Linguistic Social Robot Control by Crowd-Computing Feedback

○ Vassilis KABURLASOS, Human-Machines Interaction Laboratory (HUMAIN-Lab), Department of Computer and Informatics Engineering, Eastern Macedonia and Thrace Institute of Technology (EMaTTech), Kavala, Greece, vgkabs@teiemt.gr
Christos BAZINAS, HUMAIN-Lab, Department of Computer and Informatics Engineering, EMaTTech, Kavala, Greece
George SIAVALAS, HUMAIN-Lab, Department of Computer and Informatics Engineering, EMaTTech, Kavala, Greece
George PAPAKOSTAS, HUMAIN-Lab, Department of Computer and Informatics Engineering, EMaTTech, Kavala, Greece

While social robots become ever more popular, novel modeling and control methodologies are sought toward optimizing their engagement. This work considers a NAO robot playing games with children. A novel feedback control is proposed using a resultant sentence, induced by crowd-computing techniques, as Reference to be compared with a computer-vision induced sentence toward driving a linguistic controller. Preliminary application results have been encouraging.

Key Words: Crowd Computing, Linguistic Feedback Control, Vision-Induced Text

1. Introduction

Cyber-Physical Systems (CPSs) have been introduced to account for technical devices with certain adaptive, sensing and reasoning abilities with a varying degree of autonomous behavior within networked environments (i.e. Internet-of-Things) – with or without a human in the information and control loop. CPSs have been considered in different application domains including Disabled People, Healthcare, Agriculture & Food Supply, Manufacturing, Energy & Critical Infrastructures, Transport & Logistics and Community Security & Safety and, lately, Education & Pedagogical Rehabilitation [1]. The CRSs of interest in this work are Social Robots in children education applications. Our interest here is not in "learning about robotics", but rather it is in "robotics assisted learning" [2] as explained below.

Traditional education is delivered by a human teacher in a classroom. Typically, no advice is sought from an outside expert during a class session. Maintaining a skilled body of human teachers clearly has benefits for a nation's young generation, nevertheless the corresponding cost might be unaffordable. New technology solutions could assist in delivering high skilled education at an affordable cost as explained in the following.

Social Robots are getting established in various types of education as instruments to a teacher [3]. The next natural step, regarding the engagement of a Social Robot in education, would be to engage a Social Robot as an assistant to a teacher. To fulfill its aforementioned potential a Social Robot should be intelligent enough to interact with a human. However, an intelligent Social Robot might be short of the expertise required to handle specific educational situations. We remark that even a human teacher in the classroom might not be able to handle specific educational situations, also due to a shortage of expertise. Therefore, expert advice should be made available.

This work proposes a scheme for delivering expert advice in the classroom to a Social Robot on-demand via crowd-computing techniques. Recall that crowd-computing has been defined as "harnessing the power of people out in the web to do tasks that are hard for individual users or computers to do alone. Like cloud computing, crowd computing offers elastic, on-demand human resources that can drive new applications and new ways of thinking about technology." [4].

This paper is organized as follows. Section 2 outlines the AB-scheme. Section 3 describes the implementation of the AB-scheme. Section 4 presents preliminary application results. Finally, section 5 concludes by summarizing this contribution and discussing future work.

2. The Proposed AB-Scheme

Fig.1 displays a child playing an educational game with a Social Robot, namely NAO. Fig.2 presents, the AB-scheme which implements feedback control towards enhancing child-robot interaction as explained in the following.

A Rule Base in Fig.2 includes deterministic knowledge elicited from experts and represented by (fuzzy) rules. The Rule Base calculates a Reference (antecedent proposition) p_R as an input to a Comparator. For instance, the proposition p_R could be "A boy 5 years old is playing with a board-game, and is smiling". In other words, the Reference defines the antecedent of a Rule to make a Child happy.

NAO sensor signals, including video, are processed by Computational Intelligence as well as Machine Learning algorithms toward inducing proposition p_S. Finally, propositions p_R and p_S are compared toward driving a Linguistic Controller (embedded in the NAO robot) toward bringing the Child to the desired Reference. The induction of proposition p_S from sensory data was necessary toward seeking the verbal advice of a crowd of experts via a Social (Internet-based) Network.

On one hand, the blocks "NAO", "Child", "NAO sensors" and "Signal to Proposition Conversion" in Fig.2 were implemented in a classroom, namely *Action-station* or, alternatively, *station-A*. On the other hand, all the other blocks in Fig.2 were implemented remotely in the *Base-station* or, alternatively, *station-B*. Due to the collaborative engagement of station-A and station-B the proposed scheme was named *AB-scheme*.

3. AB-Scheme Implementation Details

The AB-scheme is based on a scenario where a child is playing a game to be recognized as well as the child's age /expression /gender need to be recognized. More specifically, the child's age is one of 14 discrete age groups, its expression is one of 5 different expressions, whereas 3 different games have been considered. Therefore, 14*2*5*3 = 420 combinations are possible.

A sentence was induced by pattern recognition techniques. In particular, for gender classification and expression recognition, a Face Recognizer was trained; for age estimation, a Convolutional Neural Network (CNN) was used [5]; for game recognition and classification, a Cascade Classifier was used [6]. Training data information is summarized in Table 1. For instance, a sentence could be "A boy 5 years old is playing with a board-game, and is smiling."

To handle these distinct propositions, Crowd Computing and Machine Learning were employed synergistically as explained in the following. On one hand, if the sentence induced by computer vision appears for the first time, it is sent to the Crowd. The Crowd within 5 minutes suggests an action which is implemented by NAO (if there are no suggestions from Crowd, a random one is chosen instead). For instance, examples of actions include: "Give the toy to an older child", "Give the toy to a child of the opposite gender", "Change Toy", etc. This process continues until the child's facial expression becomes "Happy". Then the corresponding rule is stored in the Rule Base for future use. On the other hand, if the sentence has appeared previously, then the corresponding rule is used.

4. Preliminary Application Results

Our experimental results are showing that the proposed method can be more and more effective, with less and less participation from the Crowd. As the Rule base gets populated with successful pairs of computer vision induced Propositions and Crowd References, the emotional status - expression of the participant child tends to stabilize to the condition "Happy". In time, Crowd participation will not be needed at all, as the Rule base will be filled with all unique combinations of age, gender, toy, expression and the proper Reference for each one of them.

In Fig 3, the results of 2,000 repetitions are shown. At the first observations, the Rule base is empty. This means that the Crowd must be involved in almost all times the process is run, because the system (NAO) hasn't yet learned the correct References for each Proposition. When the Crowd gets involved, some new Rules arise, and they populate the Rule base. So, here, a connection between the Crowd participation number and the new Rules number is obvious. The descending percentage of Crowd participations, and new Rules as a consequence of having more and more Rules in the Rule base is also obvious.

The most important results that were noticed, was success rate (the child ends up happy). In all the observations there is a possibility where a happy child plays with a toy, so no further participation from Crowd or NAO is needed. While we repeat the experiment, the possibility of successful result is increasing. This happens because when Crowd or a Rule is used, the child is helped so it reaches the status "Happy" (more possibilities of a successful process). And since the Crowd proposes something on a totally unknown situation, it has great chances of proposing something wrong. On the other hand, NAO's References have been used before and have been successful, so the possibilities of being successful again are quite bigger.

As continuing monitoring the process, a strong connection between the usage of previously saved Rules and final success is noticed. The possibility of having a happy child anyway, is added to the possibility of using a Rule with successful result and the possibility of Crowd participation that results "Happy Child". Hence, after 2000 times, the percentage of success reaches 88%.

5. Conclusion and Future Work

This work has presented preliminary results regarding a Social Robot that assists learning using crowd-computing feedback techniques. We have developed the software/hardware infrastructure necessary for elaborate experiments in the future. Preliminary application results, demonstrated here, have been encouraging. Our proposed techniques may enable the engagement of lower skilled personnel for delivering (much) higher skilled instruction in University Teaching/Instruction and beyond.

ACKNOWLEDGEMENT

This work has been supported, in part, by the European Commission Horizon 2020 MSCA-RISE Project no. 777720 "Cyber-Physical Systems for PEdagogical Rehabilitation in Special EDucation (CybSPEED)".

REFERENCES

[1] CybSPEED, "Cyber-physical systems for pedagogical rehabilitation in special education," European Commission Horizon 2020 MSCA-RISE Project no. 777720, 1 Dec. 2017 - 30 Nov. 2021.

[2] Salvini, P., Korsah, A., and Nourbakhsh, I., "Educational Robotics – Fostering Creativity and Problem-Solving Abilities," IEEE Robotics & Automation, vol. 23, no. 2, 2016.

[3] Conti, D., Di Nuovo, S., Buono, S., and Di Nuovo, A., "Robots in education and care of children with developmental disabilities: a study on acceptance by experienced and future professionals," International Journal of Social Robotics, vol. 9, iss. 1, pp. 51-62, 2017.

[4] Miller, R., Microsoft research talk, June 19, 2013; http://research.microsoft.com/apps/video/default.aspx?id=194501

[5] Levi, G. and Hassner, T., "Age and Gender Classification using Convolutional Neural Networks," IEEE Workshop on Analysis and Modeling of Faces and Gestures (AMFG), at the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Boston, June 2015.

[6] Viola, P. and Jones, M. J., "Robust real-time face detection," International Journal of Computer Vision, vol. 57, no. 2, pp. 137-154, 2004.

Fig.1 A child playing a game with robot NAO.

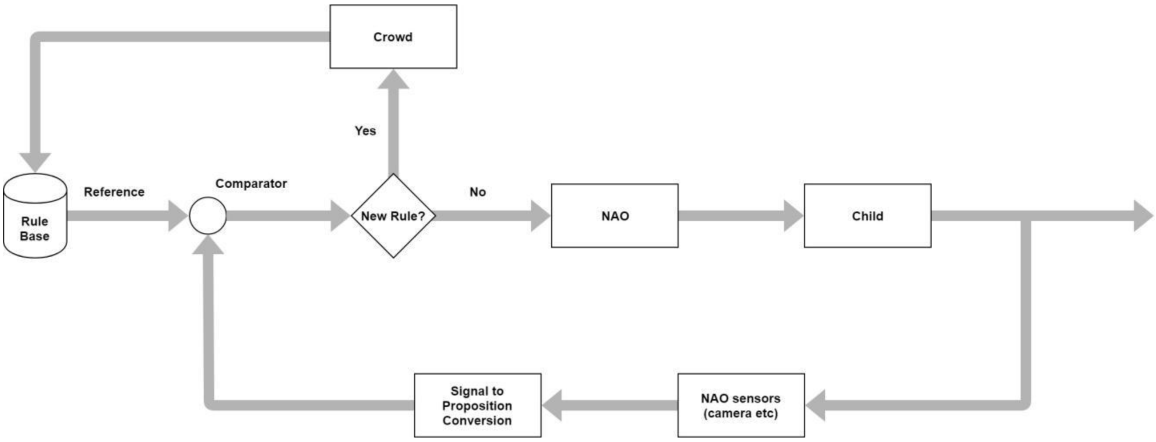

Fig.2 Block diagram of the AB-scheme toward delivering skilled education supported by crowd-computing techniques.

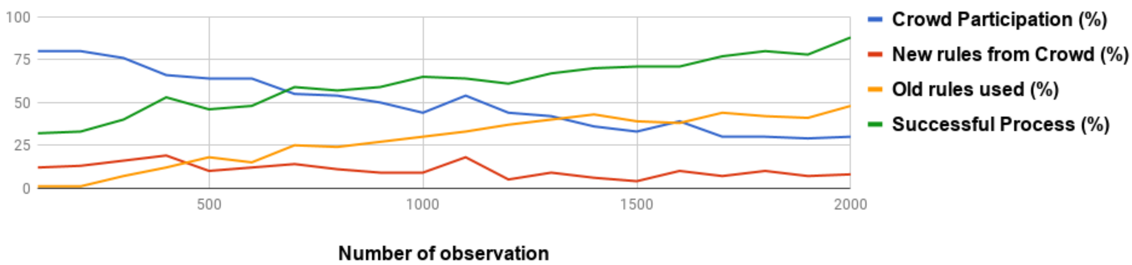

Fig.3 Crowd participation, success percentages, new and old rules usage.

Table 1: Training data information.

Feature	Dataset (number of images)	Image dimensions (px)	Training time (hours)
Age	0	0	0
Expression	400	various	4
Gender	>41,000	100×100	72
Game	>45,000	50×50	1260