Lifelong Robot Learning

Erhan Oztop[1,2] erhan.oztop@otri.osaka-u.ac.jp

Emre Ugur[3] emre.ugur@boun.edu.tr

[1]Osaka University, OTRI, SISReC
1-1 Yamadaoka, Suita, Osaka 565-0871, Japan
[2]Ozyegin University, Computer Science Department

[3]Bogazici University, Department of Computer Engineering
34342, Bebek, Istanbul, Turkey

1 Definition

Lifelong robot learning can be defined as the ability of a robot to acquire an ever-growing number of skills and knowledge over multiple domains through autonomous task engagement and learning.

2 Overview

The key characteristics of a life learning robot are autonomous continual learning and positive transfer of knowledge among learned tasks. The former indicates the ability to learn tasks in an incremental manner without forgetting the already learned skills. The latter indicates the bootstrapping leverage that can be obtained by previously learned skills. The concept of lifelong learning was initially introduced by Thrun and Mitchell (1995), where the main focus was knowledge transfer from one learning task to another in a given task domain. This was formalized by extending the Markov Decision Process framework by introducing multiple reward functions and contexts. The task of the robot was to learn individual policies for different contexts with different reward functions. Recently, the lifelong learning concept is revived with the increased popularity of deep learning (Parisi et al., 2019; Lesort et al., 2019). The developments in this front can be referred as lifelong machine learning (LML) to set it apart from lifelong robot learning (LRL). The main focus in LML is to develop a neural network architecture with a continual learning algorithm to allow transfer learning, i.e. accelerating the learning of the subsequent novel data or task without forgetting the knowledge or skill acquired by previous learning experiences (Silver et al., 2013). LRL often involves LML but it may have additional components as part of the robot cognitive architecture (Vernon et al., 2016) as embodiment and autonomous interaction with the environment is at the core of LRL (Asada et al., 2009). As such, a lifelong learning robot needs decision mechanisms to control for example which skills to improve upon and when to disengage an active task execution (see Figure 2). In this sense, with classical robot terminology, a lifelong learning robot can be considered as having 'behavior-based' control (see Arrichiello (2020) within this book) where behaviors are generated with respect to task learning and value systems.

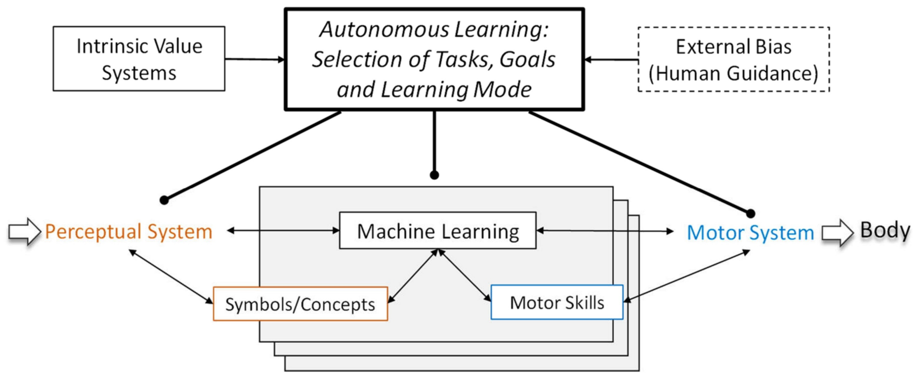

Figure 1: The general view of a lifelong robot learning (LRL) system is shown. Lifelong machine learning (LML) is an integral part of a LRL system; but it is not in general suitable for multi-domain heterogeneous learning tasks. Therefore, a LRL system often involves multiple LML systems with different architectures as the robot may be required to perform tasks that cannot be described by a single learning framework. The goal of LRL is then to create an autonomous system that can exploit multiple LML systems for effective and efficient continual learning.

In this chapter, we focus on research directions that are key for facilitating robot lifelong learning. To be concrete, in addition to LML concepts of multi-task learning and knowledge transfer, the topics of autonomous development, learning abstractions and movement representations are reviewed. In this chapter a 'task' refers to a learning task; but for modeling robot tasks in general the readers are referred to the chapter by Lima (2020) in this book. The current state of the art in LML is primarily dominated by deep learning based methods, which we sample and present representative approaches in this chapter; however for a wider coverage with additional considerations, the readers are referred to the recent reviews of Parisi et al. (2019) and Lesort et al. (2019).

3 Key Research Findings

3.1 Multiple Task Learning

A robot with lifelong learning ability must be able to learn multiple tasks that it may face in a sequential manner or in an interleaved regime according to its action selection mechanism or environment dynamics. When a learning system is trained sequentially on multiple tasks, the learning parameters that are important for a particular task may be changed because of subsequent learning for improving the performance on the novel task, thereby interfering negatively with the initial task performance. This phenomenon is classically called catastrophic interference or catastrophic forgetting (French, 1999, 1993). Thus, ability to learn tasks by focusing only on the current task is not alone sufficient for LRL because learning of the current task can destroy the earlier task knowledge violating one of the fundamental requirements of robot lifelong learning. There are several machine learning approaches aimed to tackle catastrophic forgetting which can be adopted to robotics, as briefly explained below.

Learning from interleaved task data. A straightforward solution to catastrophic forgetting is to use the data from multiple tasks in an interleaved manner, often called rehearsal (Rebuffi et al., 2017). So, this approach assumes that the data related to the tasks are available for the learner and can be rehearsed, i.e. sampled arbitrarily. In terms of a learning robot, this means that as the robot faces new learning tasks it has to also store the learning data so that it can be rehearsed while learning other tasks. Thus, this approach is not practical for learning a large number of tasks as in the case for robot lifelong learning. The workaround for this problem is so called intrinsic replay (Draelos et al., 2017), in which a generative model is learned to represent the samples experienced during task learning. The generative model is then used for emulating the rehearsal idea when learning a new task: instead of sampling from an experience buffer, the generative model is used to generate samples from the previously encountered data distribution(s). Both approaches are depicted in Figure 3.1A.

Selective updating of learning parameters. A clever strategy to avoid catastrophic forgetting is to detect the importance of learning parameters for the acquired tasks/knowledge, and use it to control the plasticity of the parameters when learning subsequent tasks . With this approach the critical parameters for the learned task are retained while other less critical parameters are allowed to change to fulfill the new task requirements (see Figure 3.1B). With this strategy Kirkpatrick et al. (2017) has shown that learning classification tasks without forgetting is possible with a single network - without neural expansion. The extension of this idea to the reinforcement learning (RL) domain was achieved by introducing higher-level mechanisms such as recognizing the task presented to the network and flexible switching between tasks, which allowed DQN, the deep neural network representing the state-action value function to learn to play multiple Atari games (Kirkpatrick et al., 2017). A similar but more biologically inspired approach is to require each network weight (i.e. synapse) to keep track of its importance based on the contribution it has in reducing the error for the learned tasks in an online fashion (Zenke et al., 2017). Both approaches yield a regularized loss function which protects the old knowledge being overwritten by subsequent task learning. Instead of computing or tracking the importance of neurons independently, Fernando et al. (2017) uses evolutionary algorithm concepts to find paths in the neural network representing a task skill that should be re-used in learning subsequent tasks. Catastrophic forgetting is avoided by the freezing of the weights on the discovered path to help preserve earlier task knowledge and facilitate skill transfer due to the weight reuse. It has been shown that this architecture is suitable for both classification and reinforcement learning (albeit with a two tasks scenario) tasks. However the application to real robotics scenarios is yet to be seen.

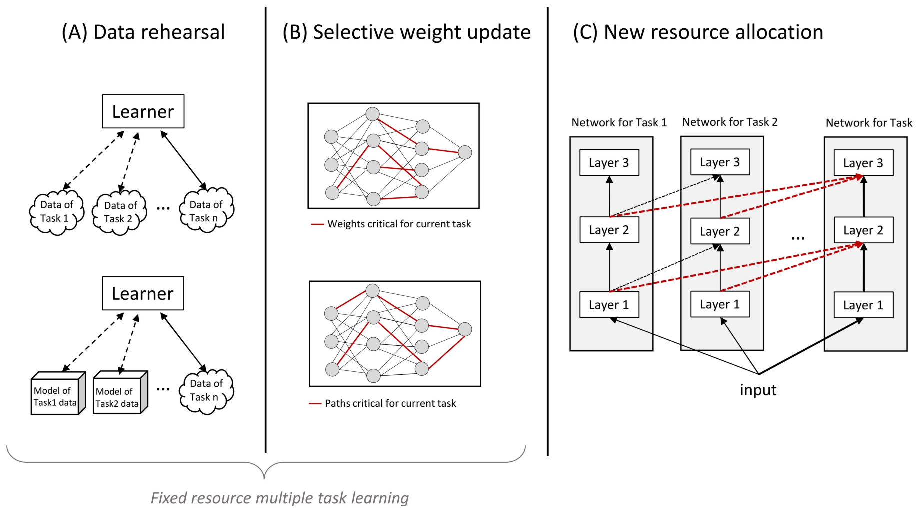

Figure 2: Several multiple task learning approaches from the literature are illustrated. The approaches in columns (A) & (B) aim to learn new tasks without growing the learned model/network; where in (C) additional resources are introduced to accommodate new task learning. (A) The data pertaining to earlier tasks are retained (upper panel, e.g. Rebuffi et al., 2017) or used to learn generative models (lower panel, e.g. Draelos et al., 2017), which are then used to perform interleaved training when learning subsequent tasks. (B) The plasticity of the weights are controlled based on their importance for the current task thereby avoiding catastrophic forgetting while learning new tasks. Importance of the weights can be set individually as shown in the upper half (e.g. Kirkpatrick et al., 2017; Zenke et al., 2017) or based on paths leading to the outputs as illustrated in the lower half (e.g. Fernando et al., 2017). (C) Additional resource allocation by creating of a new network for the new task is often accompanied by skill transfer through the creation of lateral connections to the layers of the networks that implement the earlier tasks. While training the new network, the weights related to the previous tasks may be allowed to change (e.g. Li and Hoiem, 2016) or kept fixed (e.g. Rusu et al., 2016a).

Allocating additional resources. The multi-task learning methods that use fixed computational resources are naturally bound to lose their ability to learn with excessive number of tasks. In this case, the problem is not only that new tasks cannot be learned; but the already learned ones may deteriorate with the attempt of learning a new task. Therefore approaches that grow a given learning system or create family of learners for solving ever-increasing tasks might be better suited for LRL. The straightforward approach to take here is to allocate independent resources for individual tasks to be learned. This could be a reasonable choice if the tasks are independent and limited in number. However, when tasks are logically related, learning of a new task may conflict with an earlier one. Furthermore, if the number of tasks are not limited, serious resource problems arise, unless no mechanisms are implemented for resource sharing and task similarity detection with task assimilation. Because of these reasons, usually resource allocation for new tasks is performed in moderation by sharing some of the existing knowledge and resources. In this vein, in the work of Rusu et al. (2016a) a new sub-network with random weights is created for an upcoming new task, while establishing lateral connections to the previously learned neural networks (see Figure 3.1C). To avoid the problem of destroying earlier knowledge, the network weights for the earlier tasks are not allowed to change when training for the new task takes place. This solves the problem of forgetting earlier skills while making use of learned skills for positive knowledge transfer for the novel tasks to be learned. A less radical approach taken by Li and Hoiem (2016) keeps a set of weights shared among tasks and allows the adaptation of all the weights with a special regularized cost to keep the old network outputs to be stable in the face of new training. It is shown that this approach is effective in incremental image classification tasks. However. it is not clear how this method will perform in reinforcement learning tasks, whereas the former approach i.e. the complete freezing of the earlier task weights idea has been shown to work in reinforcement learning and robotic manipulation tasks (Rusu et al., 2016a,b).

3.2 Knowledge Transfer

Knowledge transfer concept has been exploited extensively in general machine learning; here we take a neural network point of view and focus on the issues relevant for LRL. For a more general treatment of the topic the readers are referred to surveys avaliable in the literature (Pan and Yang, 2010; Zhuang et al., 2019; Weiss et al., 2016).

The knowledge transfer idea by itself does not require that the task providing the source for the knowledge transfer to be preserved. However, in the context of robot learning this is often the desired case, and luckily many continual multi-task learning approaches go hand-to-hand with knowledge transfer as hinted in the previous section. We first start with the simplest knowledge transfer ideas and then visit those that are more suitable for LRL.

Feature extraction. A straightforward way to transfer knowledge from a learned task to a similar one is to exploit the feature representations that have been obtained through an initial training with a general data set in subsequent learning task (Azizpour et al., 2016; Razavian et al., 2014). The features taken are usually the outputs of the last hidden layer of a deep neural network, where as the learning for subsequent tasks usually involves only the adaptation in a newly created output layer. This approach works well for visual recognition and classification tasks (Razavian et al., 2014; Guo et al., 2019) including learning the effect categories as the result of robot actions (Ugur and Piater, 2014); however it is not clear how this will generalize to complex robot (reinforcement) learning tasks as the features developed in the neural networks for implementing policy or value networks may be too domain specific.

Fine-tuning. Feature extraction does not change the feature representation inherited from earlier training. Thus, it only works well if the novel task is similar to the original task(s) that led to the formation of the features. If not, then one can choose to adapt the weights including also the ones producing the features, for improving the learning performance in the new task. This is often called network fine-tuning (Li and Hoiem, 2016). To avoid catastrophic forgetting and/or improving learning, several heuristics can be applied during fine-tuning such as using a lower learning rate for the weights involved in computing the features or freezing parts of the network (Rebuffi et al., 2017; Draelos et al., 2017).

Clone and fine-tune. One easy and effective approach to protect skills acquired earlier and facilitate skill transfer is to clone an existing network that represents an acquired skill, and then perform the training or fine-tuning on the cloned version. To assure positive skill transfer, the compatibility of the tasks must be assumed; otherwise, negative transfer may occur. So in a multi-task setting which network should be cloned for an upcoming new task is an issue to be solved. This problem can be circumvented by creating a new (sub)network with access to all earlier network outputs representing skills acquired so far, and by keeping only the newly created weights adaptive (e.g. Rusu et al., 2016a). This solves the problem of forgetting earlier skills while making use of the learned skills for knowledge transfer. However, the price paid for this is significant neural resource inflation as a huge network would be formed to accommodate an ever-growing number of tasks.

3.3 Autonomous lifelong learning and development

LML approaches in the literature either aim at creating a single dynamic network to represent the targeted skill set or aim at squeezing in as many tasks as possible to a fixed sized network with the goal of resource economy. Different form LML, in LRL, there is no merit in requiring a single (huge) network to capture all the skills of a robot. On the contrary, from a cognitive architecture perspective small modules that can be easily manipulated is more beneficial. Moreover, modularity is a basic principle that is believed to govern human motor control (Wolpert and Kawato, 1998; Imamizu et al., 2003; Oguz et al., 2018), and thus is a reasonable requirement for cognitive control architectures for robots. With this, we switch our attention from isolated LML systems to the requirements of full robotic architectures that are expected to engage in autonomous lifelong learning or development. In this framework, one of the critical issues is the question of among a possible set of tasks, which task the robot should choose to learn? In the same vein, within a chosen task, the learning samples that the robot experiences is not 'given' but comes out based on the robot action selection. This is so even the robot is not engaged in reinforcement learning; but simply tries to understand the effects of its actions via supervised learning. In general, the learning samples, the task order to be learned are not given but are the result of the action selection implemented within the robot cognitive architecture. Thus, robots with cognitive architectures capable of sensorimotor development need to have a value system that guides the selection of actions (Vernon et al., 2016). In the rest of the paper the concepts related to cognitive architectures and action selection/decision making in the context of LRL are analyzed.

3.3.1 Unsupervised Staged Development

Developmental robotics aims to develop lifelong learning robots following the principles inspired from infant development (Lungarella et al., 2003; Asada et al., 2009; Cangelosi and Schlesinger, 2015). Researchers in this field typically investigate various developmental stages of human infants such as emergent competencies until 12 months (Law et al., 2011) or development of tool use (Guerin et al., 2012), and apply the developed computational models on real robot platforms such as the iCub robot platform built for developmental research (Metta et al., 2010). While broad diversity of sensorimotor skills in various stages have been studied separately in the literature (Cangelosi and Schlesinger, 2015), approaches that span long-horizon staged development are missing. Exceptionally, Ugur et al. (2015) studied skill development that started from behavior primitive discovery stage, followed by learning of affordances (Gibson, 1986) and goal-emulation, and finally realized imitation learning exploiting parental scaffolding mechanisms (Berk and Winsler, 1995). However, the transition in the stages were manually designed in this work. On the other hand, for unconstrained lifelong learning, the robots should be able to autonomously decide which learning tasks to engage and disengage in different stages, with the possibility of implementing an interleaved task learning regime. In the next section, the concept of intrinsic motivation is presented as a potential developmental mechanism serving this purpose.

3.3.2 Intrinsic Motivation

In autonomous lifelong learning, robots are required to select what task to engage next, rather than being given a sequence of tasks to complete by a human tutor. More generally, the robot decision includes selecting which action to take, which skill to learn, which goal to achieve, which environment to interact, which agents to communicate etc. Especially for learning tasks where externally defined reward or assistance are not available, developmental robotics benefits from intrinsically motivated behaviors that maximize exploration, diversity, novelty, competence or learning progress. Various robot learning algorithms and applications exploited intrinsically motivated (IM) strategies in socially guided learning (Ivaldi et al., 2013; Duminy et al., 2019; Fournier et al., 2019), affordance learning (Ugur and Piater, 2014; Manoury et al., 2019; Baldassarre et al., 2019), and planning (Blaes et al., 2019). For example, a particular IM signal that maximizes learning progress (Oudeyer et al., 2007; Schmidhuber, 1991) can guide the robot to learning regions that are neither too easy nor too difficult to learn i.e., with the appropriate level of complexity which is inline with infant data (Kidd et al., 2012). While intrinsic motivation based exploration and learning might be one of the key characteristics of lifelong learning in robotics, it is a challenge to autonomously decide which type of signals (e.g. novelty or competence) to use and which learning space (e.g. space of objects, actions or goals) to explore in different stages of development.

3.4 Learning Abstractions

Through drives such as IM, the robots focus on first simple skills and then complex ones. Thus, lifelong learning requires the robots to acquire skills in different levels of complexities involving high-level ones such as abstract reasoning and symbolic communication capabilities. The building blocks of such high-level skills are composed of abstract structures including high-level concept representations, discrete symbols and logical rules.

Since the early days of intelligent robotics, researchers have studied how to bridge the representational gap between the continuous sensorimotor experience of the robots and discrete symbols and rules used by traditional AI planning systems. Traditionally, the cognitive architectures of these systems have been designed as multi-layered software architectures since the time of Shakey robot (Kuipers et al., 2017) where lower levels are used in perception and control of the robot, and higher levels are used in complex reasoning and planning.

Typically, the symbols and rules that are effective in making plans in the higher levels are manually coded first, and mappings from the higher levels to the sensorimotor experience of the robots are established through manual coding or learning (Petrick et al., 2008). However, in a life-long learning setting, the number and range of the symbols and rules are potentially unlimited and therefore pre-designing them is not possible. The alternative approach is to extract the abstract representations directly from the sensorimotor experience of the robot linking them to the goals and actions (Sun, 2000). In one of the pioneering studies, Pisokas and Nehmzow (2005) implemented a system on a simulated mobile robot that generates sub-symbolic structures that could be used in planning. Sub-symbolic planning in a more complex robotic setup was later realized by Ugur et al. (2011), where the robot learned to predict the change in object features in response to robot's own actions. The system was able to chain predicted effects, generating search trees of future states of sequence of possible actions, and finding plans in this tree to achieve given goals. While these earlier studies formed sub-symbolic structures and used them, emergence of fully symbolic structures that can be fed into AI planners has been recently studied by several groups (Ugur and Piater, 2015b,a; Konidaris et al., 2018; Konidaris, 2019; James et al., 2019). These studies generally find symbols through the organization of the continuous sensorimotor world of the robot by applying clustering or classification algorithms, which lead to perceptual classes that directly serve as the preconditions and outcomes of the action operators that are fed into AI planners. While these approaches are shown to be effective in simulation based RL settings and real robot manipulation tasks with fixed action sets, scaling-up these methods to larger action and interaction spaces with long-term learning in physical robots still stands as an elusive challenge.

3.5 Movement Representations for LRL

For robots, the movement generation ability is so central that often special mechanisms for learning movement trajectories are designed. For some cases, the acquisition of a desired skill is simply equivalent to movement trajectory learning; in other cases, a complex movement is composed of sequenced movements which must be individually learned. Thus, a brief review of popular movement representation is in order to assess their suitability for LRL.

Besides the classical trajectory representation of splines, in the last decades several movement representations with desirable properties have been proposed. In particular, approaches based on dynamic systems (Schaal, 2006) and statistical modeling (Calinon, 2016) have been popular in learning by demonstration applications (see also the chapter by Calinon (2020) in this book). The so called Dynamic Movement Primitives (DMPs) encode a demonstrated trajectory (a time varying vector) as a set of differential equations (corresponding to each component), implementing a spring-mass-damper system extended with a non-linear function. While DMPs are designed to encode and generate individual trajectories, Probabilistic Movement Primitives (ProMPs) (Paraschos et al., 2013) can represent distributions of trajectories, generating stochastic policies. Recently deep neural networks are leveraged for learning powerful and robust movement representations. Conditional Neural Movement Primitives (CNMPs) (Seker et al., 2019), for example, not only encode trajectory distributions similar to ProMPs, but also have the capability to encode multiple modes of operation for the same skill, and to learn non-linear relationships between environment properties and action trajectories from a few data samples as opposed to previous approaches.

An action encoded as one of these movement primitives should be adapted to new situations in order to achieve new goals in lifelong learning settings. For this, the robot is required to generate its own experiences (Hester et al., 2018; Vecerik et al., 2019) and improve the established trajectory generation policy based on signals about how close it is to achieve the new goals in new tasks. While adapting to new environments it is important to retain the previously learned skills (i.e. avoid catastrophic forgetting problem that was introduced earlier). Within the ProMPs framework, Stark et al. (2019) used KL divergence to stay close to the previous parameters of the model, avoiding the distortion on the shape of the original movement (Peters et al., 2010). However a new non-parametric ProMP was generated for each task variation. As a more general approach, Ewerton et al. (2019) used a single model which combines ProMP and Gaussian Processes to encode a parametric skill and condition it with the corresponding task parameters. For adaptation, RL was used to find the relations between the task parameters and the ProMP model parameters by using a trajectory relevance metric. Compared to the above studies, Adaptive Conditional Neural Movement Primitives (ACNMP) (Akbulut et al., 2020) did not require explicit optimization using metrics such as KL-divergence. Instead, ACNMP adopted a rehearsal type of approach to learn the demonstrated trajectories and the newly explored ones together, automatically preserving the old skills while extending the model to the new task parameters thanks to the robust and flexible representation ability of the network (Garnelo et al., 2018).

Although, powerful multiple trajectory learning methods have been developed, there are open questions yet to be answered. While the robot is increasing its movement capacity and learning new knowledge from the environment, it needs to automatically decide whether to assimilate the new knowledge into one of the existing skills (e.g. by modifying the related trajectory distribution) or accommodate a new movement primitive that represents this movement completely as a new skill. Thus, additional mechanisms should be developed for trajectory representation in a lifelong learning robot in analogy with the assimilation/accommodation mechanisms described for infants (Piaget and Cook, 1952).

4 Example Applications

Lifelong robot learning has a huge potential in real-word applications, although full fledged lifelong learning robots have not yet been available in commercial or industrial applications (Lesort et al., 2019). In particular, in addition to the proof-of-concept realization of lifelong learning systems in simplified settings and/or in simulation (e.g. Suro et al., 2021), applications in selected domains can be found.

Lifelong learning efforts directed towards robotic sub-tasks such as objection segmentation (Michieli and Zanuttigh, 2021) and object detection (Gepperth and Hammer, 2016) can be used to construct self-improving perception modules for various robotic architectures. For example, such modules can be used within autonomous driving systems together with end-to-end driving via (deep) reinforcement learning (Jaritz et al., 2018) and imitation learning (Codevilla et al., 2018) algorithms.

An important application domain that would benefit from lifelong robot learning is manipulation in home environments (Ersen et al., 2017). As a development in this front, Knowrob (http://knowrob.org) is proposed as a knowledge infrastructure to enable autonomous robots to perform everyday manipulation tasks by bridging the gap between vague task descriptions and the detailed information needed to actually to perform the tasks (Tenorth and Beetz, 2013).

The system can be considered a direction towards a LRL system, as it computes knowledge on demand by integrating internal or external information sources (such as internet-based knowledge bases which is open to change over time) with its perceptual processing. In the same vein, a web community for robot is proposed (Robo-brain, http://robobrain.me) for empowering autonomous robots by sharing their experience and task specific information via heterogeneous information sources such as internet, computer simulations and real-life robot trials(Saxena et al., 2015). Both Robo-brain and Knowrob are developments pointing towards cloud-based LRL architectures.

Another important application domain of LRL is social robotics, where robots are required to interact with humans and adapt to the dynamics of human behavior. Churamani et al. (2020b) argues that continual learning is an essential paradigm for affective robotics as the learning objectives may change rapidly while adapting to the behaviors of the users and their affective states and moods. As an example, Churamani et al. (2020a) propose a robot personality model for collaborative human-robot interactions, which generates personality-driven behaviour in negotiation scenarios. In the social learning approach taken by Tjomsland et al. (2020), the robot is proposed to learn to assess the social appropriateness of its behaviors in a continual manner. Yet in another social life-long learning application, the human facial data is incrementally acquired during social interactions for developing a facial recognition system that takes individual differences into consideration (Churamani and Gunes, 2020).

Although most of the current LRL applications, focus on limited domains and given as proof of concept or constrained to simulation, these endeavours are important steps towards real robotic applications. Such efforts coupled with developments in LML create a strong impetus for the research towards realizing robotic systems with continual world modeling and task learning ability that can be deployed in general task domains.

5 Future Directions for Research

As hinted in the previous section, a grand challenge for robotics research is to develop cognitive architectures combining control, learning and reasoning so as to sustain autonomous lifelong learning in general settings. In the recent years significant progress in error gradient based deep learning and symbolic/hybrid reasoning approaches have been made. In particular, the developments in flexible movement representations, symbol formation, continual learning and skill transfer form a firm basis for robot lifelong learning. It is expected that lifelong learning robots will include a variety of cognitive modules representing e.g. intrinsic motivation and other value systems to guide task engagement and information seeking behavior. For the robots in social settings, additional modules inspired from social psychology and neuroscience implementing a range of modulatory systems such as artificial pain (Asada, 2019), emotion (Michaud et al., 2000; Kirtay et al., 2019) and empathy (Asada, 2015) are required for human-like autonomous learning, for which the analysis is left for another review.

To sum up, with the transfer of growing knowledge in artificial intelligence to robotics, one may expect a breakthrough in the next decade in how robotic systems acquire knowledge and skills autonomously to become more dexterous and socially competent. Convergence of the research on developmental robotics and lifelong machine learning towards developing multi-domain, multi-task human-like learning robots will be the key for the realization of this breakthrough.

References

M. Akbulut, E. Oztop, Y. Seker, H. Xue, A. Tekden, and E. Ugur. Acnmp: Skill transfer and task extrapolation through learning from demonstration and reinforcement learning via representation sharing. In *Conference on Robot Learning (CoRL)*, 2020. accepted.

F. Arrichiello. *Behavior-Based Systems*, pages 1–7. Springer Berlin Heidelberg, Berlin, Heidelberg, 2020. ISBN 978-3-642-41610-1.

M. Asada. Development of artificial empathy. *Neuroscience Research*, 90:41–50, 2015. ISSN 0168-0102.

M. Asada. Artificial pain may induce empathy, morality, and ethics in the conscious mind of robots. *Philosophies*, 4(3):38, 2019. ISSN 2409-9287.

M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida. Cognitive developmental robotics: A survey. *IEEE transactions on autonomous mental development*, 1(1):12–34, 2009.

H. Azizpour, A. S. Razavian, J. Sullivan, A. Maki, and S. Carlsson. Factors of transferability for a generic convnet representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 38(9):1790–1802, 2016.

G. Baldassarre, W. Lord, G. Granato, and V. G. Santucci. An embodied agent learning affordances with intrinsic motivations and solving extrinsic tasks with attention and one-step planning. *Frontiers in neurorobotics*, 13:45, 2019.

L. E. Berk and A. Winsler. *Scaffolding Children's Learning: Vygotsky and Early Childhood Education. NAEYC Research into Practice Series. Volume 7.* ERIC, 1995.

S. Blaes, M. V. Pogančić, J. Zhu, and G. Martius. Control what you can: Intrinsically motivated task-planning agent. In *Advances in Neural Information Processing Systems*, pages 12541–12552, 2019.

S. Calinon. A tutorial on task-parameterized movement learning and retrieval. *Intelligent Service Robotics*, 9(1):1–29, 2016.

S. Calinon. *Learning from Demonstration (Programming by Demonstration)*, pages 1–8. Springer Berlin Heidelberg, Berlin, Heidelberg, 2020. ISBN 978-3-642-41610-1.

A. Cangelosi and M. Schlesinger. *Developmental robotics: From babies to robots.* MIT press, 2015.

N. Churamani and H. Gunes. Clifer: Continual learning with imagination for facial expression recognition. In *Proceedings of the 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pages 693–699, 2020.

N. Churamani, P. Barros, H. Gunes, and S. Wermter. Affect-driven modelling of robot personality for collaborative human-robot interactions. *arXiv preprint arXiv:2010.07221*, 2020a.

N. Churamani, S. Kalkan, and H. Gunes. Continual learning for affective robotics: Why, what and how? In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 425–431. IEEE, 2020b.

F. Codevilla, M. Müller, A. López, V. Koltun, and A. Dosovitskiy. End-to-end driving via conditional imitation learning. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4693–4700. IEEE, 2018.

T. J. Draelos, N. E. Miner, C. C. Lamb, J. A. Cox, C. M. Vineyard, K. D. Carlson, W. M. Severa, C. D. James, and J. B. Aimone. Neurogenesis deep learning: Extending deep networks to accommodate new classes. In *2017 International Joint Conference on Neural Networks, IJCNN 2017, Anchorage, AK, USA, May 14-19, 2017*, pages 526–533. IEEE, 2017.

N. Duminy, S. M. Nguyen, and D. Duhaut. Learning a set of interrelated tasks by using a succession of motor policies for a socially guided intrinsically motivated learner. *Frontiers in neurorobotics*, 12:87, 2019.

M. Ersen, E. Oztop, and S. Sariel. Cognition-enabled robot manipulation in human environments: Requirements, recent work, and open problems. *IEEE Robotics Automation Magazine*, 24(3):108–122, 2017. doi: 10.1109/MRA.2016.2616538.

M. Ewerton, O. Arenz, G. Maeda, D. Koert, Z. Kolev, M. Takahashi, and J. Peters. Learning trajectory distributions for assisted teleoperation and path planning. *Frontiers in Robotics and AI*, 6:89, 2019.

C. Fernando, D. Banarse, C. Blundell, Y. Zwols, D. Ha, A. A. Rusu, A. Pritzel, and D. Wierstra. Pathnet: Evolution channels gradient descent in super neural networks. *CoRR*, abs/1701.08734, 2017.

P. Fournier, C. Colas, M. Chetouani, and O. Sigaud. Clic: Curriculum learning and imitation for object control in non-rewarding environments. *IEEE Transactions on Cognitive and Developmental Systems*, 2019.

R. M. French. Catastrophic interference in connectionist networks: Can it be predicted, can it be prevented? In J. D. Cowan, G. Tesauro, and J. Alspector, editors, *Advances in Neural Information Processing Systems 6, [7th NIPS Conference, Denver, Colorado, USA, 1993]*, pages 1176–1177. Morgan Kaufmann, 1993.

R. M. French. Catastrophic forgetting in connectionist networks. *Trends Cogn Sci*, 3(4):128–135, 1999. ISSN 1879-307X (Electronic) 1364-6613 (Linking).

M. Garnelo, D. Rosenbaum, C. Maddison, T. Ramalho, D. Saxton, M. Shanahan, Y. W. Teh, D. Rezende, and S. M. A. Eslami. Conditional neural processes. In *ICML*, 2018.

A. Gepperth and B. Hammer. Incremental learning algorithms and applications. In *European symposium on artificial neural networks (ESANN)*, 2016.

J. J. Gibson. The ecological approach to visual perception. hills-dale. *NJ: Lawrence*, 1986.

F. Guerin, N. Kruger, and D. Kraft. A survey of the ontogeny of tool use: from sensorimotor experience to planning. *IEEE Transactions on Autonomous Mental Development*, 5(1):18–45, 2012.

Y. Guo, H. Shi, A. Kumar, K. Grauman, T. Rosing, and R. S. Feris. Spottune: Transfer learning through adaptive fine-tuning. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pages 4805–4814. Computer Vision Foundation / IEEE, 2019.

T. Hester, M. Vecerik, O. Pietquin, M. Lanctot, T. Schaul, B. Piot, D. Horgan, J. Quan, A. Sendonaris, I. Osband, et al. Deep q-learning from demonstrations. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

H. Imamizu, T. Kuroda, S. Miyauchi, T. Yoshioka, and M. Kawato. Modular organization of internal models of tools in the human cerebellum. *Proceedings of the National Academy of Sciences of the United States of America*, 100(9): 5461–6, 2003. ISSN 0027-8424 (Print) 0027-8424 (Linking).

S. Ivaldi, N. Lyubova, A. Droniou, D. Gerardeaux-Viret, D. Filliat, V. Padois, O. Sigaud, P.-Y. Oudeyer, et al. Learning to recognize objects through curiosity-driven manipulation with the icub humanoid robot. In *2013 IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, pages 1–8. IEEE, 2013.

S. James, B. Rosman, and G. Konidaris. Learning portable representations for high-level planning. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, 2019.

M. Jaritz, R. De Charette, M. Toromanoff, E. Perot, and F. Nashashibi. End-to-end race driving with deep reinforcement learning. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2070–2075. IEEE, 2018.

C. Kidd, S. T. Piantadosi, and R. N. Aslin. The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one*, 7(5):e36399, 2012.

J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska, D. Hassabis, C. Clopath, D. Kumaran, and R. Hadsell. Overcoming catastrophic forgetting in neural networks. *Proc Natl Acad Sci U S A*, 114(13):3521–3526, 2017. ISSN 1091-6490 (Electronic) 0027-8424 (Linking).

M. Kirtay, L. Vannucci, U. Albanese, C. Laschi, E. Oztop, and E. Falotico. Emotion as an emergent phenomenon of the neurocomputational energy regulation mechanism of a cognitive agent in a decision-making task. *Adaptive Behavior*, 0(0):1059712319880649, 2019.

G. Konidaris. On the necessity of abstraction. *Current Opinion in Behavioral Sciences*, 29:1–7, 2019.

G. Konidaris, L. P. Kaelbling, and T. Lozano-Perez. From skills to symbols: Learning symbolic representations for abstract high-level planning. *Journal of Artificial Intelligence Research*, 61:215–289, 2018.

B. Kuipers, E. A. Feigenbaum, P. E. Hart, and N. J. Nilsson. Shakey: From conception to history. *AI Magazine*, 38(1):88–103, 2017.

J. Law, M. Lee, M. Hülse, and A. Tomassetti. The infant development timeline and its application to robot shaping. *Adaptive Behavior*, 19(5):335–358, 2011.

T. Lesort, V. Lomonaco, A. Stoian, D. Maltoni, D. Filliat, and N. Díaz-Rodríguez. Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges, 2019.

Z. Li and D. Hoiem. Learning without forgetting. In B. Leibe, J. Matas, N. Sebe, and M. Welling, editors, *Computer Vision – ECCV 2016*, pages 614–629. Springer International Publishing, 2016. ISBN 978-3-319-46493-0.

P. U. Lima. *Robot Task Modeling*, pages 1–12. Springer Berlin Heidelberg, Berlin, Heidelberg, 2020. ISBN 978-3-642-41610-1.

M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini. Developmental robotics: a survey. *Connection science*, 15(4):151–190, 2003.

A. Manoury, S. M. Nguyen, and C. Buche. Hierarchical affordance discovery using intrinsic motivation. In *Proceedings of the 7th International Conference on Human-Agent Interaction*, pages 186–193, 2019.

G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. Von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor, et al. The icub humanoid robot: An open-systems platform for research in cognitive development. *Neural networks*, 23(8-9):1125–1134, 2010.

F. Michaud, P. Pirjanian, J. Audet, and D. Létourneau. *Artificial Emotion and Social Robotics*, pages 121–130. Springer Japan, Tokyo, 2000. ISBN 978-4-431-67919-6.

U. Michieli and P. Zanuttigh. Knowledge distillation for incremental learning in semantic segmentation. *Computer Vision and Image Understanding*, 205: 103167, 2021.

O. S. Oguz, Z. Zhou, S. Glasauer, and D. Wollherr. An inverse optimal control approach to explain human arm reaching control based on multiple internal models. *Sci Rep*, 8(1):5583, 2018. ISSN 2045-2322 (Electronic)2045-2322 (Linking).

P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner. Intrinsic motivation systems for autonomous mental development. *IEEE transactions on evolutionary computation*, 11(2):265–286, 2007.

S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010. ISSN 1558-2191.

A. Paraschos, C. Daniel, J. Peters, and G. Neumann. Probabilistic movement primitives. In *NIPS*, pages 2616–2624, 2013.

G. I. Parisi, R. Kemker, J. L. Part, C. Kanan, and S. Wermter. Continual lifelong learning with neural networks: A review. *Neural Networks*, 113:54–71, 2019. ISSN 0893-6080.

J. Peters, K. Mulling, and Y. Altun. Relative entropy policy search. In *Twenty-Fourth AAAI Conference on Artificial Intelligence*, 2010.

R. Petrick, D. Kraft, K. Mourao, N. Pugeault, N. Krüger, and M. Steedman. Representation and integration: Combining robot control, high-level planning, and action learning. In *Proceedings of the 6th international cognitive robotics workshop*, pages 32–41, 2008.

J. Piaget and M. Cook. *The origins of intelligence in children*, volume 8. International Universities Press New York, 1952.

J. Pisokas and U. Nehmzow. Experiments in subsymbolic action planning with mobile robots. In *Adaptive Agents and Multi-Agent Systems II, Lecture Notes in AI*, pages 80–87. Springer, 2005.

A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson. CNN features off-the-shelf: An astounding baseline for recognition. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR Workshops 2014, Columbus, OH, USA, June 23-28, 2014*, pages 512–519. IEEE Computer Society, 2014.

S. Rebuffi, A. Kolesnikov, G. Sperl, and C. H. Lampert. icarl: Incremental classifier and representation learning. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 5533–5542. IEEE Computer Society, 2017.

A. A. Rusu, N. C. Rabinowitz, G. Desjardins, H. Soyer, J. Kirkpatrick, K. Kavukcuoglu, R. Pascanu, and R. Hadsell. Progressive neural networks, 2016a.

A. A. Rusu, M. Vecerík, T. Rothörl, N. Heess, R. Pascanu, and R. Hadsell. Sim-to-real robot learning from pixels with progressive nets. *CoRR*, abs/1610.04286, 2016b.

A. Saxena, A. Jain, O. Sener, A. Jami, D. K. Misra, and H. S. Koppula. Robobrain: Large-scale knowledge engine for robots. *arXiv*, 1412.0691, 2015.

S. Schaal. Dynamic movement primitives-a framework for motor control in humans and humanoid robotics. In *Adaptive Motion of Animals and Machines*, pages 261–280. Springer, 2006.

J. Schmidhuber. Curious model-building control systems. In *Proc. international joint conference on neural networks*, pages 1458–1463, 1991.

M. Y. Seker, M. Imre, J. Piater, and E. Ugur. Conditional neural movement primitives. In *Robotics: Science and Systems (RSS)*, 2019.

D. L. Silver, Q. Yang, and L. Li. Lifelong machine learning systems: Beyond learning algorithms. In *Lifelong Machine Learning, Papers from the 2013 AAAI Spring Symposium, Palo Alto, California, USA, March 25-27, 2013*, volume SS-13-05 of *AAAI Technical Report*. AAAI, 2013.

S. Stark, J. Peters, and E. Rueckert. Experience reuse with probabilistic movement primitives. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.

R. Sun. Symbol grounding: A new look at an old idea. *Philosophical Psychology*, 13(149–172), 2000.

F. Suro, J. Ferber, T. Stratulat, and F. Michel. A hierarchical representation of behaviour supporting open ended development and progressive learning for artificial agents. *Autonomous Robots*, 45(2):245–264, 2021. ISSN 1573-7527. doi: 10.1007/s10514-020-09960-7. URL `https://doi.org/10.1007/s10514-020-09960-7`.

M. Tenorth and M. Beetz. Knowrob: A knowledge processing infrastructure for cognition-enabled robots. *The International Journal of Robotics Research*, 32 (5):566–590, 2013.

J. Tjomsland, S. Kalkan, and H. Gunes. Mind your manners! a dataset and a continual learning approach for assessing social appropriateness of robot actions. *arXiv preprint arXiv:2007.12506*, 2020.

E. Ugur and J. Piater. Emergent structuring of interdependent affordance learning tasks. In *4th International Conference on Development and Learning and on Epigenetic Robotics*, pages 489–494. IEEE, 2014.

E. Ugur and J. Piater. Refining discovered symbols with multi-step interaction experience. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pages 1007–1012. IEEE, 2015a.

E. Ugur and J. Piater. Bottom-up learning of object categories, action effects and logical rules: From continuous manipulative exploration to symbolic planning. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2627–2633. IEEE, 2015b.

E. Ugur, E. Oztop, and E. Sahin. Goal emulation and planning in perceptual space using learned affordances. *Robotics and Autonomous Systems*, 59(7–8): 580–595, 2011.

E. Ugur, Y. Nagai, E. Sahin, and E. Oztop. Staged development of robot skills: Behavior formation, affordance learning and imitation with motionese. *IEEE Transactions on Autonomous Mental Development*, 7(2):119–139, 2015.

M. Vecerik, O. Sushkov, D. Barker, T. Rothörl, T. Hester, and J. Scholz. A practical approach to insertion with variable socket position using deep reinforcement learning. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 754–760. IEEE, 2019.

D. Vernon, C. von Hofsten, and L. Fadiga. Desiderata for developmental cognitive architectures. *Biologically Inspired Cognitive Architectures*, 18:116–127, 2016. ISSN 2212-683X.

K. Weiss, T. M. Khoshgoftaar, and D. Wang. A survey of transfer learning. *Journal of Big Data*, 3(1):9, 2016. ISSN 2196-1115.

D. M. Wolpert and M. Kawato. Multiple paired forward and inverse models for motor control. *Neural Networks*, 11(7-8):1317–1329, 1998.

F. Zenke, B. Poole, and S. Ganguli. Continual learning through synaptic intelligence. *Proceedings of machine learning research*, 70:3987–3995, 2017. ISSN 2640-3498.

F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He. A comprehensive survey on transfer learning, 2019.