Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Multi-Agent System Cooperation Via Constructivst Learning Approach

Ahmed Mokhtari[1] , Youcef Dahmani[2] , Adel Tomouh[3] , Mokhtaria Boubabouri
[1]department of Computer Science, Sidi-bel-Abbes university , Algeria
[1]ahmed.mokhtari@univ-sba.dz
[2]department of Computer Science, Tiaret university , Algeria
[2]dahmani_y@yahoo.fr
[3]department of Computer Science, Sidi-bel-Abbes university , Algeria
[3]toumouh@gmail.com
[4]department of Computer Science, Tiaret university , Algeria
[4]hadjerboubabouri@gmail.com

ARTICLEINFO

Received: 14 Dec 2024

Revised: 16 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

Multi-agent systems have a wide range of real-world applications in many fields, where multiple agents must cooperate to achieve their global objectives in a shared environment.

Cooperative learning in multi-agent systems is a fundamental field within artificial intelligence. It aims to design autonomous agents capable of cooperation by learning and adapting to dynamic environments. As these environments become more complex, agents require learning strategies that allow them not only to react but also to evolve and build their own behavioral models to solve tasks collaboratively. Constructivist approaches consider agents as active entities capable of constructing their internal knowledge through experience and intrinsic motivation, without relying on predefined behaviors or external rewards. In this work, we propose a model that integrates constructivist learning concepts like schema and intrinsic motivation in cooperative multi-agent systems to enable agents to progressively build, refine behavioral schemas and to learn collaboration behaviors without external supervision.Furthermore, we evaluate the proposed architecture for solving the collaborative resources extraction problem in a grid environment simulation. The result shows that the agents are able to learn how to navigate and how to collaborate to extract resources and avoid obstacles.The emergence of collaborative and navigational behavior through constructivist learning and intrinsic motivation mechanisms can lead to the development of autonomous, self-motivated agents capable of cooperation.

Keywords: machine learning, cooperative learning, multi-agent system, constructivist learning, collaboration, intrinsic motivation.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

INTRODUCTION

Multi-agent systems (MAS) have a wide range of real-world applications in many fields, such as robot control [1], precision agriculture [2], traffic signal control [3], underwater exploration [4], power dispatch [5], finance [6] and robotics-based rescue tasks [7].

A multi-agent system is made up of several entities called agents located in an environment, designed to solve tasks collectively. Cooperation between agents is necessary to achieve system goals. To achieve cooperative behavior, there are two approaches, static and dynamic. In the static case, the approach consists of hard pre-programming the agents so that in a given situation they perform the appropriate pre-defined actions, but this solution has the disadvantage that agents faced with a new situation fail to adapt well.

Multi-agent learning (MAL), which is part of the dynamic approach, is a sub-domain of machine learning in which researchers are interested in developing self-learning agents capable of learning by interacting with their environment and with other agents.

MAL is an interdisciplinary field that intersects with many other domains, including evolutionary biology [8], [9] and economics [10], [11]. Social concepts like communication are important during the process of learning [12].

Among multi-agent learning solutions, we find "multi-agent reinforcement learning" (MARL) [r13}, which consists of using reinforcement learning in a collective context. Despite its success in several applications [14], [15], [16], it suffers from scalability problems (the state space in this type of learning is so large that it cannot be modeled a priori) and problems of adaptation to new situations.

Knowing that the agents are located in complex environments, where changes can happen all the time, we defend the hypothesis that endowing a multi-agent system with a constructivist learning mechanism and intrinsic motivation gives its agents a capacity for adaptation, enabling them to learn how to collaborate.

The rest of the paper consists of the following structure: the related works are presented in Section 2, the constructivist approach is presented in section 3,the proposed model is introduced in Section 4, experimental results and discussions are detailed in Section 5 and 6, and finally conclusions in Section 7.

RELATED WORK

Multi-agent reinforcement learning (MARL) remains at the center of MAS learning. Agents learn by interacting with their environments based on scalar rewards to optimize certain policies. Some MARL methods include Independent Q-Learning, where agents do not know the other agents' policies, leading to non-stationarity [17], and MAD-DPG with full information during training and local execution [18].

On the other side, several works have implemented constructivist learning in a single agent. In [19], the BEL-CA model, based on the idea that learning emerges from interaction, without explicit supervision, was presented.

CCA is an extension of BEL-CA that introduces self-motivation mechanisms to enhance exploration through autotelic curiosity that drives the agent to seek novel interactions. This addition reduces the time spent in inefficient behaviors, guiding agents more directly toward meaningful schema development [20].

The Intrinsically Motivated Schema mechanism proposes a hierarchical approach where

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Agents construct complex behaviors guided by internal motivational signals [21]. Replacing external reward signals (as in traditional reinforcement learning) with proclivity values (internal satisfaction signals) allows agents to behave in a goal-directed manner without requiring extrinsic reinforcement.

The authors of [22] present a constructivist approach to state space adaptation applied to reinforcement learning. They used the Multi-Layer Growing Neural Gas (ML-GNG) clustering algorithm; their results show that the agent was able to adapt by learning suitable state spaces.

THE CONSTRUCTIVIST LEARNING APPROACH

Intelligent systems are evolutionary and adaptive because they are endowed with learning mechanisms that enable them to learn by observing and interacting with their environments while constructing a set of internal representations.

One approach to learning is constructivist learning. It is mainly inspired by developmental theory in cognitive psychology, based on the work of Piaget [23], [24].

Internal representations, called schemas in Piaget's work [25], represent the basic units that make up a cognitive system and, at the same time, the dynamics that govern it.

This type of learning has certain characteristics that distinguish it from other approaches to learning.

firstly, learning is considered to be a progressive, incremental process that passes through levels of increasing complexity and difficulty. At each level, the agent learns to construct knowledge in the form of representations that will be used in successive levels to obtain and construct more complex and abstract representations.

Secondly, learning is an intrinsic process, i.e., motivations from within the system that drive the agent to learn and develop these skills, without necessarily having an external motivation (in the form of a reward) [26].

METHODS

1. Proposed model

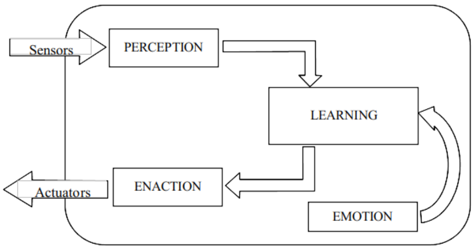
In this section, we introduce our model based on a constructivist learning approach. We start by defining the notion of the schema and its structure, then we detail the architecture of the agents.

In constructivist literature, the schema is considered the fundamental unit or the building block that constructs and forms the cognitive system. In our model, the schema is of the form <context,action,emotion>, where context represents the context in which the agent is situated, action is the action to be enacted in such context, and emotion is the new emotional state of the agent after enacting the action. We associate a numerical value with each schema that represents its valence (satisfaction) in the learning phase.

Our multi-agent system is composed of n homogeneous agents interacting with the environment and with each other. The proposed agent architecture consists of the following modules: perception, learning, enaction, and emotion (figure 1).

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article



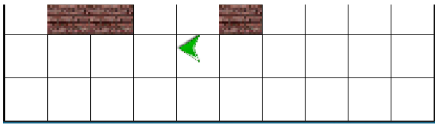


Figure 1. the proposed architecture of an agent

the perception module is responsible for collecting data regarding the state of the environment as well as the internal state of the agent, which is called a context. A context is formed by the set of percepts in the form of a vector <$p_0,p_1,p_2,\dots,p_n$> , where p_i is a percept.

the Enaction is the part responsible for enacting the actions of the agent in the environment.

Emotional states are stored in the emotion module, they are transmitted to the learning module to calculate the valences during the phase of updating the constructed schemas.

Table 1 contains the different primitive actions and associated emotions.

Action	Inaction	Description	Emotion state	Valence
Move forward	succes	Move to empty cell	happy	5
	failure	Bump obstacle	sad	-5
Turn left	succes	Turn left to empty cell	happy	5
	failure	Bump obstacle	sad	-5
Turn right	succes	Turn right to empty cell	happy	5
	failure	Bump obstacle	sad	-5
Wait	succes	Wait close to a resource	excited	10
	failure	Wait far from a resource	boring	0
Move toward	succes	Get close a resource	happy	20
	failure	Bump obstacle	boring	0
Extract	succes	Extract a resource	Very excited	50
	failure	Extraction fails	Very disoppointing	-10

Table 1. primitive actions and emotions

The last module to describe is the learning module. In this module, the agent's behavior is controlled through the construction of schemas and their enactions in the environment according to different contexts.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Algorithm 1. Schema selection algorithm

The learning proceeds as follows (see Algorithm 1):

1. The agent constructs the context vector from the different perceptions.
2. A context-based selection mechanism generates a subset of schemas called candidate schemas.
3. A second selection based on valences to choose the schema to be enacted will take place.
4. The chosen schema is sent to the enaction module, which executes the action part of the schema.
5. An update of the schemas using the intrinsic feedback from the emotion module.

2. Experience

To test the model, we use collective resource extraction as testbed, assuming that resource extraction requires the collaboration of two or more agents. The agents must learn how to collaborate to extract resources .

In this testbed, the environment in which the agents are situated is an NxN cell grid that contains obstacles in brown and resources are represented as red box (see figure 2).

Algorithm 1 schema selection algorithm

```
candidates_schemas ← φ
if mode is exploitation then
    for each sch ∈ Schemas do
        if sch.context = context then
            candidates_schemas.add(sch)
        end if
    end for
    if candidates_schemas ≠ φ then

schema = Max_valance(candidates_schemas)
    else
        schema = generate_new_schema()
    end if
else if mode is exploration then
    schema = generate_new_schema()
end if
```

Figure 2. The collective resource extraction environment

(agents in green and resource in red)

The agent is equipped with sensors to detect obstacles, resources, and other agents, and actuators that allow it to move in different directions (forward, left, right, backward) to avoid obstacles and extract resources.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

The agent is also equipped with an emotion management module in which its emotional state changes after the enaction of various actions in the environment (see Table 1).

For example, if the agent fails to enact the action of moving because it hits an obstacle while trying to move, it feels pain in the emotion module; otherwise, it feels pleasure because it succeeds in performing an action.

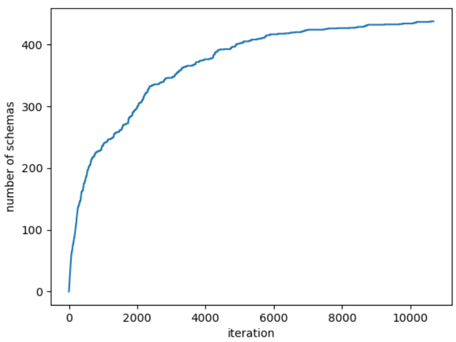
If he attempts to extract a resource and there is no other agent to help him, the enaction of the extraction action fails, and he feels very unhappy; however, if there is another agent, he succeeds in the enaction of the extraction of a resource and feels very happy.

The system is composed of two agents who learn a set of schemas without any prior knowledge about the structure of the environment or the concept of collaboration.

They must learn navigation schemas and cooperation schemas to accomplish the task of collaborative resource extraction.

RESULTS

In this section, we present results obtained by simulation of our model. Figure 3. illustrates the growth of the number of constructed schemas; we observe an exponential growth and then a convergence of the number of the stable schemes to 423 schemas after 10000 iterations.



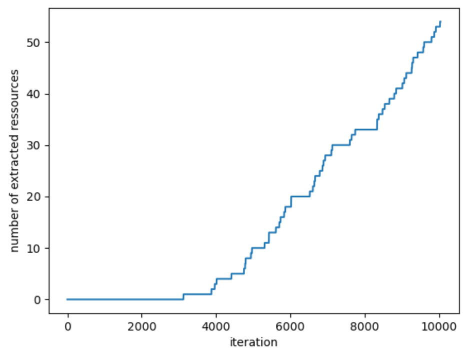


Figure 3. Number of constructed schemas
resources

Figure 4. Number of extracted

Figure 4. shows that agents needed an 3200 iterations to successfully extract the first resource, and then they started to extract more resources in a shorter time, more then 50 object in the remaining 6800.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

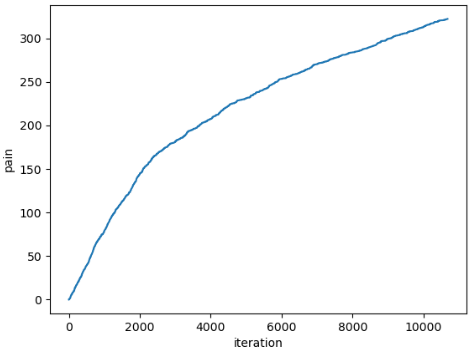


Figure 5. the curve of the pain

In Figure 5, we also see that the pain plot converges to a certain value (350) after that the agents learn to avoid obstacles and walls.

from Table 2, we cite the most relevant schemas after 10000 iterations. for example, the schema [<no_obstacle, no_resource,no_agent>,move_forward,happy] expresses a learned construction that helps the agent during navigation when there is not obstacles in his way, whereas the schema [obstacle_left, no_resource, no_agent>, turn_right, happy] and the schema [<obstacle_right,no_resource,no_agent>,turn_left, happy] expresses the common knowledge that when facing an obstacle, the agent has to change the direction.

N°	context	action	Emotion	Valence
1	<no_obstacle,no_resource,no_agent>	move_forward	happy	7460
2	<obstacle_left,no_resource,no_agent>	turn_right	happy	175
3	<obstacle_left,no_resource,no_agent>	turn_left	sad	-23
4	<obstacle_right,no_resource,no_agent>	turn_left	happy	750
5	<obstacle_right,no_resource,no_agent>	turn_right	sad	-19
6	<no_obstacle,resource_close,no_agent>	wait	exciting	7960
7	<no_obstacle, resource _left,no_agent>	move-toward	happy	2700
8	<no_obstacle,resource_right,no_agent>	move-toward	happy	8130
9	<no_obstacle,resource_close,no_agent>	extract	very disoppointing	-25
10	<no_obstacle, resource _close,agent>	extract	very exciting	1400

Table 2. snapshot of some constructed schemas

There are also the two schemas, [<no_obstacle,resource_close,no_agent>,wait, exciting] and [<no_obstacle, resource_close,agent>,extract,very_exciting], which are schemas that allow an agent to collaborate, meaning they are activated when the agent encounters a resource. The first schema expresses the notion of waiting next to a resource for other agents, and the second is used in the context of collective extraction.

Those results show the emergence of a collaborative and navigational behavior through constructivist learning and intrinsic motivation mechanisms.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

DISCUSSION

This section focuses on understanding why certain behaviors emerged, what they imply about the cognitive dynamics of each agent, and how they reinforce our central hypothesis: endowing a multi-agent system with a constructivist learning mechanism and intrinsic motivation gives its agents a capacity for adaptation, allowing them to learn how to collaborate.

The first observation is that the agents consistently construct more schemas, and they do it at a faster rate until the number of schemas stabilizes, which means that our agents start by exploring the environment (greater exploration exposes the agent to more diverse initial contexts/actions combinations); once this phase finishes, the number of schemas stabilizes and remains almost constant.

The convergence of the pain curve means the agents learn the structure of the environment, and that allows them to avoid obstacles and walls, which leads to less bumping and so less pain.

The extraction of the first resource by the collaboration of the agents took much time (3200 iterations) because they had to learn the action of waiting for the help of the other agents and also learn to enact the action of extraction simultaneously by both of the agents. In the beginning of the simulation, even if the first agent learns to find a resource and wait for help, the second agent has not yet learned the appropriate schemas.

Once the agents learn to construct the navigation schemas and collaborative schemas, and they learn to sequence behaviors, for example [move toward a resource, wait for help, extract], the collaboration of the agents takes less time, and more resources are easily extracted collaboratively. This is a sign of emerging collaborative complex behaviors. The agents are not just reacting; they are showing planning capability.

The simulation results strongly support the hypothesis that constructivist learning and intrinsic motivation in a cooperative multi-agent system enhance the learning capacity, behavioral flexibility, and long-term performance of agents in collaborative tasks.

CONCLUSION

In this article, we designed and tested a cooperative multi-agent system based on a vision deeply inspired by human cognitive development. By integrating Jean Piaget's constructivist learning principles and intrinsic motivation mechanisms, we demonstrated that it is possible to create artificial agents capable of:

- constructing their knowledge without supervision,
- self-evaluating through the satisfaction derived from their interactions with the environment and the other agents,
- collaborate to reach a common objectives.

This model breaks with traditional learning approaches strictly guided by external rewards or human demonstrations. Through a simple environment, the agents were able to explore, fail, correct themselves, construct knowledge, and cooperate without a priori knowledge or external signals.

Despite our contribution, we mention some limitations of this work, agents interact socially but without active communication or direct imitation, and also the simulation remains 2D, limiting the diversity of experiences.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Several avenues could extend and enrich this work, for example allowing agents to exchange schemas, or share experiences within a framework inspired by social or cultural learning, and application to real-world domains like robotics-based rescue tasks, and underwater exploration.

REFRENCES

[1] Orr, J. & Dutta, A. (2023). Multi-agent deep reinforcement learning for multi-robot applications: A survey. Sensors, Vol. 23, No. 7, pp. 3625. https://doi.org/10.3390/s23073625

[2] Dutta, A., Roy, S., Kreidl, O. P., & Bölöni, L. (2021). Multi-robot information gathering for precision agriculture: Current state, scope, and challenges. Ieee Access, Vol. 9, pp. 161416 161430. https://doi.org/10.1109/access.2021.3130900

[3] Xu, B., Wang, Y., Wang, Z., Jia, H., & Lu, Z. (2021). Hierarchically and cooperatively learning traffic signal control. Proceedings of the AAAI conference on artificial intelligence, volume 35, pp. 669–677. https://doi.org/10.1609/aaai.v35i1.16147

[4] Zhou, Z., Liu, J., & Yu, J. (2021). A survey of underwater multi-robot systems. IEEE/CAA Journal of Automatica Sinica, Vol. 9, No. 1, pp. 1–18. https://doi.org/10.1109/jas.2021.1004269

[5] Wang, J., Xu, W., Gu, Y., Song, W., & Green, T. C. (2021). Multi-agent reinforcement learning for active voltage control on power distribution networks. Advances in neural information processing systems,Vol. 34, pp. 3271–3284.

[6] Fang, Y., Tang, Z., Ren, K., Liu, W., Zhao, L., Bian, J., Li, D., Zhang, W., Yu, Y., & Liu, T.-Y. (2023). Learning multi-agent intention-aware communication for optimal multi-order execution in finance. Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 4003–4012. https://doi.org/10.1145/3580305.3599856

[7] Queralta, J. P., Taipalmaa, J., Pullinen, B. C., Sarker, V. K., Gia, T. N., Tenhunen, H., Gabbouj, M., Raitoharju, J., & Westerlund, T. (2020). Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision. Ieee Access, Vol. 8, pp. 191617–191643. https://doi.org/10.1109/access.2020.3030190

[8] Jaderberg, M., Czarnecki, W. M., Dunning, I., Marris, L., Lever, G., Castaneda, A. G., Beattie, C., Rabinowitz, N. C., Morcos, A. S., Ruderman, A., et al. (2019). Human-level performance in 3d multiplayer games with population-based reinforcement learning. Science, Vol. 364, No. 6443, pp. 859–865. https://doi.org/10.1126/science.aau6249

[9] Duéñez-Guzmán, E. A., Sadedin, S., Wang, J. X., McKee, K. R., & Leibo, J. Z. (2023). A social path to human-like artificial intelligence. Nature machine intelligence, Vol. 5, No. 11, pp. 1181–1188. https://doi.org/10.1038/s42256-023-00754-x

[10] Zheng, S., Trott, A., Srinivasa, S., Parkes, D. C., & Socher, R., . The ai economist: Optimal economic policy design via two-level deep reinforcement learning, 2021. URL https://arxiv. org/abs/2108.02755.

[11] Johanson, M. B., Hughes, E., Timbers, F., & Leibo, J. Z., . Emergent bartering behaviour in multi-agent reinforcement learning, 2022. URL https://arxiv. org/abs/2205.06760.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[12] Hertz, U., Koster, R., Janssen, M., & Leibo, J. Z. (2023). Beyond the matrix: Experimental approaches to studying social-ecological systems. Technical report, Center for Open Science. https://doi.org /10.1016/j.cognition.2024.105993

[13] Zhang, K., Yang, Z., & Başar, T. (2021). Multi-agent reinforcement learning: A selective overview of theories and algorithms. Handbook of reinforcement learning and control, pp. 321–384. https://doi.org/ 10.1007/978-3-030-60990-0_12

[14] Lv, Z., Xiao, L., Du, Y., Niu, G., Xing, C., & Xu, W. (2023). Multi-agent reinforcement learning based uav swarm communications against jamming. IEEE Transactions on Wireless Communications, Vol. 22, No. 12, pp. 9063–9075. https://doi.org/10.1109/twc.2023.3268082

[15] Qie, H., Shi, D., Shen, T., Xu, X., Li, Y., & Wang, L. (2019). Joint optimization of multi-uav target assignment and path planning based on multi-agent reinforcement learning. IEEE access, Vol. 7, pp. 146264–146272. https://doi.org/10.1109/access.2019.2943253

[16] Cui, J., Liu, Y., & Nallanathan, A. (2019). Multi-agent reinforcement learning-based resource allocation for uav networks. IEEE Transactions on Wireless Communications, Vol. 19, No. 2, pp. 729–743. https://doi.org/10.1109/twc.2019.2935201

[17] Claus, C. & Boutilier, C. (1998). The dynamics of reinforcement learning in cooperative multiagent systems. AAAI/IAAI, Vol. 1998, No. 746-752, pp. 2.

[18] Lowe, R., Wu, Y. I., Tamar, A., Harb, J., Pieter Abbeel, O., & Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. Advances in neural information processing systems, Vol. 30.

[19] Xue, J., Georgeon, O. L., & Hassas, S. (2020). A constructivist approach and tool for autonomous agent bottom-up sequential learning. International Journal of Educational and Pedagogical Sciences, Vol. 14, No. 10, pp. 886–894.

[20] Xue, J. (2020). Architecture cognitive constructiviste: un modèle pour concevoir un agent automotivé capable de faire du sens et de construire des connaissances de l'environnement. Ph.D. thesis, Université de Lyon.

[21] Corbacho, F. J. (2019). Towards self-constructive artificial intelligence: Algorithmic basis (part i).arXiv preprint arXiv:1901.01989.

[22] Guériau, M., Cardozo, N., & Dusparic, I. (2019). Constructivist approach to state space adaptation in reinforcement learning. 2019 IEEE 13th International Conference on Self-Adaptive and Self-Organizing Systems (SASO), IEEE, pp. 52–61. https://doi.org/10.1109/saso.2019.00016

[23] Piaget, J. (2013). The construction of reality in the child. Routledge. https:// doi.org/10.4324/ 9781315009650

[24] Huitt, W. & Hummel, J. (2003). Piaget's theory of cognitive development. Educational psychology interactive, Vol. 3, No. 2, pp. 1–5.

[25] Drescher, G. L. (1991). Made-up minds:a constructivist approach to artificial intelligence. MIT press.https://doi.org/10.7551/mitpress/4378.001.0001

[26] Haber, N., Mrowca, D., Fei-Fei, L., & Yamins, D. L. (2018). Emergence of structured behaviors from curiosity-based intrinsic motivation. arXiv preprint arXiv:1802.07461.