

Guest Editorial

Continual Unsupervised Sensorimotor Learning

I. SCOPE OF THIS SPECIAL ISSUE

The pursuit of higher levels of autonomy and versatility in robotics is arguably led by two main factors. Firstly, as we push robots out of the labs and production lines, it becomes increasingly challenging to design for all possible scenarios that a particular robot might encounter. Secondly, the cost of designing, manufacturing, and maintaining such systems becomes prohibitive.

Although machine learning algorithms continue to improve at a rapid pace enabling technologies and products such as autonomous driving cars and sophisticated image and speech recognition, it is often forgotten that these applications represent tailored solutions to specific tasks. Thus it is not clear if or how these autonomous systems can pave the road to general-purpose machines envisioned by many.

As the algorithms for learning single tasks in constrained environments are improving, new challenges have gained relevance to achieve more autonomous artificial systems. These challenges include multi-task learning, multimodal sensorimotor learning and lifelong adaptation to wear and tear. Addressing these challenges promise higher levels of autonomy and versatility of future robots.

This special issue on Continual Unsupervised Sensorimotor Learning is primarily concerned with the developmental processes involved in unsupervised sensorimotor learning in a life-long perspective [7], and in particular the emergence of representations of action and perception in humans and artificial agents in a continual learning manner [6]. These processes are done via action-perception cycle, active perception, continual sensorimotor learning, environmental-driven scaffolding, and intrinsic motivation [5]. The special issue highlights behavioural and neural data, and cognitive and developmental approaches to research in the areas of robotics, computer science, psychology, neuroscience, etc. Contributions focus on mathematical and computational models to improve robot performance and/or attempt to unveil the underlying mechanisms that lead to continual adaptation to changing environment or embodiment and continual learning in open-ended environments.

Moreover, it contains contributions from multiple disciplines including cognitive systems, cognitive robotics, developmental and epigenetic robotics, autonomous and evolutionary robotics, social structures, multi-agent and artificial life systems, computational neuroscience, and developmental psychology, on theoretical, computational, application-oriented, and experimental studies.

II. THEMES

This special issue reports state-of-the-art approaches and recent advances on Continual Unsupervised Sensorimotor

Learning with a cross-disciplinary perspective. Topics included in this special issue are:

- Intrinsic motivation
- Hierarchical policy learning
- Continual sensorimotor learning
- Motor primitives
- Goal generation
- Agent Architectures
- Curriculum learning
- Learning from demonstration
- Self-organizing behaviour
- Skill acquisition and planning
- Environmental-driven scaffolding
- Emergence of representations via continual interaction
- Deep reinforcement learning

III. CONTRIBUTIONS TO THE SPECIAL ISSUE

This special issue includes nine papers that describe work inspired by developmental processes for unsupervised sensorimotor learning, casting light on this topic from multiple angles. We briefly highlight the main contributions and the field of application of each paper as follows.

Fournier et al. presents the most theoretical article of this special issue in the paper entitled *CLIC: Curriculum Learning and Imitation for Object Control in Non-Rewarding Environments*. This article combines curriculum and imitation learning within the RL formalism of an interactive intrinsically motivated learner in a new reinforcement learning setting with a non-rewarding environment. The algorithm CLIC is a variant of the DQNfD for the case of multi-task learning and augmented by intrinsic motivation. CLIC is tested in various grid non-rewarding multi-task worlds and tutors. The learning agent can take advantage without being limited by the tutor. Moreover, its intrinsic motivation can choose what to explore and what to imitate. It can ignore both non-reproducible and mastered interactions with objects. This approach is similar to research on interactive intrinsic motivation and hierarchical learning [e.g. 1, 4], but proposes a deep reinforcement learning algorithm.

Applying deep reinforcement learning and intrinsic motivation in a real-world scenario, the paper entitled *BND*-DDQN: Learn to Steer Autonomously Through Deep Reinforcement Learning* features an end-to-end network to achieve safe autonomous steering in changing environments. This novel deep reinforcement learning model uses spatio-temporal features from a depth map and the difference between consecutive images as independent input streams to decide angular and linear actions taking into consideration smooth angular velocity. Moreover, Wu et al. also use random network distillation (RND) exploration bonuses as an intrinsic reward to encourage

exploration of dissimilar states. Further to prevent oscillatory left-right steering, the angular velocity is selected based on the estimated Q-values but also the command executed in the previous step. Finally, the use of depth map information allows for generalization as well as excellent transferability from simulated environments to cluttered real-world environments containing both static and dynamic obstacles.

The next article by Abdelfattah et al., integrate elements similar to the first two articles described so far, i.e. it shows an algorithm applied to low-level data like what is presented by Wu et al., and it proposes a hierarchical RL method based on intrinsic motivation like done by Fournier et al. Specifically, Abdelfattah et al. in *Intrinsically Motivated Hierarchical Policy Learning in Multi-Objective Markov Decision Processes* propose a novel intrinsically motivated multi-objective reinforcement learning method that can learn hierarchical policy coverage sets to better generalize to different shifts in the environment's dynamics: learn skill sets in a developmental manner using intrinsic motivation RL, and a novel multi-objective RL method for evolving policy coverage sets using hierarchical policy learning, generalize to non-stationary dynamics in the state transition distribution and reward distributions. The work is in line with developmental and epigenetic robotics research and introduces a novel multi-objective reinforcement learning framework to learn skill sets. The robot uses low-level data from camera and proximity sensors to explore the environment with intrinsic motivation. It can adapt to changing environments with the evolution of the policy coverage set for each shift in the environmental dynamics.

The next article presented in this special issue also describes an approach of learning based on raw sensory data. The article entitled *Open-Ended Continuous Learning of Compound Goals* focuses on hierarchical policy learning and goal generation. Dhakan et al. present a goal-generation mechanism that can autonomously build from raw, sensory data, target goals to be learned by an artificial agent. They propose an algorithm for continuously learning new skills in an open-ended manner, in a discrete environment, using automatic goal generation mechanisms to learn maintaining tasks of different levels of complexity, after reflecting on the state variables structure. They have tested their approach on different variations of the 2D-navigation problem of a mobile robot. This approach comprises different steps, that can be repeated indefinitely, to tackle open-ended learning: an exploration phase where a reinforcement learning agent is set to explore its environment, a state aggregation process which uses hierarchical clustering to create groups of related states for reducing their cardinality for future steps, a compound goal generation process which combines elementary tasks related to the aggregated states group, the final step uses reinforcement learning to learn the skills to maintain all the built compound goals.

Ezenkwu et al. take on this idea of self-organisation of the sensorimotor map in the article entitled *Unsupervised Temporospatial Neural Architecture for Sensorimotor Map Learning*, published in Volume 13, Issue: 1. Ezenkwu et al. present a Temporospatial Merge Grow When Required (TMGWR) network for continuous self-organisation of an agent's sensorimotor maps in noisy environments. The key features of

the proposed TMGWR Networks are a) the nodes are linked based on their sensorimotor proximity, i.e. edges between nodes are created if there is an action that causes the agent's sensorimotor experience to change from one node to the other. Such encoding allows for causal reasoning and planning like shown by Manoury et al. [3] where the sensorimotor experience builds simple tasks that can be used for planning or as subtasks for compound tasks. b) it employs recursive temporal context to keep track of the sensorimotor history, which allows the model to be used on non-Markovian problems. Moreover, Ezenkwu et al. also introduce a new metric to quantify the representation of causality in sensorimotor maps called "sensorimotor-link error" (SE). SE is the ratio of the number of impossible transitions to the total number of transitions encoded in the map. The article presents evidence of the efficiency and suitability of TMGWR for noisy and partially observable environments in comparison to other state-of-the-art algorithms.

Like most of the articles in this special issue, *Effect Regulated Projection of Robot's Action Space for Production and Prediction of Manipulation Primitives Through Learning Progress and Predictability Based Exploration* focuses on active exploration and progress-based intrinsic motivation, but this time applying to a robot arm. Bugur et al. propose an intrinsic motivation based parameter exploration mechanism that enables the formation of motor primitives based on action specialization, and allows lightweight predictive models to be formed for individual actions. The concept is realized in the manipulation domain, furthering the current state of the art in affordance learning [2]. Interesting parallels between the development of reach and grasp execution and prediction in infants are found. The key ingredient of the proposed model is the interplay between the learning progress of parameter regions and the splitting of the regions to maximize prediction accuracy. Thus, the proposed system not only tries to discover and learn in the regions where learning progress is high, but it implements load balancing over the regions in the sense that it splits regions that receive excessive amounts of learning requests. Overall, the work contributes to continual learning unsupervised learning by serving as an example to the fact that intrinsic motivation should be balanced with other mechanisms such as those that control modularity and computational resources.

Likewise, the article entitled *Autonomous Identification and Goal-Directed Invocation of Event-Predictive Behavioral Primitives* uses intrinsic motivation, but uses a surprise-based criterium. Assuming that the novel and complex behaviours of humans stem from motor primitives or elementary building blocks, Gumbesch et al. propose a learning architecture called "surprise-based behavioural modularization". In this learning architecture, the association of the primitives is obtained by the self-organizing dynamics, which was called differential extrinsic plasticity. It was also shown that the learned behaviours speed up the learning of different goal-reaching tasks. As a result, various complex and highly coordinated behaviours are shown in robotic simulations. Even though the robot does not have any prior knowledge, the authors show that complex coordinated behaviours can be assembled from the motor primitives.

More theoretical, the article entitled *Where Do I Move My Sensors? Emergence of a Topological Representation of Sensors*

Poses from the Sensorimotor Flow, presents a mathematical formalization for the construction of sensorimotor representations obtained via iterative interaction with the environment. Marcel et al. show that a naive agent, having access only to sensory input and actuator states, can build an accurate internal representation of the agent dynamics in the physical world. The obtained representation is topologically organized; if the statistics of sensory invariants are continuous. Further, such topological internal representation could be used, for instance, for path planning or obstacle avoidance. Marcel et al. note that such topological representation would be readily usable when the dynamics of the agent are faster than the dynamics of the environment. Although internal representations for environmental dynamics faster than agent dynamics could also be usable in real-world applications, these representations would be distorted and would require additional systems such as reinforcement learning to deal with repeatability of agent's movements.

Finally, in the review paper on vocal control learning entitled *Vocal Imitation in Sensorimotor Learning Models: A Comparative Review*, Pagliarini et al. propose a taxonomy of sensorimotor learning models on vocal imitation learning with a focus on song learning in birds and speech acquisition in humans. To do so, the comparative neuroanatomy of the two species are used to lay down the behavioural and physiological features of vocal imitation. Then, the components of such learning systems are charted out, and the reviewed models are analysed regarding the components identified. In particular, the authors consider the motor, sensory and perceptual space as the three main foundations that a computational model builds upon and which the learning framework operates on. The learning framework is further viewed as containing the architecture, the learning algorithm, the evaluation and exploration strategy. The brain mechanisms of sensorimotor learning are reviewed in detail, and the models are compared with respect to the key brain functions/areas. Although the vocal imitation is chosen as the target for the review, the contribution of the paper extends well beyond vocal learning and imitation and can be used as a general reference for sensorimotor learning and its plausible biological modelling.

Nicolás NAVARRO-GUERRERO, *Co-Lead Guest Editor*
Aarhus University, Aarhus, Denmark
E-mail: nicolas.navarro.guerrero@gmail.com
Homepage: <https://nicolas-navarro-guerrero.github.io>

Sao Mai NGUYEN, *Co-Lead Guest Editor*
IMT Atlantique, Ensta-Paris & Inria, France
E-mail: nguyensmai@gmail.com
Homepage: <https://nguyensmai.free.fr>

Erhan OZTOP, *Guest Editor*
Osaka University, Japan / Ozyegin University, Turkey
E-mail: erhan.oztop@{otri.osaka-u.ac.jp, ozyegin.edu.tr}
Homepage: <http://robotics.ozyegin.edu.tr/members/erhan-oztop>

Junpei ZHONG, *Guest Editor*
The Hong Kong Polytechnic University, Hong Kong
E-mail: joni.zhong@ieee.org
Homepage: <http://junpei.eu>

IV. CONCLUSION

The articles presented in this special issue represent some of the current challenges in continual unsupervised sensorimotor learning. How can heuristics such as intrinsic motivation guide the learning in multi-task settings? How can hierarchical learning take advantage of emergence of representations and goal generation? These articles show state-of-the-art results for various applications and settings: agents interacting with their environment and users, for robots navigating in dynamic environments or manipulating objects or vocalisations.

REFERENCES

- [1] Nicolas Duminy et al. “Intrinsically Motivated Open-Ended Multi-Task Learning Using Transfer Learning to Discover Task Hierarchy”. In: *Applied Sciences* 11.3 (2021). DOI: 10.3390/app11030975.
- [2] Lorenzo Jamone et al. “Affordances in Psychology, Neuroscience, and Robotics: A Survey”. In: *IEEE Transactions on Cognitive and Developmental Systems* 10.1 (2018), pp. 4–25. ISSN: 2379-8939. DOI: 10.1109/TCDS.2016.2594134.
- [3] Alexandre Manoury, Sao Mai Nguyen, and Cédric Buche. “Hierarchical Affordance Discovery Using Intrinsic Motivation”. In: *International Conference on Human-Agent Interaction (HAI)*. Vol. 7th. Kyoto, Japan: Association for Computing Machinery, Oct. 2019, pp. 186–193. ISBN: 978-1-4503-6922-0. DOI: 10.1145/3349537.3351898.
- [4] Sao Mai Nguyen et al. “Robots Learn Increasingly Complex Tasks with Intrinsic Motivation and Automatic Curriculum Learning”. In: *KI - Künstliche Intelligenz* (2021).
- [5] Pierre-Yves Oudeyer, Frédéric Kaplan, and Verena V. Hafner. “Intrinsic Motivation Systems for Autonomous Mental Development”. In: *IEEE Transactions on Evolutionary Computation* 11.2 (2007), pp. 265–286. DOI: 10.1109/TEVC.2006.890271.
- [6] Mark B. Ring. “Continual Learning in Reinforcement Environments”. Dissertation. Austin, TX, USA: University of Texas at Austin, Aug. 1994.
- [7] Sebastian Thrun and Lorien Pratt. “Learning to Learn: Introduction and Overview”. In: *Learning to Learn*. Boston, MA, USA: Springer US, 1998, pp. 3–17. ISBN: 978-1-4615-5529-2. DOI: 10.1007/978-1-4615-5529-2_1.



Nicolás Navarro-Guerrero Nicolás Navarro-Guerrero is an Assistant Professor at Aarhus University. He is co-founder and coordinator of the multidisciplinary AU Social Robotics Lab at Aarhus University. He is also the co-founder of the AU ENG Makerspace. He received a degree in Electronic Engineering from the Universidad Técnica Federico Santa María, Valparaíso, Chile, in 2010. Where he specialised in embedded systems and applied artificial intelligence and robotics. In 2016 he completed a Ph.D. in Computer Science at the University of Hamburg (Germany) under the supervision of Prof. Stefan Wermter. His dissertation concerned computational models for adaptive robot behaviour based on neural mechanisms of self-preservation. Nicolás was also a postdoc at the University of Hamburg where he expanded his research scope to Cognitive Robotics and Social Robotics. He is member of the editorial board of the journal IEEE TCDS, Frontiers in Robotics and AI, Frontiers in Neurorobotics and the chair of the Task force “Action and Perception” of the IEEE Technical Committee on Cognitive and Developmental Systems. For more information visit his webpage: <https://nicolas-navarro-guerrero.github.io>



Sao Mai Nguyen specialises in cognitive-developmental learning, reinforcement learning, imitation learning, curriculum learning for robots and human activity recognition. She received her PhD from Inria in 2013 in computer science, for her machine learning algorithms combining reinforcement learning and active imitation learning for interactive and multi-task learning. She holds an Engineer degree from Ecole Polytechnique, France and a master's degree in adaptive machine systems from Osaka University, Japan. She has enabled a robot to coach physical rehabilitation in the projects RoKInter and the experiment KERAAL she coordinated, funded by the European Union through FP-7 project ECHORD++. She has participated in project AMUSAAL, for analysing human activities of daily living through cameras, and CPER VITAAL for developing assistive technologies for the elderly and disabled. She is currently an assistant professor at Ensta-Paris and is affiliated to IMT Atlantique, France. She also acts as an associate editor of the journal IEEE TCDS and the co-chair of the Task force “Action and Perception” of the IEEE Technical Committee on Cognitive and Developmental Systems. For more information visit her webpage: <https://nguyensmai.free.fr>



Erhan Oztop earned his Ph.D. at the University of Southern California in 2002. In the same year, he joined the Computational Neuroscience Laboratories at the Advanced Telecommunications Research Institute International, (ATR) in Japan. There he worked as a researcher and later a senior research and group leader where he also served as vice department head for two research groups (2002-2011). Currently, he is a specially appointed professor at Osaka University, Japan, and also a professor in the Computer Science Department of Ozyegin University, Turkey. His research involves the computational study of intelligent behavior, human-in-the-loop systems, computational neuroscience, machine learning, cognitive and developmental robotics.



Junpei Zhong is a research assistant professor at the Hong Kong Polytechnic University. He graduated from the South China University of Technology with a Bachelor in Automation. Then he obtained his MPhil degree at the Hong Kong Polytechnic University, Hong Kong, and Ph.D. degree at the University of Hamburg, Germany. Before he returned to PolyU, he has been a researcher at the University of Hertfordshire (UK), University of Plymouth (UK), Waseda University (Japan) and the National Institute of Advanced Industrial Science and Technology (AIST, Japan), Nottingham Trent University (UK) etc. He has participated in several EU and Japanese projects centering around machine learning, assistive technologies, and cognitive robots. His current research focuses on cognitive models, intelligent robots and their usage in wellbeing and elderly healthcare. He is the organizer of the IEEE Symposium on Domestic Robotics 2019 (part of the Symposium Series of Computational Intelligence 2019) in Xiamen, China. He also organized special sessions and workshops in ICIRA 2019, IEEE-RAS Humanoids 2018, IEEE ICDL-EpiRob 2018, and guest editors of several journals such as IEEE Transaction on Cognitive and Developmental Systems, Interaction Studies and Complexity.