants' Feedback from a Video-Based Prototyping Study. Mixed Reality and Human-Robot Interaction. 47 Wang, Xiangyu (Ed.):133–151.

[11] Baroni, I., Nalin, M., Baxter, P., Pozzi, C., Oleari, E., Sanna, A. et al (2014). What a Robotic Companion Could Do for a Diabetic Child.. In 23rd IEEE International Conference on Robot and Human Interactive Communication (RoMAN 2014)

[12] S. Costa, H. Lehmann, K. Dautenhahn, B. Robins, F. Soares, "Using a humanoid robot to elicit body awareness and appropriate physical interaction in children with autism", International Journal of Social Robotics 7: 265-278. Publisher: Springer, 2014

[13] Lemon, O. & Pietquin, O. (2012), "Data-Driven Methods for Adaptive Spoken Dialogue Systems: Computational Learning for Conversational Interfaces" Lemon, O. & Pietquin, O. eds., November, 2012, Springer 177 pages.

[14] Argall, B. D., Chernova, S., Veloso, M. and Browning, B. (2009). "A Survey of Robot Learning from Demonstration". In Robotics and Autonomous Systems, 57 (5) pp. 469-483.

[15] J. Clerk Calinon, S., D'hHalluin, F., Sauser, E. L., Caldwell, D.G., and Billard, A.G. (2010). "Learning and reproduction of gestures by imitation". IEEE Robotics & Automation Magazine, 17(2): 44-54, June 2010.

[16] Thomaz, A. L., and Breazeal, C. (2008). "Teachable robots: Understanding human teaching behaviour to build more effective robot learners". In Artificial Intelligence, 172 (6), pp. 716-737.

[17] A. Vinciarelli, A. Esposito, E. Andre, F. Bonin, M. Chetouani, J. Cohn, M. Cristani, F. Fuhrmann, E. Gilmartin, Z. Hammal, D. Heylen, R. Kaiser, M. Koutsombogera, A. Potamianos, S. Renals, G. Riccardi, and A. Salah, "Open challenges in modelling, analysis and synthesis of human behaviour in human-human and human-machine interactions," *Cognitive Computation*, vol. 7, pp. 397--413, Apr. 2015.