Proceedings of Machine Learning Research 123:142–152, 2020 NeurIPS 2019 Competition and Demonstration Track

REAL-2019: Robot open-Ended Autonomous Learning competition

Emilio Cartoni EMILIO.CARTONI@ISTC.CNR.IT
Francesco Mannella FRANCESCO.MANNELLA@ISTC.CNR.IT
Vieri Giuliano Santucci VIERI.SANTUCCI@ISTC.CNR.IT
Institute of Cognitive Sciences and Technologies
National Research Council

Jochen Triesch TRIESCH@FIAS.UNI-FRANKFURT.DE
Frankfurt Institute for Advanced Studies

Elmar Rueckert RUECKERT@ROB.UNI-LUEBECK.DE
University of Luebeck

Gianluca Baldassarre GIANLUCA.BALDASSARRE@ISTC.CNR.IT
Institute of Cognitive Sciences and Technologies
National Research Council

Editors: Hugo Jair Escalante and Raia Hadsell

Abstract

Open-ended learning, also called 'life-long learning' or 'autonomous curriculum learning', aims to program machines and robots that autonomously acquire knowledge and skills in a cumulative fashion. We illustrate the first edition of the *REAL-2019 – Robot open-Ended Autonomous Learning* competition, prompted by the EU project *GOAL-Robots – Goal-based Open-ended Autonomous Learning Robots*. The competition was based on a simulated robot that: (a) acquires sensorimotor competence to interact with objects on a table; (b) learns autonomously based on mechanisms such as curiosity, intrinsic motivations, and self-generated goals. The competition featured a first 'intrinsic phase', where the robots learned to interact with the objects in a fully autonomous way (no rewards, predefined tasks or human guidance), and a second 'extrinsic phase', where the acquired knowledge was evaluated with tasks unknown during the first phase. The competition ran online on *AIcrowd* for six months, involved 75 subscribers and 6 finalists, and was presented at NeurIPS-2019. The competition revealed very hard as it involved difficult machine learning challenges usually tackled in isolation, such as exploration, sparse rewards, object learning, generalisation, catastrophic interference, and autonomous skill learning. Following the participant's positive feedback, the preparation of a second *REAL-2020* competition is underway, improving on the formulation of a relevant benchmark for open-ended learning.

Keywords: Simulated robot, autonomous open-ended learning, intrinsic motivations.

1. Introduction

Background and impact. Evolution selected mechanisms that allow mammals to gradually acquire knowledge and skills during life to behave adaptively in environments having a great complexity. While the satisfaction of biological and social needs are major drives for learning, in more sophisticated animals curiosity and intrinsic motivations represent fundamental means for learning as they can guide the acquisition of knowledge and skills when these are not immediately needed (Berlyne, 1960; Baldassarre, 2011; Gottlieb et al., 2013).

Increasingly sophisticated learning algorithms and robots have been recently developed, but their autonomy and versatility are still limited in comparison to those of animals. While this is not a problem when solving specific predefined tasks, the lack of autonomy of present robots prevents them from successfully acting in unstructured/non-engineered environments where they face situations requiring sensorimotor skills unknown at design-time. As summarised by some industrial stakeholders[1]: “Present-day robots are made for the purpose of repeating several tasks thousands of times. Future robots, on the other hand, will have to perform thousands of tasks several times”. It is impossible to directly program robots to solve those tasks because they are both many and unknown beforehand. A solution to this problem is that robots *autonomously* generate those tasks and learn to solve them. The major objective of the REAL-2019 competition is the creation of an *open-ended learning benchmark* to attract work of the community on this great challenge, to isolate its key underlying problems, and to compare alternative solutions.

In the last decades, the autonomous learning of multiple tasks has been tackled under different headings providing results relevant for open-ended learning. In the field of *developmental robotics* (Lungarella et al., 2003; Cangelosi and Schlesinger, 2015), task-agnostic signals produced by algorithms implementing different mechanisms for *intrinsic motivations* (Barto et al., 2004; Oudeyer et al., 2007; Baldassarre and Mirolli, 2013; Baldassarre, 2011) have been proposed to drive autonomous exploration and skill learning (Santucci et al., 2014a; Schembri et al., 2007; Schmidhuber, 2010; Tanneberg et al., 2019). Hierarchical reinforcement learning (Barto and Mahadevan, 2003) has been combined with IMs, and with deep learning techniques (Kulkarni et al., 2016), for the autonomous learning of skills sequences. Other works employ IMs to support the autonomous sampling/discovery/generation of goals, intended as representations of states or sets of states that the agent can pursue with its action (e.g., Santucci et al., 2013; Forestier et al., 2017; Nair et al., 2018). The autonomous identification of goals can support open-ended leaning as it allows the autonomous generation of tasks to acquire the skills directed to pursue the goals (Santucci et al., 2014b). An increasing number of works thus focuses on the development of agents able to autonomously form new goals and learn the related skills (Held et al., 2017; Meeden and Blank, 2017; Nair et al., 2018; Rolf and Asada, 2014; Santucci et al., 2016; Seepanomwan et al., 2017) based on the saliency of world states (Barto et al., 2004), the change of states (Santucci et al., 2016; Sperati and Baldassarre, 2018), eigenoptions (Machado et al., 2017), density models (Bellemare et al., 2016), entropy (Eysenbach et al., 2018), and variational inference (Achiam et al., 2018). All these approaches are relevant for the open-ended learning challenge operationalised in the REAL competition benchmark.

Novelty of the competition. While autonomous open-ended learning has been studied for several years within the developmental robotics community and is now getting increasing attention in the machine learning and robotic communities, we still do not have a standard test to compare different systems. Video games, such as Atari’s, have been profitably used in AI research, but they involve simplified actuators with respect to robotic setups and they are usually used to accomplish externally defined tasks. Some existing competitions share

1. Manfred Gundel, CEO KUKA Robotics, *World Robotics*, 2011, `http://www.worldrobotics.org`

similarities with REAL but also have relevant differences. As REAL, the *AutoML for Life-long Machine Learning*[2] competition aims to develop systems able to acquire an increasing amount of knowledge. However, contrary to REAL it focuses on learning increasing amounts of input-output data furnished externally, rather than on embodied systems interacting with a physical world to actively generate new experience. *Animal-AI Olympics*[3] are focused on simulated animal-like robots interacting with physical environments. However, it differs from REAL because robots are tested with a set of specific tasks defined through reward functions, rather than requiring the autonomous creation of tasks. As REAL, the *ICDL MODELbot Challenge*[4] is focused on developmental processes. However, it also involve social tasks (e.g., imitation, social learning), and is not focused on a specific benchmark but rather on an article-based jury's evaluation of the scientific quality of the reproduction of a target empirical database chosen from three possible experiments.

2. The REAL challenge

Setup. REAL involves a simulated robot that has to autonomously learn how to act in a 'kitchen-like' environment to later solve some tasks unknown while learning (Fig. 1). The robot is formed by: a Kuka robotic arm with 7 degrees of freedom; a gripper with 2 degrees of freedom; and a camera viewing the scene from the top. The environment is a simplified kitchen scenario formed by: a table; a shelf on the table; a cube and two kitchen objects – a mustard bottle and a tomato can – derived from the *YCB standard object set*[5].

Structure of the competition. Each competing system undergoes two phases: an intrinsic phase of learning and an extrinsic phase of testing. In the *intrinsic phase*, the robot autonomously interacts with the environment for 10,000,000 simulation steps and should acquire as much knowledge and skills as possible to best solve the tasks in the extrinsic phase. During the *extrinsic phase*, the knowledge acquired during the intrinsic phase is evaluated with tasks unknown during the intrinsic phase. The robot can also learn in the extrinsic phase, but this is of little help for the limited time available to solve each task.

During the extrinsic phase, the robot has to solve 3 types of tasks ('challenges'; Fig. 1). For each task: (1) the robot is shown a certain configuration of the 3 objects in the environment ('overall goal'); (2) the objects are set in a different position and orientation in the environment; (3) the robot is given some time to bring the objects to the overall-goal configuration. The three challenges involve tasks with different types of object configurations: (1) *2D challenge*: the overall goal is a configuration of the 3 objects placed on the table plane; compared to the initial configuration, 1, 2, or 3 objects have to be moved on the table plane; the orientation of the objects in the initial and final configurations is the same, and the objects are separated from each other. (2) *2.5D challenge*: the overall goal is a configuration of the 3 objects placed either on the table plane or on the shelf; compared to the initial configuration, 1, 2, or 3 objects have to be moved from the plane to shelf or vice-versa; the orientation of the objects in the initial and final configurations is the same,

2. http://automl.chalearn.org/life-long-learning
3. http://animalaiolympics.com
4. https://icdl-epirob2019.org/modelbot-challenge
5. https://arxiv.org/pdf/1502.03143.pdf

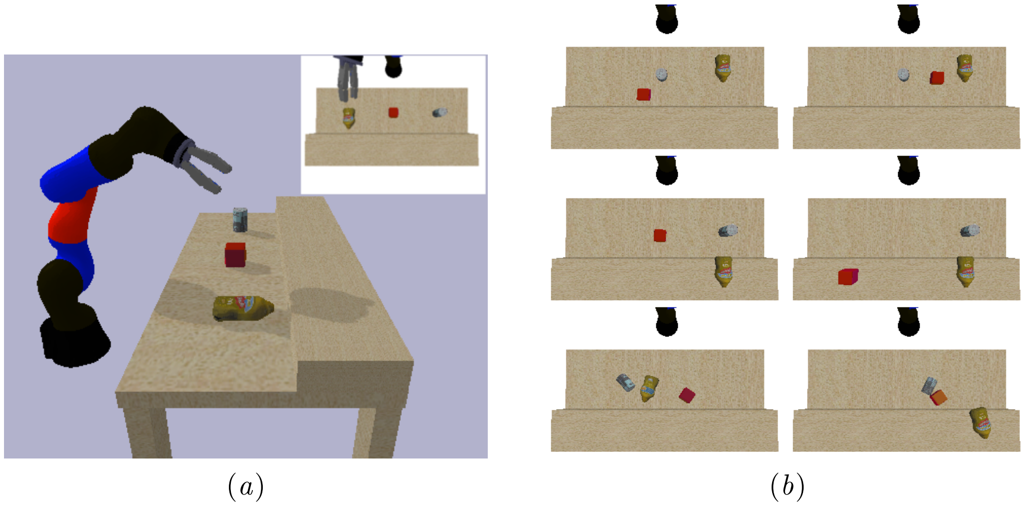

Figure 1: (a) The environment with the robot, the table, the three objects, and the shelf. The inset shows the environment as seen from the robot top view camera. (b) Initial (left) and overall-goal (right) configurations of three types of goals of the extrinsic phase. From top to bottom: 2D goal; 2.5D goal; 3D goal.

and the objects are separated from each other. (3) *3D challenge*: the overall goal is a configuration of the 3 objects placed anywhere on the table plane or on the shelf; the objects can have any initial/final orientation and may touch/overlap. The extrinsic phase involved 350 different overall goals each to be solved in 2,000 simulation steps: 150 goals for the 2D challenge (50 for each of the 1/2/3-object tasks); 150 goals for the 2.5D challenge (50 for each of the 1/2/3-object tasks); 50 goals for the 3D challenge (3-object tasks).

Note how the only regularities ('structure') shared between the intrinsic and extrinsic phases involve the environment and objects: in the intrinsic phase the robot has no knowledge about which tasks it will have to solve in the extrinsic phase. Therefore, in the intrinsic phase the robot should undergo an autonomous open-ended learning process to acquire, in the available time, as much knowledge and as many skills as possible.

Metrics. The performance of the extrinsic phase for an overall goal g was scored according to the following metrics M_g:

$$M_g = \frac{1}{3}\sum_{o}^{3}\left[w \cdot e^{-p||\mathbf{p}_o^* - \mathbf{p}_o||} + (1-w) \cdot e^{-a||\mathbf{a}_o^* - \mathbf{a}_o||}\right] \tag{1}$$

where $\mathbf{p}^*$ is the goal position (x,y,z) vector of the mass center of the object indexed by $o = \{1, 2, 3\}$ in the overall goal, $\mathbf{p}$ is the position of the object at the end of the task after the robot attempts to bring it to the goal position, p is a constant ensuring that this part of the score will be 0.25 if the distance to the goal position is 0.05 (5 cm); $\mathbf{a}^*$, $\mathbf{a}$, and a, are the analogous variables for the object orientation, with a set so that the orientation component of the score will be 0.05 if the difference between the desired and the obtained orientation, expressed as quaternions, is 0.30; $w = 0.75$ for $3D$ tasks and 1 otherwise. Note that the metric ranges in $(0, 1]$, and the two weighted components (position and orientation) also range in $(0, 1]$ with 1 being assigned if the objects are exactly in the goal position and orientation, and the score decaying exponentially the further they are from them. All objects in the final configuration were scored, regardless if they had to be moved compared to the initial configuration or not. The score for each challenge was calculated by averaging the score of all goals of that challenge, and a final score was obtained by averaging the scores of all challenges.

Human check for compliance with the rules. Organisers inspected the code of sampled systems submitted during Round 1, and the code of all Round 2 final submissions, to ensure the compliance with the rules and the competition 'spirit' (see Rules below).

Software and material. We provided an *environment*[6] and a *software kit*[7] through *GitHub* repositories. These repositories included the instructions to use the software kit; the software to run the simulation of the robot and environment; the interface to run the intrinsic phase learning and the extrinsic phase scoring. The environment was simulated using *PyBullet* and was created as an *OpenAI Gym* environment. OpenAI Gym environments are a standard in the machine learning community thus making easy to reuse existing algorithm implementations in the competition. The software kit provided an example of how to run the environment and evaluate a system. The environment could be run with the GUI on, showing the environment as in Fig. 1, or with the GUI off, for faster execution. When running the tests in local machines, the participants could also freely set the duration of the intrinsic and the extrinsic phases.

3. Organization of the competition

Protocol and schedule. The participants registered in AIcrowd web-site [8] , downloaded the competition *software kit*, and installed the robot and environment simulator in their machines. The participants managed their submission code as a private Git repository in AIcrowd. The repository had to contain the code of the systems to submit for the official evaluation. To make a submission, the participants created a new *Git tag* and pushed it to the system repository. The automatic evaluator picked up the code, built a docker image of it, and computed the performance score by running the simulation using cloud servers on Google Cloud Platform.

The competition was divided into two rounds: (1) *Round 1*: submissions were evaluated by running only the extrinsic phase while participants run the intrinsic phase on their machines; the top 20 ranking participants could participate to Round 2; (2) *Round 2*: submissions were evaluated by running both the intrinsic and extrinsic phase in AIcrowd; all final submissions were checked for coherence with the competition rules and 'spirit'. The winners were publicly declared during the NeurIPS-2019 conference.

6. https://github.com/AIcrowd/real_robots
7. https://github.com/AIcrowd/neurips_goal_real_robots_starter_kit
8. https://www.aicrowd.com

Beta testing at IMOL-2019 conference. The software kit of the competition was beta tested during a *Summer School on Open-Ended Learning Robots*[9] linked to the *International Conference on Open-Ended Learning (IMOL-2019)*[10].

Rules. The rules explained the competition software kit, phases, Rounds, etc. Rules also contained the 'spirit' of the competition: that is, that the competition is directed to develop systems that have close to zero knowledge on the world when they start the intrinsic phase and so they should autonomously acquire all the knowledge they need to solve the tasks in the extrinsic phase. In line with this 'spirit', a *Golden Rule* was estabilished that forbid the use of the extrinsic-phase scoring function, or variants of it, as a reward function during the intrinsic phase. However, given the difficulty of the competition, in Round 1 we allowed participants to violate in part the competition spirit, except the Golden Rule above, by using hardwired or pre-trained models for recognising the identity and position of objects. Round 2 allowed no exception instead, so the systems had to learn everything from scratch during the intrinsic phase.

4. Results

Final ranking. During the competition, 171 users subscribed to REAL on AICrowd, with 78 submissions in Round 1 and 46 submissions in Round 2. The submissions in Round 2 were from 6 participants involving both individual researchers and research groups. The final top scores were all very similar (Table 1).

Table 1: The participants of the competition and their final scores.

Final Submissions	**Score**	**2D**	**2.5D**	**3D**
mrrobot	0.235	0.352	0.330	0.021
tky	0.235	0.352	0.330	0.021
CIIRC-Incognite	0.235	0.352	0.330	0.021
isi	0.220	0.330	0.307	0.024
AutoLearningMPI	0.219	0.329	0.307	0.021
bryan_lincoln	0.208	0.310	0.287	0.025

With a maximum score of 1.0, the score of 0.235 (3 top ranking teams) was quite low and could be achieved by not moving any object. Indeed, many of the overall goals had 1 or 2 objects already in the desired place, so trying to adjust the remaining ones risked to impair their correct position by accidentally hitting them. As a result, the submissions where the robot moved the objects achieved a lower score. Achieving a score higher than 0.235 required to consistently improve the position of objects out of place while not deteriorating the position of objects already in place, a result not achieved by any submission of any team.

9. http://www.goal-robots.eu/announcements/events/summer-school-2019-2
10. http://www.imol-conf.org

What we learned. The major thing we learned from the competition is that it involves many challenges at the core of machine learning, and that facing them at the same time made the competition very hard. Here we illustrate a few of those challenges, some solutions that were attempted by participants, and other solutions that could be explored in the future. *Exploration: touching the objects.* A major challenge is that random exploration rarely leads to touch objects as most of the time the robot arm moves far from the table and objects. So, some sort of intrinsic motivation is needed to guide exploration towards the interesting parts of the environment. Some groups developed their own intrinsic motivation mechanism, e.g. the Incognite Group crafted a 'Reward Mixer' that combined motorics, touch, and latent distance to self-generate a reward signal. Another possibility could be to use 'curiosity-driven exploration by self-supervised prediction' (Pathak et al., 2017), where the robot is attracted to things that it cannot initially predict. *Abstraction: object recognition.* To successfully interact with the environment the robot needs a suitable abstraction, here of objects. This can be explicit, based on a separate abstraction module, or implicit, based on end-to-end approaches. In Round 2, we did not allow models where object recognition/localisation was pre-trained, so the robot had to face the intrinsic phase without even knowing that objects existed. One natural solution for explicit abstraction would be to employ approaches, such as 'variational autoencoders' (Kingma and Welling, 2013), allowing the abstraction of objects as latent variables. However, to be trained such approaches need a variety of images of the object in different positions, but to generate this variety the robot needs to be able to move the objects, which in turn needs to abstract them. This boot-strapping problem is one of the major challenges of the competition. Also, when the arm moves an object it covers a large part of the image, thus impairing the abstraction of objects. *Training control: no rewards, no goals.* A core feature of the challenge is that in the intrinsic phase the agent is not given any reward or goal, thus it has to self-generate tasks autonomously. Some approaches that could be used to this purpose are as follows: 'GRAIL – Goal-discovering Robotic Architecture for Intrinsically-motivated Learning' (Santucci et al., 2016), that self-generates goals/tasks when it succeeds to cause a novel change in the environment; 'Visual Reinforcement Learning with Imagined Goals' (Nair et al., 2018), that generates goals with an autoencoder; 'AGME - Autonomous Goal Manifold Exploration' (Cartoni and Baldassarre, 2018), using a non-parametric method to progressively discover the goal sub-space, and 'Automatic Goal Generation for Reinforcement Learning Agents' (Florensa et al., 2018), using a notion of 'Goal of Intermediate Difficulty', i.e. goals that are not too hard or too easy to learn. *Interdependent challenges.* All these challenges, plus others not reported here for brevity, are interdependent. For example, learning to move an object requires to know where it is, but learning to know where it is requires to move it. More in general, exploration, abstraction, and control have to be acquired at the same time as any of them cannot succeed without the others. A related issue is that control might take as input the abstract representations and these change with the progression of learning: this creates a non-stationary problem for control. *Architectures.* Facing the above problems possibly requires rather complex architectures formed by multiple modules, and coordinating these modules is in itself a difficult engineering challenge. These challenges were also exacerbated by the lack of a baseline model to start from, and by the score that, as we saw above, punished initial success. This impaired the participant's motivation and decreased the number of entrants that participated to Round 2.

5. Next steps: REAL-2020

We are now preparing the second edition of REAL. Based on what we learned from the first edition, we will lower the competition difficulty in multiple ways by: (a) introducing easier classes of goals, e.g. involving single objects; (b) furnishing parameterised motor actions to facilitate the contact with objects; (c) elaborating a metric rewarding initial performance; (d) furnishing an architecture blueprint to organise the multiple components needed to solve the challenge; (e) allowing the robot controller to decide online when to acquire the visual input, thus having much faster simulations. We will also furnish a baseline model based on which the participants could develop their own systems. These new features will allow an easier 'take-off' for the participants, even the less skilled ones, and so will allow the community to fully appreciate the fascination and challenges posed by open-ended learning.

6. Conclusions

The first edition of the REAL competition, REAL-2019, demonstrated that autonomous open-ended learning in machines and robots is still unsolved. It is a fundamental research area that it involves central challenges for artificial intelligence, such as exploration, sparse rewards, object learning, generalisation, catastrophic interference, and autonomous skill learning. Open-ended learning has also a great potential for applications involving unstructured environments posing problems unforeseeable at design time.

The competition is creating a standard benchmark for open-ended learning, useful to foster the production and comparison of new algorithms and architectures. In this respect, the core of the proposed benchmark is its structure based on the intrinsic and extrinsic phases. Indeed, while in real organisms and future robots intrinsic and extrinsic learning are and should be intermixed, the two-fold structure of the competition allows the use of the extrinsic phase to rigorously measure the knowledge acquired in the intrinsic phase on the basis of autonomous intrinsically-motivated open-ended learning.

REAL-2019 proved to be very hard, beyond the state of the art. The reason is that open-ended learning requires robots to face a number of key machine learning challenges at the same time – a difficult bootstrapping problem. Despite, but also for, these challenges, the competition received a lot of positive comments during its implementation and the NeurIPS-2019 conference. One of the NeurIPS reviewers of the competition proposal observed: "It may take an initial year of the competition for both the organizers and participants to hone this competition into a truly exciting second competition. And that is OK." Encouraged by this interest, and motivated by the fascinating challenges posed by open-ended learning, we are thus working to prepare an exciting, enhanced *Second REAL-2020 Competition*.

Acknowledgments

This work was co-funded by the European Union, Horizon 2020 Research and Innovation Program, G.A. no. 713010, Project 'GOAL-Robots – Goal-based Open-ended Autonomous Learning Robots' [11]. The competition was also supported free of charge by *AIcrowd* [12]: we

11. http://www.goal-robots.eu and https://cordis.europa.eu/project/id/713010
12. https://www.aicrowd.com/challenges/neurips-2019-robot-open-ended-autonomous-learning

thank Sharada P. Mohanty for assistance. The competition was supported by the program *Google Cloud Platform Credits for Education*[13]: we thank Manuel del Verme and Sylvain Gelly for guidance on this. We thank Svenja Stark who contributed to setup the competition rules, and Andrew Barto and Bruno Castro da Silva who contributed to elaborate the competition core ideas.

References

J. Achiam, H. Edwards, D. Amodei, and P. Abbeel. Variational option discovery algorithms. *arXiv*, 1807.10299, 2018.

G. Baldassarre. What are intrinsic motivations? A biological perspective. In *Proceedings of the International Conference on Development and Learning (ICDL-2011)*. 2011.

G. Baldassarre and M. Mirolli, editors. *Intrinsically Motivated Learning in Natural and Artificial Systems*. Springer & Verlag, Berlin, 2013.

A.G. Barto and S. Mahadevan. Recent advances in hierarchical reinforcement learning. *Discrete event dynamic systems*, 13(1-2):41–77, 2003.

A.G. Barto, S. Singh, and N. Chentanez. Intrinsically motivated learning of hierarchical collections of skills. In *ICDL-2004*, 2004.

M. Bellemare, S. Srinivasan, G. Ostrovski, T. Schaul, D. Saxton, and R. Munos. Unifying count-based exploration and intrinsic motivation. In *NIPS*, pages 1471–1479, 2016.

D. E. Berlyne. *Conflict, arousal, and curiosity.* McGraw-Hill Book Company, 1960.

A. Cangelosi and M. Schlesinger. *Developmental robotics*. MIT Press, 2015.

E. Cartoni and G. Baldassarre. Autonomous discovery of the goal space to learn a parameterized skill. *arXiv*, 1805.07547, 2018.

B. Eysenbach, A. Gupta, J. Ibarz, and S. Levine. Diversity is all you need: Learning skills without a reward function. *arXiv*, 1802.06070, 2018.

C. Florensa, D. Held, X. Geng, and P. Abbeel. Automatic goal generation for reinforcement learning agents. In *ICML-2018*, 2018.

S. Forestier, Y. Mollard, and P. Oudeyer. Intrinsically motivated goal exploration processes with automatic curriculum learning. *arXiv*, 1708.02190, 2017.

J. Gottlieb, P. Oudeyer, M. Lopes, and A. Baranes. Information-seeking, curiosity, and attention. *Trends in Cognitive Science*, 17(11):585–593, 2013.

D. Held, X. Geng, C. Florensa, and P. Abbeel. Automatic goal generation for reinforcement learning agents. *arXiv*, 1705.06366, 2017.

D.P. Kingma and M. Welling. Auto-encoding variational Bayes. *arXiv*, 1312.6114v10, 2013.

13. https://edu.google.com/programs/credits/research/

T.D. Kulkarni, K. Narasimhan, A. Saeedi, and J. Tenenbaum. Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation. In *NIPS*, pages 3675–3683, 2016.

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

M. C. Machado, M. G. Bellemare, and M. Bowling. A laplacian framework for option discovery in reinforcement learning. *arXiv*, 1703.00956, 2017.

L. Meeden and D. Blank. Developing grounded goals through instant replay learning. In *ICDL-2017*, 2017.

A. Nair, V. Pong, M. Dalal, S. Bahl, Steven Lin, and S. Levine. Visual reinforcement learning with imagined goals. In *LLRLA2018 (at FAIM2018)*, 2018.

P. Oudeyer, F. Kaplan, and V. Hafner. Intrinsic motivation systems for autonomous mental development. *IEEE transactions on evolutionary computation*, 11(6), 2007.

D. Pathak, P. Agrawal, A. A. Efros, and T. Darrell. Curiosity-driven exploration by self-supervised prediction. In *CVPRW-2017*, 2017.

M. Rolf and M. Asada. Autonomous development of goals: From generic rewards to goal and self detection. In *ICDL-2014*, 2014.

V. G. Santucci, G. Baldassarre, and M. Mirolli. Cumulative learning through intrinsic reinforcements. In *Evolution, Complexity and Artificial Life*. Springer, 2014a.

V. G Santucci, G. Baldassarre, and M. Mirolli. Autonomous selection of the "what" and the "how" of learning: an intrinsically motivated system tested with a two armed robot. In *ICDL-2014*, 2014b.

V. G. Santucci, G. Baldassarre, and M. Mirolli. GRAIL: A goal-discovering robotic architecture for intrinsically-motivated learning. *IEEE Transactions on Cognitive and Developmental Systems*, 8(3):214–231, 2016.

V.G. Santucci, G. Baldassarre, and M. Mirolli. Intrinsic motivation signals for driving the acquisition of multiple tasks: a simulated robotic study. In *ICCM-2013*, 2013.

M. Schembri, M. Mirolli, and G. Baldassarre. Evolving internal reinforcers for an intrinsically motivated reinforcement-learning robot. In *ICDL-2007*. 2007.

J. Schmidhuber. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, 2(3):230–247, 2010.

K. Seepanomwan, V.G. Santucci, and G. Baldassarre. Intrinsically motivated discovered outcomes boost user's goals achievement in a humanoid robot. In *ICDL-2017*, 2017.

V. Sperati and G. Baldassarre. Bio-inspired model learning visual goals and attention skills through contingencies and intrinsic motivations. *IEEE Transactions on Cognitive and Developmental Systems*, 10(2):326–344, 2018.

D. Tanneberg, J. Peters, and E. Rueckert. Intrinsic motivation and mental replay enable efficient online adaptation in stochastic networks. *Neural Networks*, 109:67–80, 2019.