Towards Anchoring Self-Learned Representations to Those of Other Agents

Imperial College London

Martina Zambelli, Tobias Fischer, Maxime Petit, Hyung Jin Chang, Antoine Cully and Yiannis Demiris

{m.zambelli13, t.fischer, m.petit, hj.chang, a.cully, y.demiris}@imperial.ac.uk imperial.ac.uk/**PersonalRobotics**

Motivations & Objectives

- Robots need to be able to learn how to deal with new situations autonomously [1] without requiring the intervention of an engineer.
- Robots need to understand the point of views and the different abilities of other people for better interaction. Perspective taking abilities reduce ambiguities in interactions [2].

Introduction

- We present a robotics architecture for **anchoring representations** that are **autonomously learned** by the robot into the **perspective of other agents** (see Fig. 1) based on five components:
  1. **Multimodal Sensorimotor Representations** [3]: self-learn sensorimotor representations from visual, proprioceptive and tactile stimuli. The acquired knowledge is employed by components 2 and 3.
  2. **Symbolic Representations** [4]: anchor on linguistic level (*e.g.* joint number to body part name).
  3. **Kinematic Structure Correspondences** [5]: anchor on sensorimotor level (*e.g.* body part correspondences between two agents).
  4. **Perspective Taking** [6]: use spatial reasoning algorithms learned from robot's perspective to also reason from human's perspective.
  5. **Long-term Autobiographical Memory** [7]: common interface to exchange streaming data from robot, internal representations, augmented versions.
- Follow developmental principle shown by Baraglia *et al.* [8]: **action production alters action perception**.
- Towards an implementation of the **"others like me"** paradigm [9], based on self-exploration and self-others mapping.

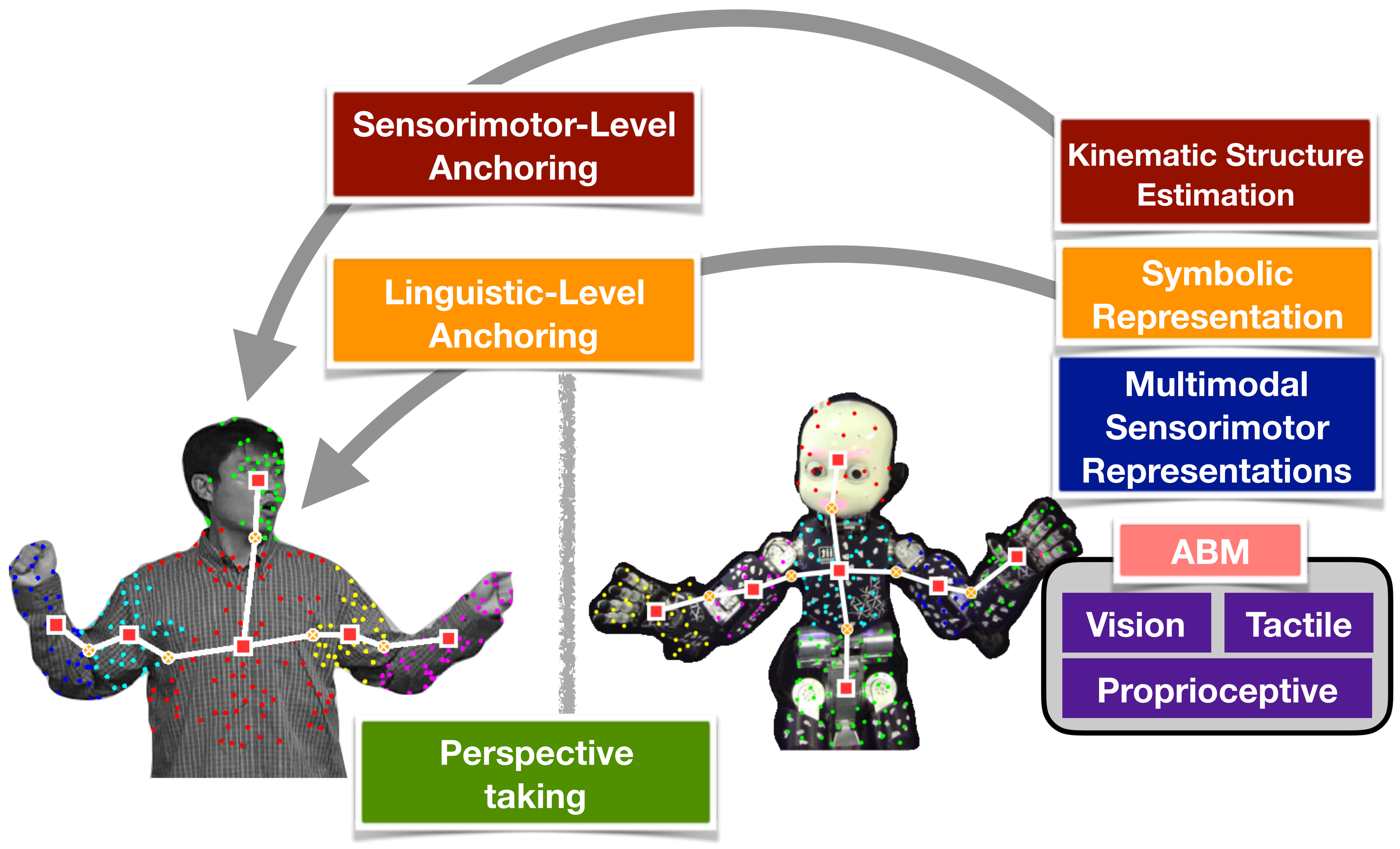

Figure 1: Overview of the proposed architecture. The robot learns **multimodal sensorimotor representations**, and employs this knowledge to **anchor self-knowledge to that of other agents** on the linguistic and sensorimotor levels. This can then be used by the **perspective taking** component to reason about the other agents' perspective. The common interface is provided by an **autobiographical memory** (ABM).

Material and Methods

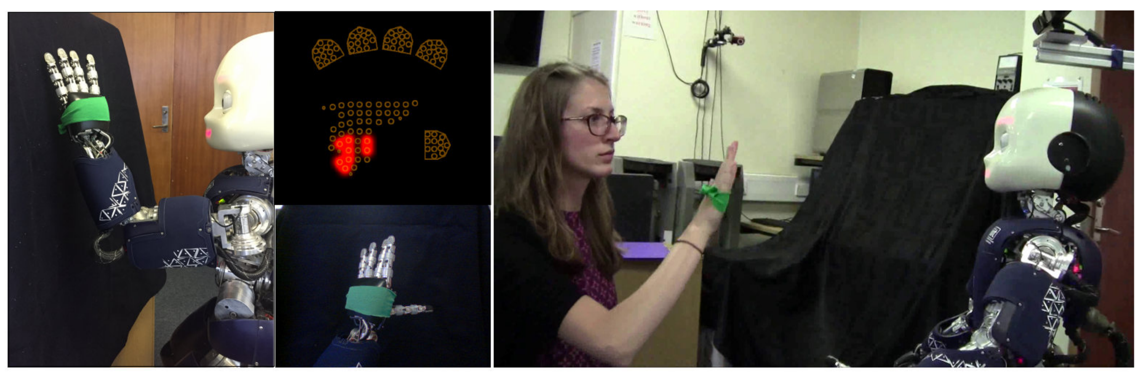

Figure 2: **Multimodal Sensorimotor Repr.** (Left) iCub faces a surface and performs movements in order to touch it. (Centre top) The palm tactile sensor is activated during touch events. (Centre bottom) View from the robot's left eye camera. (Right) iCub learns to imitate a circular trajectory demonstrated by a human.

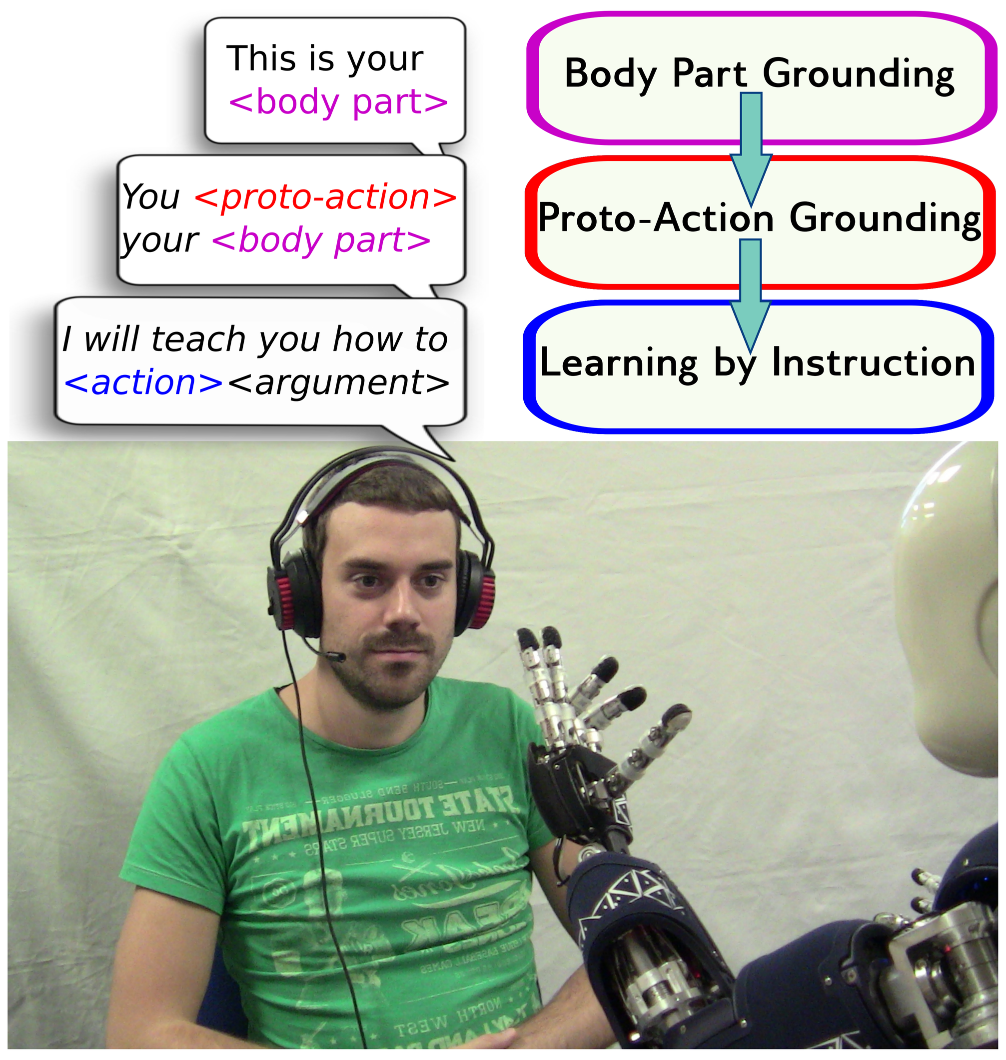

Figure 3: **Symbolic Repr.** Hierarchical components to 1) learn body part names from human labelling after a robot motor babbling, 2) discover proto-actions (*i.e.* a single joint's position command) with on-the-fly descriptions by the human, and 3) learning by instructions with the human as teacher to scaffold the acquired skills into primitives or complex actions.

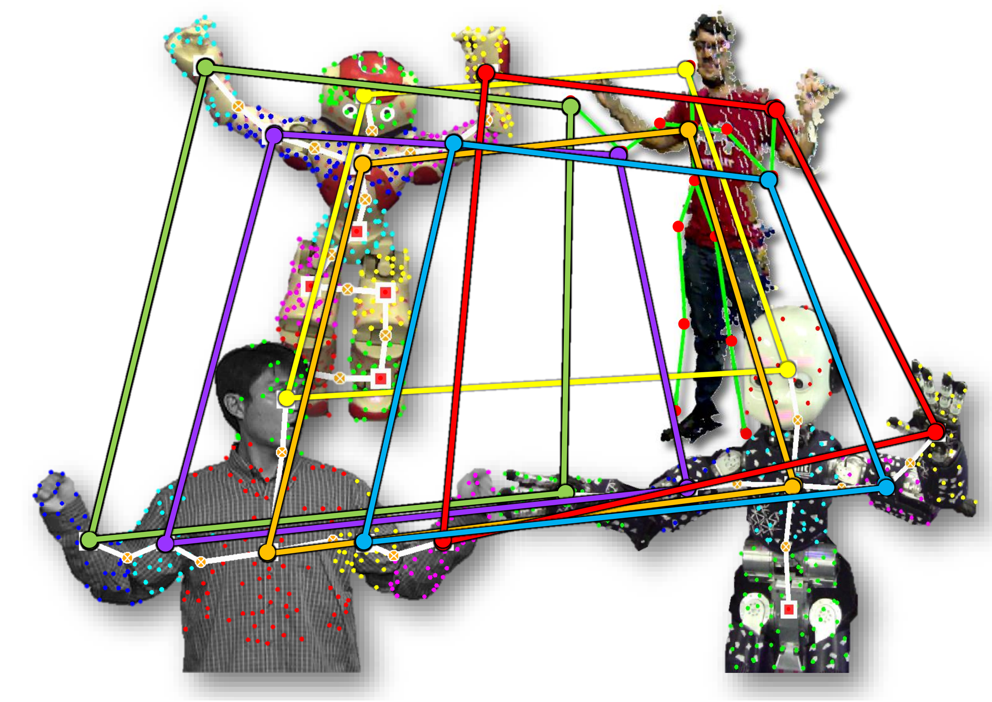

Figure 4: **Kinematic Structure Corresp.** Various kinematic structure correspondence matching results. The iCub humanoid robot (bottom right) can find correspondences to human partners, either sensed through the iCub eyes (bottom left) or a RGB-D camera (top right). Also, correspondences to other humanoid robots like the NAO robot (top left) can be found.

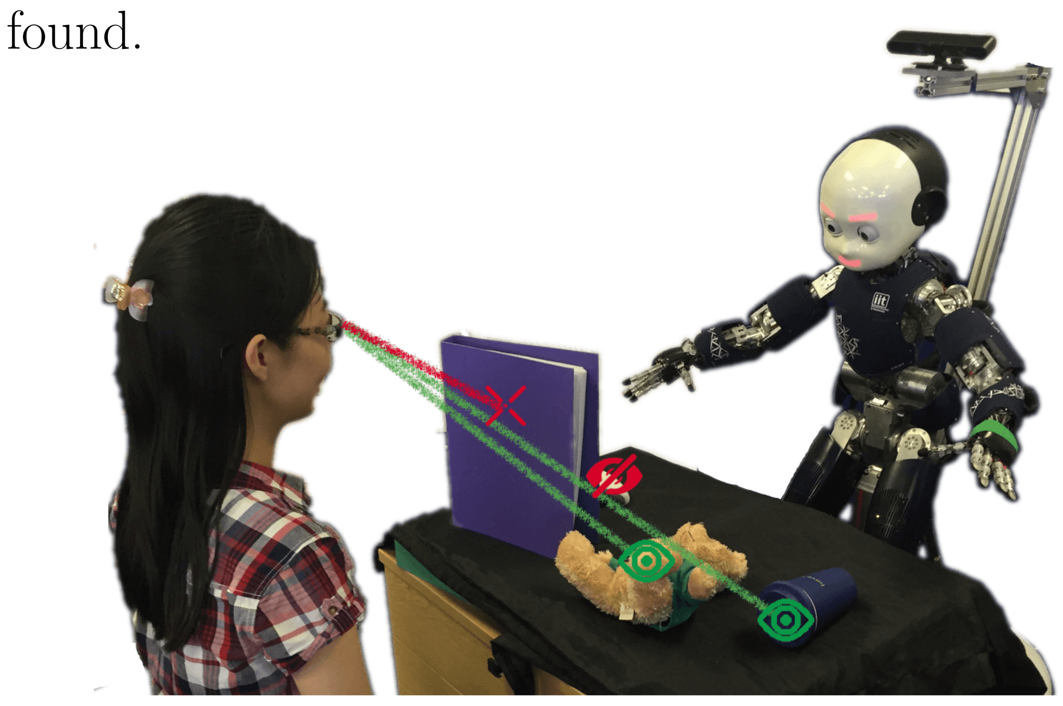

Figure 5: **Perspective Taking.** Typical set-up in a perspective taking scenario. The world perceived by the human and that of the robot differ in various aspects. For example, in this figure, one object is occluded to the human but visible by the robot. Furthermore, the blue cup is to the left of the robot, but to the right of the human.

Conclusion

With this framework, the robot is able to **predict the consequences of movements** performed by other agents, based on the **learned consequences** of its own movements, and **assist humans** effectively in cooperative tasks. The key property of this system is that it does **not use any prior knowledge** (*e.g.* about the body schema of the human).

[1] A. Cully, J. Clune, D. Tarapore, and J.-B. Mouret, "Robots that can adapt like animals," *Nature*, vol. 521, no. 7553, pp. 503–507, 2015.
[2] C. Breazeal, M. Berlin, A. Brooks, J. Gray, and A. L. Thomaz, "Using perspective taking to learn from ambiguous demonstrations," *Robotics and Autonomous Systems*, vol. 54, no. 5, pp. 385–393, 2006.
[3] M. Zambelli and Y. Demiris, "Multimodal Imitation Using Self-Learned Sensorimotor Representations," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2016.
[4] M. Petit and Y. Demiris, "Hierarchical Action Learning by Instruction Through Interactive Body Part and Proto-Action Grounding," in *IEEE International Conference on Robotics and Automation*, 2016, pp. 3375–3382.
[5] H. J. Chang, T. Fischer, M. Petit, M. Zambelli, and Y. Demiris, "Kinematic Structure Correspondences via Hypergraph Matching," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 4216–4425.
[6] T. Fischer and Y. Demiris, "Markerless Perspective Taking for Humanoid Robots in Unconstrained Environments," in *IEEE International Conference on Robotics and Automation*, 2016, pp. 3309–3316.
[7] M. Petit, T. Fischer, and Y. Demiris, "Lifelong Augmentation of Multi-Modal Streaming Autobiographical Memories," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 8, no. 3, pp. 201–213, 2016.
[8] J. Baraglia, J. L. Copete, Y. Nagai, and M. Asada, "Motor Experience Alters Action Perception Through Predictive Learning of Sensorimotor Information," in *Int. Conf. Developmental Learning and Epigenetic Robotics*, 2015, pp. 63–69.
[9] A. N. Meltzoff, "'Like me': a foundation for social cognition," *Developmental Science*, vol. 10, no. 1, pp. 126–134, 2007.