Guest Editorial
Special Issue on 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. First, 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. Second, 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 multitask learning, multimodal sensorimotor learning, and lifelong adaptation to wear and tear. Addressing these challenges promise higher levels of autonomy and versatility for 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 [item 7) in the Appendix], and in particular, the emergence of representations of action and perception in humans and artificial agents in a continual learning manner [item 6) in the Appendix]. These processes are done via action–perception cycle, active perception, continual sensorimotor learning, environmental-driven scaffolding, and intrinsic motivation [item 5) in the Appendix]. The special issue highlights behavioral 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 environments 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, multiagent and artificial life systems, computational neuroscience, and developmental psychology, on theoretical, computational, application-oriented, and experimental studies.

Digital Object Identifier 10.1109/TCDS.2021.3082880

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:

1) intrinsic motivation;
2) hierarchical policy learning;
3) continual sensorimotor learning;
4) motor primitives;
5) goal generation;
6) agent architectures;
7) curriculum learning;
8) learning from demonstration;
9) self-organizing behavior;
10) skill acquisition and planning;
11) environmental-driven scaffolding;
12) emergence of representations via continual interaction;
13) 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 article as follows.

Fournier *et al.* present the most theoretical article of this special issue in "CLIC: Curriculum learning and imitation for object control in nonrewarding 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 nonrewarding environment. The algorithm CLIC is a variant of the DQNfD for the case of multitask learning and augmented by intrinsic motivation. CLIC is tested in various grid nonrewarding multitask 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 nonreproducible and mastered interactions with objects. This approach is similar to research on interactive intrinsic motivation and hierarchical learning (e.g., [item 1) in the Appendix] and [item 4) in the Appendix]), but proposes a deep reinforcement learning algorithm.

Applying deep reinforcement learning and intrinsic motivation in a real-world scenario, the article titled "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 spatiotemporal 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.* integrates 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 multiobjective Markov decision processes" propose a novel intrinsically motivated multiobjective 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 multiobjective RL method for evolving policy coverage sets using hierarchical policy learning, generalize to nonstationary dynamics in the state transition distribution and reward distributions. The work is in line with developmental and epigenetic robotics research and introduces a novel multiobjective reinforcement learning framework to learn skill sets. The robot uses low-level data from the 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 titled "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 test their approach on different variations of the 2-D 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, and 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-organization of the sensorimotor map in the article titled "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-organization of an agent's sensorimotor maps in noisy environments. The key features of the proposed TMGWR networks are as follows.

1) 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.* [item 3) in the Appendix] where the sensorimotor experience builds simple tasks that can be used for planning or as subtasks for compound tasks.
2) 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. This 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 [item 2) in the Appendix]. 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 of the fact that intrinsic motivation should be balanced with other mechanisms such as those that control modularity and computational resources.

Likewise, the article titled "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 behaviors of humans stem from motor primitives or elementary building blocks, Gumbsch *et al.* propose a learning architecture called "surprise-based behavioral modularization." In this learning architecture, the association of the primitives is obtained by the self-organizing dynamics, which is called differential extrinsic plasticity. It is also shown that the learned behaviors speed up the learning of different goal-reaching tasks. As a result, various complex and highly coordinated behaviors are shown in robotic simulations. Even though the robot does not have any prior knowledge, the authors show that complex coordinated behaviors can be assembled from the motor primitives.

More theoretical, the article titled "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 will 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 can also be usable in real-world applications, these representations will be distorted and will require additional systems such as reinforcement learning to deal with the repeatability of agent's movements.

Finally, in the review article on vocal control learning titled "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 is used to lay down the behavioral and physiological features of vocal imitation. Then, the components of such learning systems are charted out, and the reviewed models are analyzed 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, and 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 this article extends well beyond vocal learning and imitation and can be used as a general reference for sensorimotor learning and its plausible biological modeling.

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 multitask settings? How can hierarchical learning take advantage of the 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 vocalizations.

Nicolás Navarro-Guerrero, *Co-Lead Guest Editor*
Department of Electrical and Computer Engineering
Aarhus University
8200 Aarhus, Denmark
E-mail: nicolas.navarro.guerrero@gmail.com

Sao Mai Nguyen, *Co-Lead Guest Editor*
U2IS, IMT Atlantique
44307 Nantes, France
Ensta-Paris
91120 Palaiseau, France
Inria
33405 Rocquencourt, France
E-mail: nguyensmai@gmail.com

Erhan Oztop, *Guest Editor*
Symbiotic Intelligent Systems Research Center (SISReC)
Osaka University
Osaka 565-0871, Japan
Computer Science Department
Özyeğin University
34794 İstanbul, Turkey
E-mail: erhan.oztop@otri.osaka-u.ac.jp

Junpei Zhong, *Guest Editor*
Department of Rehabilitation Sciences
The Hong Kong Polytechnic University
Hong Kong
E-mail: joni.zhong@ieee.org

Appendix
Related Work

1) N. Duminy, S. M. Nguyen, J. Zhu, D. Duhaut, and J. Kerdreux, "Intrinsically motivated open-ended multi-task learning using transfer learning to discover task hierarchy," *Appl. Sci.*, vol. 11, no. 3, p. 975, 2021, doi: 10.3390/app11030975.
2) L. Jamone *et al.*, "Affordances in psychology, neuroscience, and robotics: A survey," *IEEE Trans. Cogn. Develop. Syst.*, vol. 10, no. 1, pp. 4–25, Mar. 2018, doi: 10.1109/TCDS.2016.2594134.
3) A. Manoury, S. M. Nguyen, and C. Buche, "Hierarchical affordance discovery using intrinsic motivation," in *Proc. Int. Conf. Hum. Agent Interact. (HAI)*, vol. 7. Kyoto, Japan, Oct. 2019, pp. 186–193, doi: 1145/3349537.3351898.
4) S. M. Nguyen, N. Duminy, A. Manoury, D. Duhaut, and C. Buche, "Robots learn increasingly complex tasks with intrinsic motivation and automatic curriculum learning," *Künstliche Intelligenz*, vol. 35, pp. 81–90, Feb. 2021, doi: 10.1007/s13218-021-00708-8.
5) P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, "Intrinsic motivation systems for autonomous mental development," *IEEE Trans. Evol. Comput.*, vol. 11, no. 2, pp. 265–286, Apr. 2007, doi: 10.1109/TEVC.2006.890271.
6) M. B. Ring, "Continual learning in reinforcement environments," Ph.D. dissertation, Dept. Doctor Philos., Univ. Texas Austin, Austin, TX, USA, Aug. 1994.
7) S. Thrun and L. Pratt, "Learning to learn: Introduction and overview," in *Learning to Learn*. Boston, MA, USA: Springer, 1998, pp. 3–17, doi: 10.1007/978-1-4615-5529-2 1.

Nicolás Navarro-Guerrero received the degree in electronic engineering from the Universidad Técnica Federico Santa María, Valparaíso, Chile, in 2010, where he specialized in embedded systems and applied artificial intelligence and robotics and the Ph.D. degree in computer science from the University of Hamburg, Hamburg, Germany, in 2016, under the supervision of Prof. S. Wermter.

His dissertation concerned computational models for adaptive robot behavior based on neural mechanisms of self-preservation. He is an Assistant Professor with Aarhus University, Aarhus, Denmark, where he is the Co-Founder and the Coordinator of the multidisciplinary AU Social Robotics Lab. He is also the Co-Founder of the AU ENG Makerspace. He was also a Postdoctoral Fellow with the University of Hamburg, where he expanded his research scope to cognitive robotics and social robotics.

Dr. Navarro-Guerrero is a member of the editorial board of IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS, *Frontiers in Robotics and AI*, and *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 received the engineering degree from École Polytechnique, Palaiseau, France, the master's degree in adaptive machine systems from Osaka University, Osaka, Japan, in 2010, and the Ph.D. degree in computer science from Inria, Rocquencourt, France, in 2013, for her machine learning algorithms combining reinforcement learning and active imitation learning for interactive and multitask learning.

She specializes in cognitive-developmental learning, reinforcement learning, imitation learning, and curriculum learning for robots and human activity recognition. 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 analyzing 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 with Ensta-Paris, Palaiseau, and is affiliated to IMT Atlantique, Nantes, France.

Dr. Nguyen acts as an Associate Editor of IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS 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 received the Ph.D. degree from the University of Southern California, Los Angeles, CA, USA, in 2002.

In 2002, he joined the Computational Neuroscience Laboratories, Advanced Telecommunications Research Institute International, Seika, Japan, where he worked as a Researcher and later a Senior Research and Group Leader and also served as the Vice Department Head for two research groups from 2002 to 2011. He is currently a specially appointed Professor with Osaka University, Osaka, Japan, and also a Professor with the Computer Science Department, Özyeğin University, İstanbul, Turkey. His research involves the computational study of intelligent behavior, human-in-the-loop systems, computational neuroscience, machine learning, and cognitive and developmental robotics.

Junpei Zhong received the bachelor's degree in automation from the South China University of Technology, Guangzhou, China, in 2006, the M.Phil. degree from The Hong Kong Polytechnic University, Hong Kong, in 2009, and the Ph.D. degree from the University of Hamburg, Hamburg, Germany, in 2015.

He has been a Researcher with the University of Hertfordshire, Hatfield, U.K.; the University of Plymouth, Plymouth, U.K.; Waseda University, Tokyo, Japan; and the National Institute of Advanced Industrial Science and Technology, Tokyo; and Nottingham Trent University, Nottingham, U.K. He is a Research Assistant Professor with The Hong Kong Polytechnic University. 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.

Dr. Zhong 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, and IEEE ICDL-EpiRob 2018, and was a Guest Editor for several journals, such as IEEE TRANSACTION ON COGNITIVE AND DEVELOPMENTAL SYSTEMS, *Interaction Studies*, and *Complexity*.