Je-LKS
Journal of e-Learning and Knowledge Society
The Italian e-Learning Association Journal
PEER REVIEWED PAPERS
EDUCATIONAL ROBOTICS: RESEARCH AND PRACTICES OF ROBOTS IN EDUCATION
Vol. 15, n.2, 2019
ISSN: 1826-6223 | eISSN: 1971-8829

ADVANCED PROGRAMMING OF INTELLIGENT SOCIAL ROBOTS

Donato Malerba, Annalisa Appice, Paolo Buono, Giovanna Castellano, Berardina De Carolis, Marco De Gemmis, Marco Polignano, Veronica Rossano, Lynn Margaret Rudd

Dipartimento di Informatica
Università degli Studi di Bari Aldo Moro, Italy
{name.surname@uniba.it}

Keywords: Social Robot Programming, Machine Learning, Human–Robot Interaction, Education Technology, Affective Computing

Robotics in education is a promising new area: social robots have started to move into schools as part of educational/learning technologies, playing roles in educational settings that range from tutors, teaching assistants and learners, to learning companions and therapeutic assistants. This paper provides an overview of the main computational methods required to program a social robot and equip it with social intelligence. Some applications of social robots in the field of education are reported to show how the use of educational robots may innovate the learning process at different levels and in various contexts.

for citations:
Malerba D. *et al.* (2019), *Advanced Programming of Intelligent Social Robots,* Journal of e-Learning and Knowledge Society, v.15, n.2, 13-26. ISSN: 1826-6223, e-ISSN:1971-8829
DOI: 10.20368/1971-8829/1611

1 Introduction: the context and the target

Social Robots are embodied autonomous intelligent entities that interact with people in everyday environments, following social behavior typical of humans. Since they are autonomous, robots should be able to interpret human behavior properly, to react to changes during the interaction, make decisions, behave in a socially plausible manner, and learn from a user's feedback and previous interactions. Social robots are mainly used to improve people experience in diverse application domains. In education, for example, robots have been shown to be successful in diverse contexts, such as math tutoring (Kennedy *et al.*, 2015), social skill training for children with autism (Wainer *et al.*, 2014) and language teaching (Schodde *et al.*, 2017). In particular, the use of robots may increase attention, engagement, and compliance, which are critical components of successful learning (Ramachandran *et al.*, 2018).

This paper provides an overview of the main computational methods required to program a social robot and equip it with intelligence, so that it can be employed successfully to enhance the learning process both in formal and informal settings. The first part of the paper introduces the fundamental skills that should be implemented in a robot to achieve social intelligence, namely sensing, dialogue management, emotion recognition and user modeling. During the interaction, that is achieved through human-oriented sensing tasks and multimodal dialog management, social robots can use machine learning methods to modify the way they provide information, according to the user's needs and emotional reactions. Currently, educational robotics is one of the most promising technologies to improve teaching and learning effectiveness. The second part of the paper focuses on applications that can take advantage of social robots in the realm of education. We report some experiences with the design and the use of educational robots that can inspire new ways to innovate learning processes in various contexts.

2 Skills for Social Intelligence

Social intelligence factors increase the complexity of programming a socially interactive robot. A social robot is expected to sense its surrounding, to handle natural and multimodal dialogs, to recognize and express emotions, and to adapt the interaction to some characteristics of the user. All these skills are the basis for human social behavior models.

2.1 Sensing

Robots can sense the environment by means of integrated sensors or computer vision. For example, the Pepper humanoid robot (Lafaye, 2014) is equipped with both sensors that enable it to perceive the surrounding environment and sensors, like microphones, cameras and touch. Moreover, the tablet PC on the robot's torso allows interaction by means of a touch screen.

Speech recognition techniques are widely used in social robots (Amodei, 2016; Zhang, 2017). Many organizations have launched their own Deep CNN models to improve the accuracy of voice recognition (Xiong, 2017). Advances in this field have paved the way for the development of high-level tasks, such as semantic recognition and semantic understanding, i.e. how to formally represent the meaning of a text, that form the basis for the robot's dialog abilities, as explained in Section 2.2.

Computer vision is fundamental for the recognition of human facial expressions and movements (Canal *et al.*, 2016). The main challenge is to find a representation that can be adapted to a new task with few training data available. State-of-the-art pipelines for single-view action recognition are hand-crafted dense trajectory features and 3D CNN-based features. The CNN-based features are extracted from these intermediate layers and then fed into an SVM for the final classification. In the field of education, problems such as occlusions or poor camera view point can often occur, since the interaction can change abruptly. Therefore, robust multi-view action recognition systems are required (Efthymiou *et al.*, 2018).

Other common computer vision tasks are face recognition and detection of facial expressions and emotions, which are useful to convey the user's feelings to the robot. Specifically, emotion recognition is very important to allow the robot to adapt its behavior (e.g. showing empathy). More details are discussed in Section 2.3.

In general, social robot sensing skills leverage machine learning methods to learn models of human social cognition starting from features extracted automatically from sensory data. The challenge is a fast processing of sensory data, in order to draw conclusions, which may help in the decision of the actions to be performed. Time series algorithms to discover recurrent patterns are mainly investigated in machine learning, in order to address problems of gesture discovery, synchrony discovery, differential drive motion pattern discovery and motor primitive discovery from observations of human behavior (de Jong *et al.*, 2018).

2.2 Dialog Management

In order to obtain a reasonable level of interaction in the conversation, the robot should be able to handle dialogs, that is, modulate the initiative, handle communication interferences, make inferences related to the sentences pronounced by people, plan, organize and maintain the discourse. Starting from the perception (sensing) of the multimodal user input, according to the extracted meaning (semantic interpretation), a dialog manager has to decide (reason) how to respond to the user in a socially believable way, in order to handle the dialog flow (Figure 1). An example of a grounding-based model, implemented to this aim, is given with BIRON (Spexard *et al.*, 2006).

Usually dialog management is based on transition networks, frames (McTear, 1993), information state models (Larsson & Traum, 2000), rule-based models, or planning techniques typical of Belief–Desire–Intention (BDI) agents (Wong *et al.*, 2012). Finite state models are the simplest way to handle the dialog and are suited to applications where the dialog flow coincides with the task structure, however, they lack flexibility. Frame-based approaches are based on the structure of the entities in the application domain. The information state approach has been widely used in conversational systems. It is based on the idea that the dialog flow changes according to the dialog state, that is represented by the current topic, the recent dialog moves and information about the beliefs of the dialog participants. Planning is a more complex approach, but it can deal with changes in behavior that are required when reacting to real-world interaction. A good compromise is a mixture of the two approaches: follow a predefined path, presented as transition networks and replan only when necessary.

Finally, socially guided machine learning can use natural interaction to teach a machine new knowledge and skills (de Greeff & Belpaeme, 2015), while deep learning methods have recently been employed for activity recognition (Mohammad *et al.*, 2018). Machine learning techniques can be used to infer behavioral patterns and interaction protocols. They are explored to identify utterance vectors, typical utterances, stopping locations, motion paths and spatial formations of both human and robot participants in the environment and to train a robot to generate multimodal actions (Liu *et al.*, 2018).

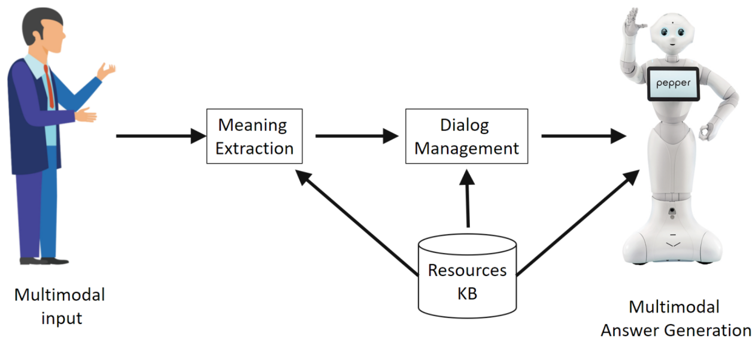

Fig. 1 - Overview of a generic dialog system

2.3 Emotion Recognition

In HRI facial expressions and speech are the most important communication channels that can be analyzed to detect and recognize emotions (Mavridis, 2015).

Ekman *et al.* (2002) defined six basic emotions that are universally recognized from facial expressions, regardless of culture. Other models, such as a Facial Action Coding System (FACS) (Ekman, 1999), can be taken into account when referring to Facial Expression Recognition (FER). A FER system has to be trained on suitable datasets (Ko, 2018) and should include the steps of a pipeline which are typical of this application domain: preprocessing, face detection and registration, feature extraction and classification. Recently deep learning methods have been used in this context, achieving state-of-the-art recognition accuracy and exceeding previous results by a long chalk (Kahou, 2015; Walecki, 2017).

Speech conveys affective information through the explicit linguistic message (what is said) and the implicit paralinguistic features of the expression (how it is said). Speech Emotion Recognition (SER) is basically performed through pure sound processing without linguistic information (Schuller, 2018). Features can be of several types and related to voice prosody, acoustic properties and transformations. In particular, Mel-Frequency Cepstral Coefficients (MFCCs), formants, energy, fundamental frequency (pitch) and temporal features have been used successfully in emotion recognition (Schuller, 2018). Then, as for FER systems, a classifier is trained using machine learning techniques. Also in this domain deep neural networks significantly boosted the performance of emotion recognition models (Fayek *et al.*, 2017).

Face and speech can be analyzed simultaneously and combined to obtain a more robust emotion recognition system by means of fusion techniques (Busso *et al.*, 2004; Zeng *et al.*, 2007; Haq *et al.*, 2008 Sebe *et al.*, 2006; De Carolis *et al.*, 2017a).

Causality analysis is also important for social robots because it allows them to discover the causal structure of a human’s behavior during the interaction (Yamashita *et al.*, 2018).

2.4 User modeling

User-adaptive systems rely on a user model, which is a structured representation of user characteristics that may be relevant for personalized interaction (Fischer, 2001; Kobsa, 2001). In the context of social robots, the user profile includes several dimensions, such as age, gender, level of expertise in a given task, emotions, personality and past interactions (Ahmad, 2017), that allow the robot to make decisions. In general, the user model is explicitly designed to facilitate decision-making in the specific field where the robot is involved. For instance, in the health care domain the user profile could store information about the user performance in a given task. In (Tapus, 2009) the robot collected data on the user's reaction time and number of correct answers in a cognitive task and adapted the dialog to motivate the user, according to the results.

Moreover, in the education domain user performance is used to adapt decision-making as well as the verbal and non-verbal behavior of the robot. For instance, in (Brown, 2013) the robot modifies its supportive feedback according to the user's behavioral state (e.g. "unmotivated"), determined by monitoring the student's interactions with the robot when performing a mathematics test.

The application of social robots in public spaces has several challenges from the point of view of user modeling, because the robot is involved in multiple short-term interactions with unknown people rather than in a long-term interaction with a known user. In public areas, such as malls, the user profile stores the information available about the user who is interacting with the robot. For instance, the profile could be acquired by an RFID reader or swiping a fidelity card (Iocchi, 2015). The communication activities and actions are personalized according to the profile, in order to increase the robot's social acceptability.

3 Applications in Education

In the following, we provide some recent examples of the usage of social robots in education (Figure 2), in order to provide a guide for researchers who consider using social robots for different educational purposes:

- Effectiveness: to support knowledge and skill acquisition;
- Engagement: to make children more involved in learning activities;
- Special needs: to support learners with specific difficulties;
- Empowerment of young patients: to educate patients to adopt a healthy lifestyle and to support patients and caregivers in managing specific medical situations;
- Language learning: to support vocabulary learning.

3.1 Effectiveness

One of the main goals investigated in the literature is the use of a human robot as a teacher or tutor to encourage active learning (Bonwell & Eison, 1991), where the teacher becomes a tutor, thus enhancing students' self-confidence and independence. Active learning is used to model a learning agent that can shape its learning experience through interaction with its teacher. Active learning between a robot learner and a human teacher leads to more effective faster learning (de Greeff *et al.*, 2012).

Fig. 2 - Social robots commonly used in education

Leyzberg *et al.* (2012) investigate if the presence of a robot tutor, Keepon (Figure 2), can influence the student's learning gain. The results confirm that the physical presence may imbue the robot with more perceived authority than an on-screen agent.

NAO robot (fig. 2) is used as a tutor in a basic arithmetic learning task. Here again students show a higher level of motivation, which usually results in better learning gain (Janssen *et al.*, 2011). Other researchers have explored the effects of social robots as tutors versus teachers. Howley *et al.* (2014) measure how the social role of both a human and a robot affects help-seeking behaviors and learning outcomes in a one-to-one tutoring setting. The results confirm that the use of pedagogical robotic agents can be more beneficial for learning than the use of human pedagogical agents. Some research investigates the effectiveness of social robots using the peer-tutoring approach. In this setting the social robot acts as the learner's companion. Tanaka *et al.* (2015) develop an application for Pepper to enable it to learn together with children. Baxter *et al.* (2017) propose a robot peer with personalized behavior in collaborative learning tasks with individual children. Both solutions facilitate the learning process and allow children to improve their knowledge and skills.

3.2 Engagement

The effectiveness of using social robots to support user engagement has been widely proved in the literature. Engagement has been identified as a key aspect in Technology Enhanced Learning research, in order to sustain interest, participation and involvement during a learning process (Carini *et al.*, 2006). To be engaging a social robot should proactively involve people in a social context by expressing and perceiving emotions. Another factor that is relevant to engagement is anthropomorphism, since it recalls in the user conversational patterns typical of human-to-human interactions (Bainbridge *et al.*, 2008). In educational contexts student engagement improves learning, therefore, social robots have recently been employed to do so. In (Castellano *et al.*, 2009), the iCat robot (fig. 2) during a chess game displayed affective reactions, in order to improve the user's engagement. To improve the engagement of students, humanoid robots have been used as pro-active tutors (Gudi *et al.*, 2019). The robot can influence the pace of the interaction in a social learning task by increasing the students' learning experience, and thus their engagement even if they do not perceive it (Ivaldi *et al.*, 2014).

3.3 Special Needs

Social robots are often applied in education of students with special needs. Teaching and learning of disabled youngsters pose unique and distinctive challenges. These students demand more time and patience. They require specialized instructional strategies in structured environments, in order to support and enhance their learning potential. Social robots are widely applied to teach basic social skills to children with autism, since they resemble humans but are less complex and seem to be able to manage these issues successfully (Palestra *et al.*, 2017). Wainer *et al.* (2014) employ the KASPAR robot to improve their cooperation skills. Pennazio (2017) used the IROMEC platform to improve human interaction skills. Boccanfuso *et al.* (2017) and Alemi *et al.* (2015) investigate the acquisition of communication skills in a language learning scenario. Pale Social robots are successfully used also in other contexts, for example, to support children affected by dyslexia (Pistoia *et al.*, 2015), or to stimulate social interaction in children with Down's syndrome (Lehmann *et al.*, 2014). They are also applied to people with speech and hearing impairments in sign language learning (Gudi *et al.*, 2019).

3.4 Empowerment

The success and effectiveness of the use of technology to support therapeutic education is increasing. Various ICT solutions address patient empowerment

(Di Bitonto *et al.*, 2012). For example, in medical contexts empowerment refers to the patient's acquisition of knowledge about his/her clinical conditions and the acquisition of a suitable lifestyle to ensure a good quality of life. Social robots, such as NAO, have been applied to support knowledge acquisition in children with Type I Diabetes who have to learn basic knowledge about diet and management of their illness (Coninx *et al.*, 2016; Cañamero & Lewis, 2016). Other experiments have been conducted with children who interact with a social robot to improve their knowledge and habits with regards to a healthy life-style (Ros *et al.*, 2016). A current challenge (Share *et al.*, 2018) is to use social robots as Assistive Technology in the field of care for the elderly.

3.5 Language Learning

Robot Assisted Language Learning (RALL), particularly for L2 (second language) learning, has proved to be more effective in boosting learner performance and motivation compared with just 2D screen-based technologies (Belpaeme T. et al, 2011). The type of applications described in the literature regard mainly vocabulary learning of the L2 language through various models (e.g. robot as storyteller, robot asking questions and checking learners' answers and robot playing charade games with learners). RALL through games has proved to be the most relaxing and enjoyable interaction and therefore the most profitable for L2 learners (Mubin O. et al, 2013). The role of the robot is usually either as a peer tutor or a teacher's assistant. The former role is the most common, but the latter role has often proved to be the most effective for L2 learners, where the teacher is present to explain any possible misunderstandings and the interactions are based on the curriculum already in use in class (Lee S. et al, 2010). The learners' performance gain has also been evaluated by pre-tests and post-tests and their motivation has been assessed by questionnaires (Schodde et al, 2017).

Conclusions

This paper outlines the key components of social intelligence and proposes a framework of design issues for the advanced programming of social robots. The state-of-the-art covers various aspects related to social robot programming and highlights the importance of sensing, user modeling and emotion recognition in accomplishing fundamental tasks related to social robot behavior. The experiences described in several educational scenarios show that the integration of social robots in education may improve student engagement and empowerment, especially students with special needs.

We are already experimenting the role of the robot in education in various forms. With reference to the list of educational applications in education we experimented the use of several robots, and specifically two Pepper humanoid robots to verify:

- effectiveness through a game where the pupil shows a disposable item and competes with the robot guessing how to correctly separate waste;
- engagement by entertaining people during big events at the University of Bari, providing people with directions or information related to the event or the courses and services offered by the Computer Science Department;
- special needs using puppies such as a dog robot, a dragon or NAO humanoid robot to overcome problems related to autism (Palestra *et al.*, 2017);
- empowerment of young patients, using a storytelling approach Pepper introduces juvenile diabetes to young patients and their classmates to make them aware of the lifestyle required by the disease;
- language learning through a game to teach pupils new terms, as well as associations between terms, while playing. We developed an artificial player for a challenging language game: the player is given a set of five words - the clues - each linked in some way to a specific word that represents the unique solution of the game. Words are unrelated to each other, but each of them has a hidden association with the solution (Basile *et al.*, 2016).

Despite the many open questions and challenges that are still to be faced in programming social robots, it is expected that soon robots will have great impact in various areas of education. This increasing impact will not replace human teachers, but will provide added value in the form of a stimulating and instructive teaching support.

Acknowledgments

Funding for this work was provided by the Fondazione Puglia that supported the Italian project "Programmazione Avanzata di Robot Sociali Intelligenti".

REFERENCES

Ahmad M.I., Mubin O. & Orlando J. (2017), *A Systematic Review of Adaptivity in Human-Robot Interaction*, Multimodal Technologies and Interactions, 3(14), 1–25.

Alemi M., Meghdari A., Basiri N. M. & Taheri A. (2015), *The effect of applying humanoid robots as teacher assistants to help Iranian autistic pupils learn English* *as a foreign language*, in Tapus A., André E., Martin J.C., Ferland F., Ammi M. (eds), Proc. of ICSR 2015. LNCS, vol 9388, 1-10, Paris, France, Springer, Cham

Amodei D., Ananthanarayanan S., Anubhai R. *et al.* (2016), *Deep speech 2: end-to-end speech recognition in English and Mandarin*, in Balcan M.F., Weinberger, K.Q. (eds), Proc. of ICML 2016. 173-182, New York, USA, JMLR.org

Bainbridge W. A., Hart J., Kim E. S., & Scassellati B. (2008), *The effect of presence on human-robot interaction*, in Buss M., Kühnlenz K. (eds), Proc. of RO-MAN 2008. 701-706, Muich, Germany, IEEE.

Basile P., de Gemmis M., Lops P. & Semeraro G. (2016), *Solving a Complex Language Game by using Knowledge-based Word Associations Discovery*, IEEE Transactions On Computational Intelligence and AI in GAMES, 8(1), 13-26.

Baxter P., Ashurst E., Read R., Kennedy J. & Belpaeme T. (2017), *Robot education peers in a situated primary school study: Personalisation promotes child learning*, PloS one, 12(5), 1-23.

Belpaeme T., Kennedy J., Baxter P. & Deblieck T. (2011), *L2TOR- Second Language Tutoring using Social Robots*, in ADFA, Springer, 1-1.

Boccanfuso L., Scarborough S., Abramson R.K., Hall V.H., Wright H. & O'Kane J.M. (2017), *A low-cost socially assistive robot and robot-assisted intervention for children with autism spectrum disorder: field trials and lessons learned*, Auton. Robots, 41(3), 637-655.

Bonwell C. & Eison J. (1991), *Active Learning: Creating Excitement in the Classroom*, AEHE-ERIC Higher Education Report, No. 1. Washington, D.C., Jossey-Bass.

Brown L., Kerwin R. & Howard A.M. (2013), *Applying behavioral strategies for student engagement using a robotic educational agent*, in SMC 2013. 4360–4365, Manchester, UK, IEEE.

Busso C., Deng Z., Yildirim S., Bulut M., Lee C.M., Kazemzadeh A. Lee S., Neumann U. & Narayanan S. (2004), *Analysis of emotion recognition using facial expressions, speech and multimodal information*, in Proc. of ICMI 2011. 205-211, State College, PA, USA ACM.

Canal H., Escalera S. & Angulo C. (2016), *A real-time human-robot interaction system based on gestures for assistive scenarios*, Computer Vision and Image Understanding, 149, 5–77.

Cañamero L. & Lewis M. (2016), *Making New "New AI" Friends: Designing a Social Robot for Diabetic Children from an Embodied AI Perspective*, Int J of Soc Robotics, 8(4), 523–537.

Carini R.M., Kuh G.D. & Klein S.P. (2006), *Student engagement and student learning: testing the linkages*, Res. High. Educ, 47, 1–32.

Castellano G., Pereira A., Leite I., Paiva A. & McOwan P. W. (2009), *Detecting user engagement with a robot companion using task and social interaction-based feature*s, in Proc. of ICMI-MLMI 2009. 119-126, Cambridge, Massachusetts, USA, ACM.

Coninx A., Baxter P., Oleari E. *et al.* (2016), *Towards long-term social child-robot interaction: using multi-activity switching to engage young users*, Journal of

Human-Robot Interaction, 5(1), 32-67.

De Carolis B., Ferilli S. & Palestra G. (2017), *Simulating empathic behavior in a social assistive robot, Multimedia Tools Appl.*, 76(4), 5073-5094.

de Jong M., Zhang C.K., Roth A. M., Rhodes T., Zhou C., Ferreira S., Cartucho J. & Veloso M. (2018), *Towards a Robust Interactive and Learning Social Robot*, in Proc. of AAMAS 2018. 883-891, Vol. 2, Stockholm; Sweden.

de Greeff J. & Belpaeme T. (2015), *Why Robots Should Be Social: Enhancing Machine Learning through Social Human-Robot Interaction*, PLoS ONE, 10(9), 1-26.

de Greeff J., Delauna F. & Belpaeme T. (2012), *Active Learning between a Robot Learner and a Human Teacher*, in AAAI Fall Symposium on Robots Learning Interactively from Human Teachers. Arlington, USA, AAAI.

Di Bitonto, P., Roselli, T., Rossano, V., Frezza, E. & Piccinno, E. (2012), *An educational game to learn type 1 diabetes managemen*t, in Proc. of DMS, pp. 139-143.

Efthymiou N., Koutras P., Filntisis P. P., Potamianos G. & Maragos P. (2018), *Multi-View Fusion for Action Recognition in Child-Robot Interaction*, in Proc. of ICIP 2018. 455-459, IEEE.

Ekman P. (1999), Basic Emotions, in Dalgleish T., Power M. (eds), *Handbook of Cognition and Emotion*. Sussex, UK: John Wiley & Sons.

Ekman P., Friesen W. V. & Hager J. C. (2002), *Facial Action Coding System, Manual and Investigator's Guide*, Salt Lake City, UT, Research Nexus.

Fayek H., Lech M. & Cavedon L. (2017), *Evaluating deep learning architectures for Speech Emotion Recognition*, Neural Networks, 92, 60-62.

Fischer G. (2001), *User Modeling in Human–Computer Interaction*, User Modeling and User-Adapted Interaction, 11(1-2), 65-86.

Gudi S. L. K. C., Ojha S., Johnston B. & Williams M. A. (2019), *A Proactive Robot Tutor Based on Emotional Intelligence*, in Kim JH. *et al.* (eds), Robot Intelligence Technology and Applications, RiTA 2017, Advances in Intelligent Systems and Computing. 113-120, vol 751, Springer.

Haq S., Jackson P. J. & Edge J. (2008), *Audio-visual feature selection and reduction for emotion classification*, in Haq S., Jackson P.J.B., Edge J. (eds), in Proc. of AVSP 2008. 185-190, Moreton Island, Australia, AVISA.

Howley I., Kanda T., Hayashi K. & Rosé C. (2014), *Effects of social presence and social role on help-seeking and learning*, in Proc. of HRI 2014. 415-422, Bielefeld, Germany, ACM/IEEE.

Iocchi L., Lazaro M.T., Jeanpierre L. & Mouaddib A.I. (2015), *Personalized short-term multi-modal interaction for social robots assisting users in shopping malls*, in Tapus A., André E., Martin JC., Ferland F., Ammi M. (eds), Proc. of ICSR 2015. 264-274, LNCS, vol 9388, Springer

Ivaldi S., Anzalone S. M., Rousseau W., Sigaud O. & Chetouani M. (2014), *Robot initiative in a team learning task increases the rhythm of interaction but not the perceived engagement*, Frontiers in neurorobotics, 8(5), 1-16.

Janssen J.B., van der Wal C.C., Neerincx M.A. & Looije R. (2011), *Motivating children to learn arithmetic with an adaptive robot game*, in Mutlu B., Bartneck C., Ham J., Evers V., Kanda T. (eds), Proc. of ICSR 2011. 415-422, LNCS, vol 7072, Springer.

Kahou S. E., Michalski V., Konda K., Memisevic R. & Pal C. (2015), *Recurrent Neural Networks for Emotion Recognition in Video*, in Proc. of ICMI 2015. 467-474, New York, USA, ACM.

Kennedy J., Baxter P., Senft E. & Belpaeme T. (2015), *Higher nonverbal immediacy leads to greater learning gains in child-robot tutoring interactions*, in Tapus A., André E., Martin JC., Ferland F., Ammi M. (eds), Proc. of ICS-R 2015. 327-336, LNCS, vol 9388, Springer.

Ko B. C. (2018), *A Brief Review of Facial Emotion Recognition Based on Visual Information*, Sensors, Basel, Switzerland, 18(2), 401.

Kobsa, A. (2001), *Generic user modeling systems, User Modeling and User-Adapted Interaction*, 11 (1-2), 49-63.

Lafaye J., Gouaillier D. & Wieber P. B. (2014), *Linear model predictive control of the locomotion of Pepper, a humanoid robot with omnidirectional wheels*, in Proc. of International Conference on Humanoid Robots. 336-341, IEEE-RAS.

Larsson S. & Traum D.R. (2000), *Information state and dialogue management in the TRINDI dialogue move engine toolkit*, Nat. Lang. Eng. 6(3-4), 323-340.

Lee S., Noh H., Lee J. Lee K. & Lee G.G. (2010), *Cognitive Effects of Robot-Assisted Language Learning on Oral Skills*, in Proc. of INTERSPEECH 2010 Satellite Workshop on Second Language Studies: Acquisition, Learning, Education and Technology.

Lehmann H., Iacono I., Dautenhahn K., Marti P. & Robins B. (2014), *Robot companions for children with down syndrome A case study*, Interaction Studies, 15, 99-112.

Leyzberg D., Spaulding S., Toneva M. & Scassellati B. (2012), *The physical presence of a robot tutor increases cognitive learning gains*, in Proc. of the Annual Meeting of the Cognitive Science Society. Vol. 34(34).

Liu P., Glas D.F., Kanda T. *et al.* (2018), *Learning proactive behavior for interactive social robots, Autonomous Robots*, 42(5), 1067-1085.

Mavridis N. (2015), *A review of verbal and non-verbal human–robot interactive communication*, Robotics and Autonomous Systems, 63(1), 22-35.

McTear M.F. (1993), *User modelling for adaptive computer systems: a survey of recent developments*, Artificial Intelligence Review, 7(3-4), 157-184.

Mohammad Y., Matsumoto K. & Hoashi K. (2018), *Deep feature learning and selection for activity recognition*, in Proc. of SAC 2018. Pau, France, 930-939, ACM.

Mubin O., Shahid S. & Bartneck C. (2013), *Robot Assisted Language Learning through Games: A Comparison of Two Case Studies*, Australian Journal of Intelligent Information Processing Systems, 13(3), 9-14.

Palestra G., De Carolis B. & Esposito F. (2017), *Artificial Intelligence for Robot-Assisted Treatment of Autism*, in Proc. of Workshop on Artificial Intelligence with Application in Health co-located with AI*IA 2017. 17-24, Ceur vol. 1982.

Pennazio, V. (2017), *Social robotics to help children with autism in their interactions through Imitation*, REM - Research on Education and Media, 9(1), 10-16.

Pistoia M., Pinnelli S. & Borrelli G. (2015), *Use of a robotic platform in dyslexia-affected pupils: the ROBIN project experience*, International Journal of Education and Information Technologies, 9, 46-47.

Ramachandran A., Huang C. M., Gartland E. & Scassellati B. (2018), *Thinking Aloud with a Tutoring Robot to Enhance Learning*, in Proc of HRI 2018. 59-68, ACM/IEEE.

Ros R., Oleari E., Pozzi C. *et al.* (2016), *A motivational approach to support healthy habits in long-term child–robot interaction*, International Journal of Social Robotics, 8(5), 599-617.

Sebe N., Cohen I., Gevers T. & Huang T. S. (2006), *Emotion Recognition Based on Joint Visual and Audio Cues*, in Proc. of ICPR 2006, 1136-1139, IEEE.

Share P. and Pender J. (2018), *Preparing for a Robot Future? Social Professions, Social Robotics and the Challenges Ahead*, Irish Journal of Applied Social Studies, 18(1), 45-62.

Spexard T., Li, S., Wrede B. *et al.* (2006), *BIRON, where are you? Enabling a robot to learn new places in a real home environment by integrating spoken dialog and visual localization*, in Proc. of Int. Conf. on Intelligent Robots and Systems. 934-940, Beijing, China, IEEE/RSJ.

Schodde T, Bergmann K. & Kopp S. (2017), *Adaptive Robot Language Tutoring Based on Bayesian Knowledge Tracing and Predictive Decision-Making*, in Proc. of HRI 2017. 128-136, New York, USA, ACM/IEEE.

Schuller B.W. (2018), *Speech emotion recognition: two decades in a nutshell, benchmarks, and ongoing trend*s, Commun. ACM, 61(5), 90-99.

Tanaka F., Isshiki K., Takahashi F., Uekusa M., Sei R. & Hayashi, K. (2015), *Pepper learns together with children: Development of an educational application*, in Proc. of Humanoid Robots. 270-275, IEEE-RAS.

Tapus A. (2009), *Improving the quality of life of people with dementia through the use of socially assistive robots*, in Proc. of the Advanced Technologies for Enhanced Quality of Life. 81–86, Iasi, Romania, IEEE.

Wainer J., Dautenhahn K., Robins B. & Amirabdollahian F. (2014), *A pilot study with a novel setup for collaborative play of the humanoid robot KASPAR with children with autism*, International Journal of Social Robotics, 6(1), 45-65.

Walecki R., Rudovic O. (2017), *Deep structured learning for facial expression intensity estimation*, Image Vis. Comput., 259, 143–154.

Wong W., Cavedon L., Thangarajah J. & Padgham L. (2012), *Flexible conversation management using a BDI agent approach*, in Nakano Y., Neff M., Paiva A., Walker M. (eds.), Proc. of IVA 2012. 464-470, LNCS, vol. 7502, Springer.

Xiong W., Droppo J., Huang X. *et al.* (2017), *The Microsoft 2016 conversational speech recognition system*, in Proc. of ICASSP 2017, 5255-5259, IEEE.

Yamashita Y., Ishihara H., Ikeda T. *et al.* (2018), *Investigation of Causal Relationship Between Touch Sensations of Robots and Personality Impressions by Path Analysis*, International Journal of Social Robotics, 1-10, Springer.

Zeng Z., Tu J., Liu M., Huang T. S., Pianfetti B., Roth D. & Levinson S. (2007), *Audio-Visual Affect Recognition*, IEEE Transactions on Multimedia, 9(2), 424–428.

Zhang Y., Chan W. & Jaitly N. (2017), *Very deep convolutional networks for end-to-end speech recognition*, in Proc. of ICASSP 2017. 4845-4849, IEEE.