Grounding Artificial Intelligence in the Origins of Human Behavior

Eleni Nisioti
Flowers Team, Inria and Ensta ParisTech
Bordeaux, France
eleni.nisioti@inria.fr

Clément Moulin-Frier
Flowers Team, Inria and Ensta ParisTech
Bordeaux, France
clement.moulin-frier@inria.fr

ABSTRACT
Recent advances in Artificial Intelligence (AI) have revived the quest for agents able to acquire an open-ended repertoire of skills. However, although this ability is fundamentally related to the characteristics of human intelligence, research in this field rarely considers the processes that may have guided the emergence of complex cognitive capacities during the evolution of the species.

Research in Human Behavioral Ecology (HBE) seeks to understand how the behaviors characterizing human nature can be conceived as adaptive responses to major changes in the structure of our ecological niche. In this paper, we propose a framework highlighting the role of environmental complexity in open-ended skill acquisition, grounded in major hypotheses from HBE and recent contributions in Reinforcement learning (RL). We use this framework to highlight fundamental links between the two disciplines, as well as to identify feedback loops that bootstrap ecological complexity and create promising research directions for AI researchers.

2 ECOLOGY GROUNDS OPEN-ENDEDNESS
Despite this progress, research in RL does not often acknowledge that open-ended skill acquisition is fundamentally related to the characteristics of human intelligence [6]. Couldn’t the study of the processes that may have guided the emergence of complex cognitive capacities during the evolution of the human species benefit the study of open-endedness in intelligent agents?

In this work, we take a step back from our computer-scientific lens and turn our attention towards Human Behavioral Ecology (HBE), a field concerned with the effect that environmental conditions have had on the evolution of the human species [10, 50, 64]. Works in this field have studied the formation of cooperative groups [13], the management of shared resources [32], tool use [22], the emergence of communication [31] and culture [69], whether humans hunt in optimal group sizes [66] and how speciation, extinction and dispersal arose in the human history [50].

Admittedly, there are many paths to the acquisition of open-ended skills in AI; grounding our study in human ecology seems to be but one of the options. But there’s a number of reasons that may persuade us to explore it: (i) examining all possible ecologies is infeasible considering our modern and foreseeable computational power [15]; (ii) ecologies that are more familiar to ours make it easier to define evaluation criteria. For example, human-ecology inspired metrics such as equality, sustainability and social welfare have been employed to evaluate agents on their ability to forage [61], find optimal taxation strategies [78] and play games [5]; (iii) Darwinian evolution offers an existence proof for human-like open-ended skill acquisition [15], as well as empirical data and testable hypotheses; (iv) similar attempts at grounding AI research in a non-computational field have already proven to be a fruitful approach. For example, concepts from Development Science such as intrinsic motivation [54, 56] and embodied language acquisition [12] have had a significant impact on modern AI research; (v) the link between HBE and RL has already been recognized in [30], where the transfer of ideas has the opposite direction from the one proposed here. A proposal to study major evolutionary transitions in ecology in order to understand the general laws that underlie innovation and transfer insights to artificial evolution is presented in [67]. Our proposal follows a similar direction but focuses on highlighting the overlap between concepts in RL and HBE. In addition, a subset of the vast range of hypotheses proposed in HBE and aims at mapping them to key research questions in RL in order to identify gaps...
and propose promising research directions. In Section 3, we provide a brief description of current trends in RL research. Section 4 provides a short introduction to the field of HBE. In Section 5, we structure our proposed transfer of ecologically-inspired insights based on three recent key research areas in RL: the adaptability of agents, multi-agent dynamics of groups and their cultural repertoire.

3 ENVIRONMENTAL COMPLEXITY IN RL: CURRENT TRENDS

The interplay between environmental complexity and open-ended skill acquisition in intelligent agents has been investigated from various perspectives. Below, we point to different works in the areas of single-agent, multi-agent and meta RL.

Single-agent settings have focused on two elements of an agent’s learning architecture: (i) the neural networks employed for function approximation, whose generalization abilities are an ongoing debate [49, 52]. Observations in [36] suggest that the ability of a single neural network to generalize emerges in a complex environment, characterized by multi-modal signals situated in temporally and physically rich spaces that allow for diversity in the agent’s perspective; (ii) the cost function or type of intrinsic motivation considered. In [8], useful skills emerge as an agent minimizes future surprise, attempting to counteract the uncertainty inherent in its environment. In curiosity-driven exploration, learning progress generates intrinsic rewards that push an agent to explore and create its own learning curricula [17, 55].

In the multi-agent reinforcement learning (MARL) literature, the automatic discovery of new environments is achieved by multi-agent autocurricula, where environmental complexity arises due to the co-existence of multiple agents [5, 46, 47, 58]. In addition to self-play originally used in two-player problem settings [65], the presence of multiple agents can give rise to an arms race [5] or create population dynamics that lead to the emergence of cooperation [61] and exploration [47].

Meta RL aims at equipping agents with the ability to generalize to tasks or environments that have not been encountered during training. Two nested processes of adaptation are traditionally considered: the inner level is a standard RL algorithm operating on a given environment, analog to a developmental learning process. The outer level is tuning the parameters of the inner loop such that it performs well on a distribution of environments, analog to an evolutionary process. Mechanisms are either gradient-based [27] or memory-based [74].

4 A BIRD’S EYE VIEW OF HUMAN BEHAVIORAL ECOLOGY

HBE emerged from the field of anthropology in the 70s and is today closely related to evolutionary psychology and cultural evolution [50]. A key assumption of HBE is that human behavior is highly flexible and can potentially adapt to changes in its environment [10, 50]. A profound challenge in studying behavioral adaptation is that the relationship between genes and behavior is to date not clear [62]. For this reason, HBE adheres to the assumption of the "phenotypic gambit," which posits that genetic, psychological or social constraints can be ignored when studying how optimal behaviors arise in a given environment [4].

Which factors contributed to the manifested ability of humans to generalize? What differentiated the human species from others that went extinct due to their inability to adapt to novel environments? These are two important questions that have occupied HBE, leading to a variety of hypotheses. The spotlight is on the Rift Valley at East Africa approximately 7 million years ago, as it is hypothesized that it constitutes a turning point in our evolutionary trajectory characterized by the first appearance of modern humans and their expansion to other geographical areas [50]. This evolutionary leap was originally studied under theories that layed emphasis on specific environmental changes. The Savannah hypothesis, for example, suggests that the change in fauna favored bipedal walking, which enabled migration and the creation of new niches for humans [50]. Later hypotheses under the pulsed climate variability framework, however, suggest that the key change was instead the general environmental complexity characterizing that period [19, 50, 60].

In Figure 1, we introduce a conceptual framework that recognizes important ecological components, as well as the feedback and feedback links that relate them. In the remainder of this section, we discuss hypotheses studying these relationships in the human ecosystem and, in Section 5, associate them with research questions in the study of artificial ecosystems. Under the proposed framework, environmental complexity is essentially driven by climate variability, which implies instability in the ecological conditions, in particular through changes in resource availability and exposition to predators [10, 30, 59]. This complexity has a strong influence on two major phenomena. First, it drives adaptability both at the evolutionary time scale, through speciation and extinction, [60, 64] and at the developmental time scale through cognitive mechanisms for learning and abstraction [30, 37]. Second, varying the levels of resource availability and exposition to predators has a strong influence on multi-agent dynamics through the modulation of cooperation and competition pressures [24, 71].

The influence of environmental complexity on adaptation and multi-agent dynamics can then have feedback and feedforward effects on the ecological system. First, increased morphological and cognitive complexity due to adaptation, as well as increased complexity in the multi-agent dynamics, feed back to environmental complexity through the modification of resource availability and predation pressure [29, 72]. For example, the Red Queen hypothesis [57] proposes that competition among different species is a major drive of evolution, possibly driving an arms race between co-adapting species. Second, adaptation and multi-agent dynamics can bootstrap in a feedforward manner the emergence of more advanced behaviors related to technology (e.g. tool use [22], communication (e.g. language [31, 71]) and culture (e.g. social norms [71], institutions [71] and religions [9]). Here again, the emergence of these new behaviors feeds back into environmental complexity through the process of social niche construction [29], thus creating a positive feedback loop potentially driving the ever-expanding social complexity of human ecology [43, 59].
5 AN ECOLOGICAL PERSPECTIVE ON RL

There seems to be a significant overlap between the questions that RL research poses in its study of the acquisition of open-ended behavior discussed in Section 3 and the hypotheses examined by HBE, presented in Section 4. In this section, we focus on three emergent phenomena attracting the interest of the RL community: adaptability of individual agents, multi-agent dynamics of groups and their cultural repertoire. By referring to concepts that we originally introduced in Figure 1 and highlight in this section, we initiate a dialogue between the two fields and identify key links that we believe deem further investigation by the RL community.

5.1 Adaptability

*Insights from ecology.* Under the pulsed climate variability framework discussed in Section 4, environmental factors such as climate variability, resource availability and predation pressures have served as a drive for the ability of humans to adapt to complex environments. Adaptability is achieved through mechanisms whose form depends on properties of the environment. If the environment is constant across time and space, natural selection may favor innate behaviors. By contrast, if the environment varies, natural selection might favor behavioral plasticity: based on environmental observations an agent may be able to switch between different behaviors following innate, and not learned instructions [25, 30, 37]. In cases where the environment changes noticeably across generations but slowly enough within a generation, behavioral plasticity is guided by a process of developmental selection, an example of which is the learning process, where an agent’s past behavior guide its future behavior. Thus, adaptation to environmental conditions operates on two scales: the evolutionary one drives speciation and extinction, while the developmental one drives learning. Adaptation feeds back into environmental complexity by affecting how environmental changes affect species, equipped with different skill repertoires. For example, during dry periods, the extinction rates of generalist species would reduce as they would be better able to find resources, while specialist species would struggle having lost their environmental niche and their competitive advantage [10].

*State of the art in RL.* The outer and inner loop optimization procedure that meta RL algorithms are following matches well with the aforementioned biological mechanisms of adaptation. There is a lot of flexibility in the choice of algorithms used to optimize the two loops: in [33, 53], an evolutionary algorithm is used in the outer loop and gradient descent in the inner loop, while in [27, 40] gradient descent is used in both loops. This comes in agreement with recent proposals to view evolution as equivalent to learning [76] and development [26]. Based on ecological insights, we can indicate the following research directions for investigating the effect of environmental variability in adaptability: (i) the rate of change is an important hyper-parameter: it should be high enough across generations to trigger mechanisms for adaptability but exhibit similarity patterns in order for innate behaviors to be useful; (ii) studying speciation and extinction in groups of intelligent agents may lead to interesting insights in meta RL, which is currently focusing on the continuous adaptation of a single learning mechanism.

5.2 Multi-agent dynamics

*Insights from ecology.* The emergence of cooperation has given birth to a variety of hypotheses in HBE. Under the big mistake hypothesis [11], altruism emerged in small-scale groups due to kin selection or reciprocity, and is today needlessly manifested because evolution has not yet adapted this mechanism. In contrast, the interdependence hypothesis [72] proposes a theory for the emergence of cooperation that replaces altruism with mutualistic collaboration. According to it, the need for foraging led to the selective helping of those who were needed as collaborative partners in the future. In sufficiently small groups, social selection was performed based on reputation. The size and structure of groups was...
dynamically shaped by their need to maintain stability and defend themselves against other groups. Competition between co-existing groups and species also gives rise to arms races, where reciprocal selection and adaptation lead to co-evolution [7]. Even at this small scale, the multi-agent dynamics feed back into environmental complexity through process of niche construction: predation, nutrient excretion and habitat modification populations alter their environment and further influence future populations [59].

State of the art in RL. The emergence of cooperation has attracted significant interest in the MARL community. In particular, multi-agent autocrucilla have leveraged the feedback effect that multi-agent dynamics have on resource availability [46], as well as arms races between competing groups [5]. In [61], agents learn how to fairly access a common pool of resources following simple trial-and-error learning. Recent works have studied the role of intrinsic motivation based on the theories of assortative matching and group selection [73], inequity aversion based on fairness norms [38] and social influence [39]. In [2], ecology-inspired hierarchical organizations are used to facilitate decentralized learning. The feedback effect that population dynamics have on the environment was investigated in [47], where increase in population size indirectly lead to exploration. As our brief discussion of related HBE literature however reveals, there exist a number of hypotheses and observations that researchers can leverage to further advance research in MARL: (1) according to the inter-dependence hypothesis, the human drive to cooperate was born neither in scenarios that required altruism nor in social dilemmas, which have served as an application ground for the majority of works in MARL. Rather, cooperation arose in Stag hunt type situations, which favored mutualistic collaboration [72]; (2) group properties such as size and structure are directly related to the multi-agent dynamics of stability and competition. Thus, their influence on the emerging multi-agent autocrucilla requires investigation.

5.3 Cultural repertoire

Insights from ecology. Non-human species often exhibit impressive behavioral repertoires [41]. However, human ecology is characterized by a uniquely large behavioral repertoire: engineering, language, social norms, institutions and religious beliefs constitute a complex cultural ecosystem that has lead scientists on the search of factors that differentiated us from other species [9, 70, 72]. According to the inter-dependence hypothesis, social norms and institutions emerged to counteract the fact that reputation alone could no longer alleviate the problem of free riding in large groups. In addition, the social complexity hypothesis [41] states that language worked as a bonding mechanism that replaced grooming, practiced in small-scale societies, and thus helped with maintaining group stability in larger groups [24]. The feedforward and feedback links associated with tool use have also been investigated under a number of, often contesting, hypotheses. Based on the data analysis in [29], environmental variability such as risk of resource failure, mobility and climate characteristics correlate significantly with tool use in food-gathering societies. However, it is the group size and not these factors that affect tool use in food-producing societies. It is therefore conjectured that the feedback link of societies with a larger cultural repertoire has a stabilizing effect, dampening the forward impact of environmental variability [18].

Another important link at this level is the relationship between tool use, language and adaptability. Studies of biological motor systems and language acquisition in infants have revealed that action and language representation share a similar compositional structure [12]. To understand how this similarity between two apparently distinct systems arose, one needs to turn to the origins of this relationship in human ecology. According to the Corbally hypothesis [20], the ability of primates’ to manipulate tools may have played a pivotal role in the evolution of language by creating the cognitive representations that compositionality requires. At the same time, the compositional structure of language is hypothesized to be an enabler of flexible and adaptable behavior, thus feeding back to adaptability [12].

State of the art in RL. Recent works in MARL have studied the emergence of communication [28, 44] as well as social norms and conventions [42, 48]. The effect that the type of social network organization, its average degree, and local connectivity has on communication learning in groups of deep reinforcement learning agents is investigated in [23]. This work constitutes an important first step in the realm of the social complexity hypothesis [24], but there remain a number of research directions lying at the intersection of MARL and meta RL: (1) the feedback effect that the cultural repertoire has on environmental variability through cultural niche construction [29] can potentially create more powerful autocrucilla than the already studied ones based on niche construction in small-scale groups [5, 61]; (2) studying the stabilization effect of cultural niche construction can provide important insights to the problem of scaling up artificial multi-agent systems; (3) the relationship between action/language compositionality and the ability of agents to generalize and adapt needs to be further investigated in order to transfer insights from human language acquisition to intelligent agents [17].

6 DISCUSSION

Our proposal is but a preliminary step towards realizing the potential of a cross-disciplinary dialogue between the HBE and RL communities, a glimpse of which has already been offered by recent works. For example, it has been proposed that RL methods can enable experimentation in large state spaces with unknown dynamics, overcoming the limitations of stochastic dynamic programming approaches currently dominating HBE [30]. The family of RL methods is, however, constantly expanding, with MARL in particular offering insights into emergent complexity. Thus, our discussion reveals that the potential of RL as a computational tool for enriching the analytical toolbox of HBE has not been fully realized. This potential has already been recognized for language development, an important component of human ecology [51].

At the other end of the spectrum, recent works in RL are increasingly drawing inspiration from ecology [73], as well as psychology [39] and economics [38]. As our proposal illustrates, however, this attempt is currently limited due to the lack of an overall picture and a thorough examination of feedforward and feedback links taking place at different ecological levels. We believe that conceptual
frameworks, such as the one that guided our current analysis in Figure 1, can serve as an important basis for this inter-disciplinary dialogue, with different questions zooming in on different sub-parts and potentially revealing lower-level relationships.

ACKNOWLEDGEMENT

This work is funded by the Inria Exploratory Action ORIGINS (2020-2022) and the ANR JCJC ECOCURL project (2020-2025).

REFERENCES

[1] Abbas Abdolmaleki, Jost Tobias Springenberg, Yuval Tassa, Remi Munos, Nicolas Heess, and Martin Riedmiller. 2018. Maximum a Posteriori Policy Optimisation. arXiv:1806.06920 [cs, math, stat] (June 2018). http://arxiv.org/abs/1806.06920

[2] Sanjeevan Ahilan and Peter Dayan. 2019. Feudal Multi-Agent Hierarchies for Cooperative Reinforcement Learning. CoRR abs/1901.08492 (2019). arXiv:1901.08492 http://arxiv.org/abs/1901.08492

[3] Maruan Al-Shedivat, Trapti Bansal, Yuri Burda, Ilya Sutskever, Igor Mordatch, and Pieter Abbeel. 2018. Continuous Adaptation via Meta-Learning in Non-Stationary and Competitive Environments. arXiv:1710.03641 [cs] (Feb. 2018). http://arxiv.org/abs/1710.03641

[4] Grafen Alan. 1984. Natural selection, kin selection and group selection. Behavioral ecology: an evolutionary approach (1984).

[5] Bowen Baker, Ingmar Kanitscheider, Todor Markov, Yi Wu, Glenn Powell, Bob McGrew, and Igor Mordatch. 2020. Emergent Tool Use From Multi-Agent Autocurricula. https://openreview.net/forum?id=Sk5psJRKs w tex=Baker2019

[6] Wolfgang Banzhaf, Bert Baumgaertner, Guillaume Beslon, René Doursat, Wolfgang Banzhaf, and Peter Dayan. 2019. Feudal Multi-Agent Hierarchies for Cooperative Reinforcement Learning. In Proceedings of the 36th International Conference on Machine Learning, Stockholm, Sweden. https://proceedings.mlr.press/v80/haarnoja18b.html

[7] C. Berkenkamp, T. Parchmann, Amanda Favis, and A. Steipelaki. 2003. Reciprocal Selection Causes a Coevolutionary Arms Race between Crossbills and Lodgepole Pine. The American Naturalist 162 (2003), 182 – 194.

[8] Glen Berseth, Daniel Geng, Coline Devin, Nicolas Rhinehart, Chelsea Finn, Dinesh Jayarajan, and Sergey Levine. 2016. Learning to Communicate with Deep Multi-Agent Reinforcement Learning. arXiv:1605.06676 [cs] (May 2016). http://arxiv.org/abs/1605.06676

[9] Willem E. Frankenhuis, Karthik Pananchathan, and Andrew G. Barto. 2019. Enriching behavioral ecology with reinforcement learning methods. Behavioral Processes 161 (2019), 94–100. https://doi.org/10.1016/j.bepros.2018.01.008

[10] Todd M. Freeberg, Robin I. M. Dunbar, and Terry J. Ord. 2012. Social complexity as a proximate and ultimate factor in communicative complexity. Philosophical Transactions of the Royal Society B: Biological Sciences 367, 1597 (July 2012), 1785–1801. https://doi.org/10.1098/rstb.2011.0213

[11] J. A. J. Gowlett. 2016. The discovery of fire by humans: a long and convoluted process. Philosophical Transactions of the Royal Society B: Biological Sciences 371, 1696 (June 2016), 20150164. https://doi.org/10.1098/rstb.2015.0164

[12] Jytte Grbic and Sebastian Risi. 2019. Towards Continual Reinforcement Learning through Evolutive Meta-Learning. In Proceedings of the Genetic and Evolutionary Computation Conference Companion (Prague, Czech Republic) (GECCO ’19). Association for Computing Machinery, New York, NY, USA, 119–120. https://doi.org/10.1145/3319619.3322044

[13] Tinoua Haartoja, Aurick Morcos, Pieter Abbeel, and Sergey Levine. 2018. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor (Proceedings of Machine Learning Research, Vol. 80). Jennifer Dy and Andreas Krause (Eds.). PMLR, Stockholmsmässan, Stockholm Sweden, 1861–1870. http://proceedings.mlr.press/v80/haartoja18b.html

[14] Jose Hernandez-Orallo. 2017. The measure of all minds: Evaluating Natural and Artificial Intelligence. Cambridge University Press.

[15] Felix Hill, Andrew Lampinen, Rosalia Schneider, Stephen Clark, Matthew Botvinick, James L McClelland, and Adam Santoro. 2020. ENVIRONMENTAL DRIVERS OF SYSTEMATICITY AND GENERALISATION IN A SITUATED AGENT. 20 (20).

[16] Dean Frederick Hougen and Syed Naveed Hussain Shah. 2019. The Evolution of Reinforcement Learning in Dynamic Environments. arXiv:1905.10985 [cs, q-bio] (July 2019). http://arxiv.org/abs/1905.10985

[17] Felix Hill, Andrew Lampinen, Rosalia Schneider, Stephen Clark, Matthew Botvinick, James L McClelland, and Adam Santoro. 2020. ENVIROMENTAL DRIVERS OF SYSTEMATICITY AND GENERALISATION IN A SITUATED AGENT. 20 (20).

[18] Mark Collard, Briggs Buchanan, April Ruttle, and Michael J O’Brien. [n.d.]. Niche Construction and the Toolkits of Hunter–Gatherers and Food Producers. (n.d.) 9.

[19] Mark Collard, Michael Kemeny, and Samantha Banks. 2005. Causes of Toolkit Variation Among Hunter-Gatherers: A Test of Four Competing Hypotheses. (2005). 20.

[20] Charles Darwin. 1859. On the Origin of Species by Means of Natural Selection or the Natural Selection of Favoured Races in the Struggle for Life. New York: D. Appleton and Company.

[21] Sophie A. de Beaune. 2004. The Invention of Technology: Prehistory and Cognition. Current Anthropology 45, 2 (April 2004), 159–162. https://doi.org/10.1086/383045

[22] R. I. M. Dunbar. 1993. Coevolution of neocortical size, group size and language in humans. Behavioral and Brain Sciences 16, 4 (Dec. 1993), 681–694. https://doi.org/10.1017/S0140525X00032525

[23] B. E. Erskridge and D. F. Hougen. 2012. Nurturing promotes the evolution of learning in uncertain environments. In 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (EDLR). 1–6. https://doi.org/10.1109/DevLRn.2012.6400847

[24] C. Fields and M. Levin. 2020. Does Evolution Have a Target Morphology? Organisms. Journal of Biological Sciences Vol 4 (Sept. 2020), 57–76 Pages. https://doi.org/10.13133/journal.of biological.sciences.2020.57-76

[25] Moreno Díaz, Daniel Geng, Celine Deniz, Nicolas Rhinehart, Bob McGrew, and Igor Mordatch. 2020. Emergent Tool Use From Multi-Agent Autocurricula. https://openreview.net/forum?id=Sk5psJRKsw tex=Baker2019

[26] Cédric Colas, Tristan Karch, Nicolas Lair, Jean-Michel Dossoux, Clément Moulin-Frier, Peter Ford Dominy, and Pierre-Yves Oudeyer. 2020. Language as a Cognitive Tool to Imagine Goals in Curiosity-Driven Exploration. arXiv:2002.09253 [cs, AI].
