IEEE CIS
AMD Newsletter
Robotics
Psychology
Neuroscience
Machine Intelligence
The Newsletter of the Autonomous Mental Development Technical Committee

AUTHOR INFORMATION

Adriana Tapus
Robotics and Computer Vision Lab
ENSTA-ParisTech, Paris, France

DIALOG COLUMN

Can We Trust Closed-Skull Socially Assistive Robots?

Robots are more and more present in our daily life; they have to move into human-centered environments, to interact with humans, to obey some social rules so as to produce an appropriate social behavior in accordance with human's profile (i.e., personality, state of mood, and preferences), and to provide assistance to vulnerable users. More and more research works are trying to step up from short-term and task specific approaches towards the development of emerging solutions for life-long and adaptive interactions [1], [2], [3].

First of all, one of the requests of Prof. Weng was the definition of "open skull" and "closed skull". The idea behind "closed skull" is how to continuously and robustly learn from almost nothing a variety of unconstrained tasks and reliably adapt to unseen contexts similarly to the way humans do it. With this in mind, I understand "closed skull" as an autonomous life-long learning and decision process with no human intervention (i.e., continuously learn and decide what to learn, when to learn, and how to behave appropriately). On the other hand, I see an "open skull" system as a system that can autonomously act, learn, and interact with human peers for a variety of specific tasks. However, if the context for which the system was tailored is highly dynamic, it is allowed for the engineer to freely "open the skull at night" to study and improve the system.

As previously stated in my response [4] to the point raised by Rohlfing and Wrede in [5], I agree that the robot has to be capable of life-long incremental learning in order to better adjust its behaviors and provide an appropriate respond as a function of various situations and tasks. Humans and environmental factors can be unpredictible and of course the brittle effect can appear. A system cannot be built and handcrafted to be perfect and this is an aspect that I totally share with Prof. Weng. Nevertheless, I believe that a system that can evolve with time and incrementally learn new inputs can work perfectly. In his research works [6], Brooks advocates that a complex and sophisticated intelligent robotic system can be developed by the incremental addition of individual layers of situation-specific control systems. Or maybe as suggested in [7], develop a solution for perturbation-tolerant autonomous system whereby the robot can infer whether or not it is achieving its goals, and if not, trying a potential "Plan B", or more or less random variations in behaviour, and ask for help, or use trial-and-error.

Assistive robots have to provide assistance to vulnerable users in a context-specific scenario, and need to respond to both short-term changes that represent individual differences and long-term changes that allow the interaction to continue to be engaging over a period of months and even years. By using a "closed-skull" development, the system must know robustly when, who, and how to assist the individuals [2]. I believe that machines' perception still lacks reliability and robustness in complex natural environments and that's a bottleneck for performing tasks that humans do effortlessly. So, at the question raised by Weng: "Is skull-closed development inappropriate for socially assistive robots that serve vulnerable users?", my concern is about how are we able to measure what the skull-closed robot has learned and its evolution in time and how can we guarantee an appropriate safe robot behavior that will serve the users' needs. One fundamental concern is wheather such robots can be trusted (trustworthiness quantified in terms of usefulness, safety, and predictability).

The tasks and the trainings required by the therapists/nurses are very different from one medical condition to another and from one user to another, making the interaction even harder and challenging. While the validation domains and thus some user populations (e.g., the elderly, post-stroke patients, children with autism) are quite diverse, the principal facet of the human-robot interaction needed for Socially Assistive Robotics (SAR) care are quite similar. The system complexity has to be variable within a wide range of individual needs. The interactions require adaptation over time so as to enable training behavior (cognitive or physical) toward long-term performance goals. With this view in mind, I believe "open-skull" interactions provide a sufficient and yet more successful route toward robust and safer assistive robots for vulnerable users.

References

[1] P. Baxter, T. Belpaeme, L. Canamero, P. Cosi, Y. Demiris, and V. Enescu, "Long-term human-robot interaction with young users," in *Proceedings of the ACM/IEEE Human-Robot Interaction conference (HRI-2011) (Robots with Children Workshop)*, Lausanne, Switzerland, 2011.

[2] Y. Demiris, "Knowing when to assist: Developmental issues in lifelong assistive robotics," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Minneapolis, MN, 2009, pp. 3357–3360.

[3] A. Tapus, C. Tapus, and M. J. Matarić, "Long term learning and online robot behavior adaptation for individuals with physical and cognitive impairments," in *In Field and Service Robotics, Results of the 7th International Conference*, A. Kelly, K. Iagnemma, and A. Howard, Eds. Series: Springer Tracts in Advanced Robotics, 2010, vol. 62, pp. 389–401.

[4] A. Tapus, "Is developmental robotics a solution for socially assistive robotics?" *Dialog column of IEEE CIS AMD Newsletter*, vol. 9, no. 1, pp. 9–10, 2012.

[5] K. Rohlfing and B. Wrede, "What novel scientific and technological questions does developmental robotics bring to hri? are we ready for a loop?" *Dialog column of IEEE CIS AMD Newsletter*, vol. 9, no. 1, pp. 3–5, 2012.

[6] R. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal of Robotics and Automation*, vol. 2, no. 1, pp. 14–23, 1986.

[7] M. L. Anderson and D. R. Perlis, "Logic, self-awareness and self improvement: The metacognitive loop and the problem of brittleness," *Journal of Logic and Computation*, vol. 14, 2004.