Successive Refinement in Continual Learning: A Study on Spatial Representations

Nicola Catenacci Volpi[1], Hippolyte Charvin[1] and Daniel Polani[1*]

Humans' capacity for perpetual learning and adjustment in response to novel circumstances throughout their lifespan is exceptional. This cognitive aptitude, known as Continual Learning (CL), involves the ongoing and progressive refinement of increasingly complex behaviors. It entails the construction of advanced behavioral patterns that build upon pre-existing ones, allowing for the utilization, adaptation, and generalization of previously acquired skills when confronted with new situations [8]. For these reasons, open-ended learning systems can benefit from CL to facilitate their progressive assimilation and integration of novel knowledge and skills over extended periods of time, and ensuring the retention of formerly acquired information while concurrently exploring.

While CL is a distinctive trait of human intelligence, it poses a substantial challenge within the realm of Artificial Intelligence and, in particular, Reinforcement Learning (RL) [9], because most of the employed strategies revolve around an agent that focuses on acquiring expertise in a specific and limited task [5]. Achieving success in CL necessitates the constant incorporation of novel information while preserving valuable past knowledge. Present-day incremental machine learning methods have been unable to master this ability due to their lack of careful management of the data they acquire, retain, or discard. The obstacles of implasticity (the failure to incorporate useful new information) and catastrophic forgetting (the unintended loss of valuable information from memory) are widely acknowledged as significant challenges [6]. A widely used measure to assess CL agents evaluates their predictive accuracy on tasks encountered in the past, examining how effectively an agent preserves previously learned information [10]. In the case of embodied agents, their ability to store and process this past information is further restricted by their finite available computational resources [7]. Furthermore, one of the primary hurdles encountered by continual RL agents is the ability to extract and compress pertinent information from an immense stream of sensory data. Hence, for effective CL, a trade-off must be achieved between the agent's computational limitations and its representations fidelity of the increasing volume of incoming data.

In this paper, to start investigating how increasingly complex models are incrementally obtained in CL, we specifically focus on the question of how an agent's spatial representations are refined as additional informational resources are made available. To this aim, we use the information-theoretical notion of *Successive Refinement (SR)* [4], which we applied to a geometric instance of rate-distortion theory [2] called *Geometric Rate-Distortion (GRD)* [1]. Let us consider an embodied agent with limited informational resources that needs to approximately represent its location in a space S (e.g., a grid world-like environment). GRD optimally solves the trade-off between minimizing the amount of Shannon information used to describe locations in the space $\mathcal{S}$ and decreasing the location error arising from the resulting approximate spatial representation. In particular, given two 2D positions $a = (a_x, a_y) \in \mathcal{S}$ and $b = (b_x, b_y) \in \mathcal{S}$, the distortion function $d_{L1}(a, b) = |a_x - b_x| + |a_y - b_y|$ measures the L1 distance between two locations in the grid. In the GRD framework, this distance also quantifies the cost of representing the location a with the location b. Given the distortion d_{L1}, GRD constructs the most compressed spatial representation of the agent location $S1^*$ with an average error from its true position no larger than D. The minimum average number of bits necessary to represent the agent location in $S1^*$, and incurring an average distortion no larger than D, is called the rate-distortion function $R(D)$. Let $p(s_1^*|s)$ be the conditional probability of representing the location $s \in \mathcal{S}$ with the codeword s_1^*. Formally, $R(D)$ is defined as the solution of the following constrained optimization

[*1] The authors are with the Adaptive Systems Research Group, School of Physics, Engineering, and Computer Science, University of Hertfordshire, Hatfield, AL10 9AB, UK. E-mail corresponding author: n.catenacci-volpi@herts.ac.uk

problem

$$R(D) = \min_{p(s_1^*|s):\mathbb{E}[d_{L1}(S,S_1^*)]\leq D} I(S; S_1^*) \tag{1}$$

where $\mathbb{E}[d_{L1}(S, S_1^*)] = \sum_{s,s_1^* \in \mathcal{S}} p(s)p(s_1^*|s)d_{L1}(s, s_1^*)$ and I denotes the mutual information.

Let's imagine the agent starting by operating with a GDR coarse representation S_1^* of its environment at rate R_1, resulting in an average distortion of D_1. Let us assume that the situation evolves and the agent now aims to improve the previous representation of the space towards a finer description with rate R_2 bits (i.e., $R_2 > R_1$). According to the aforementioned CL criteria, instead of creating a new description S_2^* from the ground up, SR takes advantage of the already existing coarser representation S_1^* refining it and combining it with an "addendum" of information ΔS^* with rate $\Delta R = R_2 - R_1$ bits. The two combined representations attain a total distortion of $\hat{D}_2$. In line with the spirit of life-long adaptation, our ultimate goal is to maintain the same notion of optimality for the refined description acquired in two incremental steps as if this was obtained in only one step yielding directly the finer representation S_2^* with GRD at rate R_2. We say that the original space S is successively refinable if the following two conditions are satisfied: $R_1 = R(D_1)$ and $R_2 = R(\hat{D}_2)$. Hence, the first coarse representation and the refined one both achieve the best rates for their respective distortions D_1 and $\hat{D}_2$. In other words, using SR, agents can switch from a coarse spatial representation to a more sophisticated one utilising minimum informational effort and no information needs to be discarded for adapting the representation's granularity. Starting from the Markovian characterization $S \to S_2^* \to S_1^*$ of SR [4], in [1] we provided an equivalent information-theoretic characterization to determine whether SR is achievable, which is the case if $I(S; S_1^*|S_2^*) = 0$. Given that the relationship among the representation S, S_1^* and S_2^* is not generally fully Markovian, the latter characterization allows, according to our knowledge for the first time, to introduce a "relaxed" version of successive refinement when $I(S; S_1^*|S_2^*) << 1$, which in this case quantifies the loss of information optimality induced by the two-stage refinement. Here, by *relaxed successive refinement* we mean that rather than strictly insisting on the condition $R_2 = R(\hat{D}_2)$, we also allow for $R(\hat{D}_2) \gtrsim R_2$, meaning refinements that are slightly suboptimal. We have observed that in the case of the GRD setting described here, for certain pairs of rates R_1 and R_2 the spatial representations can be optimally refined in a successive manner, while for other pairs the representations can still be nearly optimally refined in a relaxed manner and characterized by the aforementioned "soft" form of Markovianity. In simpler terms, if minor compromises are allowed, a fairly effective cascade of successive refinements, at least with one intermediary step, can be realized.

While we are confident that the SR of GRD could pave the way for innovative and principled information-theoretic approaches to assimilating new information within existing knowledge, there are still hurdles to overcome for its integration in continual Reinforcement Learning (RL). First,in order to adapt our formalism to the full RL framework, the actions that were only implicitly assumed in GRD as the operations enabling agents to directly transition from one state to another, should be explicitly introduced. Beyond adding an action space, also incorporating stochastic transitions would facilitate the integration with the RL framework, surpassing the successor representation [3] lying behind the GRD formalism. Then, while here we have illustrated SR with only two stages of refinements, the introduced methodology could intuitively be expanded to chains of multiple refinements, more in line with the spirit of CL. In upcoming research, we aim to utilize the tree structure inherent in the successive refinement method to yield a cascade of spatial refinements connected by Markovian or near-Markovian links. Consequently, when sequences of refinements are implemented, the outlined formulation could potentially be expanded towards a hierarchical model of refinement. We believe that implementing the SR formalism within the RL framework, and its extension towards multi-stage information processing, could offer a fresh viewpoint on hierarchical planning and learning, where the information cost of shifting from an abstract representation to a more detailed one could be minimized.

References

[1] Nicola Catenacci Volpi and Daniel Polani. Space emerges from what we know—spatial categorisations induced by information constraints. *Entropy*, 22(10):1179, 2020.

[2] Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.

[3] Peter Dayan. Improving generalization for temporal difference learning: The successor representation. *Neural computation*, 5(4):613–624, 1993.

[4] William HR Equitz and Thomas M Cover. Successive refinement of information. *IEEE Transactions on Information Theory*, 37(2):269–275, 1991.

[5] Khimya Khetarpal, Matthew Riemer, Irina Rish, and Doina Precup. Towards continual reinforcement learning: A review and perspectives. *Journal of Artificial Intelligence Research*, 75:1401–1476, 2022.

[6] Saurabh Kumar, Henrik Marklund, Ashish Rao, Yifan Zhu, Hong Jun Jeon, Yueyang Liu, and Benjamin Van Roy. Continual learning as computationally constrained reinforcement learning. *arXiv preprint arXiv:2307.04345*, 2023.

[7] Ameya Prabhu, Hasan Abed Al Kader Hammoud, Puneet K Dokania, Philip HS Torr, Ser-Nam Lim, Bernard Ghanem, and Adel Bibi. Computationally budgeted continual learning: What does matter? In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3698–3707, 2023.

[8] Mark B Ring. Child: A first step towards continual learning. *Machine Learning*, 28:77–104, 1997.

[9] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[10] Liyuan Wang, Xingxing Zhang, Hang Su, and Jun Zhu. A comprehensive survey of continual learning: Theory, method and application. *arXiv preprint arXiv:2302.00487*, 2023.