Multi-Agent Reinforcement Learning as a Computational Tool for Language Evolution Research: Historical Context and Future Challenges

Clément Moulin-Frier, Pierre-Yves Oudeyer

Multi-Agent Reinforcement Learning as a Computational Tool for Language Evolution Research: Historical Context and Future Challenges

Clément Moulin-Frier and Pierre-Yves Oudeyer
Flowers team, Inria and Ensta ParisTech, France
clement.moulin-frier@inria.fr, pierre-yves.oudeyer@inria.fr

Abstract

Computational models of emergent communication in agent populations are currently gaining interest in the machine learning community due to recent advances in Multi-Agent Reinforcement Learning (MARL). Current contributions are however still relatively disconnected from the earlier theoretical and computational literature aiming at understanding how language might have emerged from a prelinguistic substance. The goal of this paper is to position recent MARL contributions within the historical context of language evolution research, as well as to extract from this theoretical and computational background a few challenges for future research.

Origins, formation and forms

There is a wide variety of approaches to studying the conditions in which human language might have emerged (Christiansen and Kirby 2003). As we will see, computer simulations have historically played an important role in the field. We can divide the problem in three sub-parts (Oudeyer 2006). Firstly, the study of the **forms** of language, i.e. of the structure of the phonemic, semantic, syntactic or pragmatic systems constituting it. Secondly, the study of its **formation**, i.e. of the genesis of these forms through sensory-motor, cognitive, environmental, social, cultural or evolutionary processes. Thirdly, the study of the **origins**, i.e. of the biological and environmental conditions that could have bootstrapped the formation process.

Under the infinite variety of its **forms**, human language is characterized by obvious regularities, the universals of language, which we find for example at the phonemic level (with vowels present in almost all languages of the world, (Maddieson and Precoda 1989)) and syntactic level (all languages have a recursive hierarchical structure, see e.g. (Pinker and Bloom 1990)). A fundamental research question concerns the **origins** of these regularities. Three main arguments are proposed in the literature. In the Chomskyan view of a genetically specified language acquisition device (Chomsky 1965), a common innate language competence shared by all humans would explain the regularities observed in the different languages. Another view about the

universal properties of human languages may be found in the hypothesis of a common origin, by which human languages would derive from an African mother tongue (Ruhlen 1996), imposing some common traces in spite of further cultural evolution producing their diversity. A third view considers that the forms of human language are the emergent product of an optimization process, inducing some commonality in the achieved solutions because of commonality in the cognitive mechanisms at hand, and because of common exterior constraints. This is the view first popularized by (Lindblom 1984), through a proposal to *"derive language from non-language"*. This last proposal opened a whole research program aiming at understanding the **formation** of human language, i.e. how a non-linguistic substance consisting in all the biological, cognitive and environmental mechanisms present before language, could both bootstrap its emergence and shape its universal properties, its form.

Theories on the formation of language

A large proportion of these theories postulate of a joint evolution of cooperative and communicative behaviors (Smith 2010; Gärdenfors 2002; Ghazanfar and Takahashi 2014; Tomasello et al. 2012). It is in particular the central thesis of the theory developed by Michael Tomasello, who proposes that *"humans' species-unique forms of cooperation –as well as their species-unique forms of cognition, communication, and social life—all derive from mutualistic collaboration (with social selection against cheaters)"* (Tomasello et al. 2012) . In this view, it is the constraints imposed by the ecological niche occupied by human beings that has forced them to jointly develop complex collaborative and communicative behaviors, in a context of interdependence requiring the sharing of intentions. We also find compatible arguments in the mirror system hypothesis developed by Michael Arbib (Arbib 2005) proposing that language evolution is grounded in the sensory-motor integration required for the execution and the observation of transitive actions towards objects, enabling other's intention recognition and providing the bases of a syntactic structure (Roy and Arbib 2005) (see also (Iriki and Taoka 2012) for theoretical propositions on the coevolution of tool use and language in humans). Finally, the social complexity hypothesis suggests that groups with complex

social structures require more complex communication systems to regulate interactions between group members (Freeberg, Dunbar, and Ord 2012).

Other theories highlight the role of sensory-motor learning and exploration as a key element to understand how speech communication could emerge from pre-existing morphological, perceptual and behavioral constraints (Lindblom 1984; MacNeilage 1998; Schwartz et al. 2012). A few theoretical contributions have proposed a potential role of curiosity-driven exploration in both language acquisition (Oller 2000) and evolution (Oudeyer and Smith 2015).

From verbal to computational descriptions

A major limitation of most of the theories mentioned above is that they are described in a verbal form. They are of course supported by experimental data but the description of the underlying hypotheses regarding the formation of linguistic structures mostly relies on a verbal explanation. This can be problematic because the aim of those theories is precisely to describe a complex dynamical process where linguistic structures emerge from multiple constraints in a prelinguistic environment (e.g. morphological, sensory-motor, cognitive, developmental, evolutionary or cultural constraints). Computer simulation is required to study the emergent properties of such a complex dynamical system.

For this reason, computational modeling has played a major role in language evolution research. Already in the 70s, Lindblom's "Dispersion Theory" (Liljencrants and Lindblom 1972) proposed that human phonological systems are optimized for maximizing auditory distances between phoneme pairs in order to enhance distinguishability. In these early contributions, language forms (e.g. the form of vowel systems) are considered as the equilibrium of a macroscopic system, analog to how thermodynamics describes changes in macroscopic physical quantities. In the 90s, these "global" approaches were completed by "local" approaches, were the equilibrium emerges from the interaction of "microscopic" elements, analog to statistical mechanics showing how the concepts from macroscopic observations are related to the description of microscopic states. These local approaches usually involve interacting prelinguistic agents and study how properties of human language can emerge from these interactions. A well-known example is the naming game paradigm showing how a shared communication system, associating signals emitted by the agents with semantic references to the external world, can self-organize out of a decentralized learning process from the local interactions between the agents (Steels 1997) (see (de Boer 2000; Moulin-Frier et al. 2015) for extensions to vocal communication and (Oudeyer 2005a; de Boer and Zuidema 2010) for extensions to combinatorial communication). However, these naming game models rarely address the issue of the functionality of communication (i.e. why to communicate?). Models from the field of evolutionary robotics (Quinn 2001; Grouchy et al. 2016) have the advantage of considering more realistic interaction scenarios than naming games but they specifically focus on genetic evolution algorithms, which do not consider the role of sensory-motor learning processes.

Computational models of emergent communication in agent populations are currently gaining interest in the machine learning community, due in particular to recent advances in Multi-Agent Reinforcement Learning (MARL) (see (Hernandez-Leal, Kartal, and Taylor 2019) for a survey). These new possibilities have allowed to overcome certain limitations of earlier contributions in two main directions. On the one hand, the paradigm of naming games presented above has been extended to more realistic references to the external world, learning directly from observations of raw images (Lazaridou et al. 2018). On the other hand, recent contributions based on the paradigm of partially-observable cooperative Markov games (Littman 1994; Leibo et al. 2017) have shown how a communication system can emerge to solve cooperative tasks in sequential environments (Sukhbaatar, Szlam, and Fergus 2016; Mordatch and Abbeel 2017; Foerster et al. 2016). These contributions adopt an utilitarian view of communication, where communication emerges as a way to solve complex cooperative tasks (Gauthier and Mordatch 2016).

Extracting future challenges for MARL

The utilitarian approach relying on partially observable cooperative Markov games provides a powerful conceptual and computational framework for modeling emergent communication as a way to solve complex problems in sequential environments. However, existing contributions are still relatively disconnected from the earlier literature presented in the previous section. In this section, we will extract from this theoretical and computational background a few challenges for future MARL research.

Decentralized learning

As mentioned in the previous section, the first models attempting to predict language forms from a prelinguistic substance adopted a global, macroscopic approach. This global approach has then be complemented by a local, microscopic approach where language forms emerge from the repeated interactions between individual agents.

A large proportion of current MARL contributions rely on *centralized learning decentralized execution* algorithms (Sukhbaatar, Szlam, and Fergus 2016; Mordatch and Abbeel 2017; Foerster et al. 2016), analog to a global macroscopic approach. While centralized learning is able to efficiently solve complex problems, the lack of biological plausibility strongly limits its use in language evolution research. Contributions relying on decentralized learning (Jaques et al. 2019) are less efficient from a performance point of view but have the advantage of highlighting important issues regarding the unstable nature of cooperative and communicative behavior in multi-agent settings, due e.g. to the non-stationarity it induces. Solving such issues is an important challenge in both MARL and language evolution research.

Role of morphological and sensory-motor constraints

Current MARL contributions mostly rely on an idealized communication channel where the signal produced by an

agent is directly broadcasted to other agents (Sukhbaatar, Szlam, and Fergus 2016; Mordatch and Abbeel 2017; Foerster et al. 2016), similar to earlier contributions based on the naming game paradigm. In contrast, speech communication is strongly shaped by sensory-motor constraints, involving the control of vocal articulators (e.g. the jaw, the tongue, the lips) for modulating a sound wave resulting in the perception of acoustic features. Vocal control is actually a classical robotic problem, where the agent has to decide how to move vocal articulators to reach acoustic targets. This control problem is a difficult one due to the complex morphology of the vocal tract, the highly nonlinear nature of the articulatory-to-acoustic transformation, as well as the presence of acoustic noise in the environment. Earlier contributions have studied how vocal communication can emerge from the interaction of sensory-motor agents equipped with articulatory synthesizers, i.e. computer models of the human vocal tract able to generate sound waves from articulator trajectories (Moulin-Frier et al. 2015; Moulin-Frier, Nguyen, and Oudeyer 2014). This resulted in multi-agent simulations able to predict the statistical tendencies of the phonological systems used in world languages (Oudeyer 2005b), as well as to test hypotheses regarding the influence of prelinguistic orofacial behaviors on the syllabic structure of speech communication ((Moulin-Frier et al. 2015), following an hypothesis from (MacNeilage 1998)). Introducing biologically plausible sensory-motor abilities of signal production and perception in MARL models would allow to extend the aforementioned results to more complex environments and learning abilities.

Role of intrinsic motivation

A few theoretical contributions have proposed a potential role of curiosity-driven exploration in both language acquisition (Oller 2000) and evolution (Oudeyer and Smith 2015). Active exploration can spontaneously generate diverse behaviors from modality-independent and task-independent internal drives. Such spontaneous behavior can result in vocal activity that may have bootstrapped the emergence of communication. This hypothesis is supported by computational simulations showing a role of curiosity-driven exploration in vocal development (Moulin-Frier, Nguyen, and Oudeyer 2014), social affordance discovery (Oudeyer and Kaplan 2006) and the active control of complexity growth in naming games (Schueller and Oudeyer 2015).

Despite recent progress in curiosity-driven RL (Pathak et al. 2017; Colas et al. 2019), very few MARL contributions have used such algorithms for studying emergent communication (see (Jaques et al. 2019) but which is specific to social interactions on a single task). It is a promising direction of research to explore how general-purpose curiosity-driven multi-task reinforcement learning algorithms (Colas et al. 2019) can be integrated in multi-agent environments to encourage the discovery of complex communication systems supporting the acquisition of an open-ended repertoire of cooperative skills. A key step in this direction has recently been proposed in the IMAGINE architecture (Colas et al. 2020), where an agent uses language compositionality to generate new goals by composing known ones.

Emergent complexity

Earlier contributions in language evolution modeling has often been limited by the use of simplistic environments and learning abilities. Recent advances in MARL can allow to overcome these limitations to show how language complexity can emerge as a way to optimize behavior in complex cooperative environments. In particular, recent contributions in MARL have shown how an autocurriculum of increasingly complex behaviors can emerge from agent's coadaptation in mixed cooperative-competitive environments (Bansal et al. 2018; Baker et al. 2019). Can such an auto-curriculum through coadaptation favor the emergence of increasingly complex communicative systems? In turn, can complex communication favor the emergence of increasingly complex cooperative strategies? Addressing these open questions can potentially help to understand the processes that have shaped the impressive complexity of human language.

Conclusion

Recent advances in MARL provides a powerful conceptual and computational framework for modeling emergent communication as a way to solve complex problems in sequential environments. There are however important differences in the methodology and the objectives between 1) implementing efficient and robust multi-agent systems learning how to communicate for solving complex problems (as it is the case in the majority of recent MARL contributions), vs. 2) using multi-agent learning as a computational tool for better understanding human language evolution (an approach which has historically played an important role in language evolution research, see (Oudeyer 2006) for an epistemological analysis). In this paper we have reviewed earlier computational contributions and have extracted from them a few future challenges for MARL research.

References

[Arbib 2005] Arbib, M. A. 2005. From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28:105–167.

[Baker et al. 2019] Baker, B.; Kanitscheider, I.; Markov, T.; Wu, Y.; Powell, G.; McGrew, B.; and Mordatch, I. 2019. Emergent Tool Use From Multi-Agent Autocurricula.

[Bansal et al. 2018] Bansal, T.; Pachocki, J.; Sidor, S.; Sutskever, I.; and Mordatch, I. 2018. Emergent Complexity via Multi-Agent Competition. In *International Conference on Learning Representations*.

[Chomsky 1965] Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

[Christiansen and Kirby 2003] Christiansen, M. H., and Kirby, S. 2003. Language evolution: Consensus and controversies. *Trends in Cognitive Sciences* 7:300–307.

[Colas et al. 2019] Colas, C.; Sigaud, O.; Oudeyer, P.-Y.; Fournier, P.; Chetouani, M.; Sigaud, O.; and Oudeyer, P.-Y. 2019. CURIOUS: Intrinsically Motivated Multi-Task, Multi-Goal Reinforcement Learning. In *Proceedings of the 36th International Conference on Machine Learning*, 1331–1340.

[Colas et al. 2020] Colas, C.; Karch, T.; Lair, N.; Moulin-Frier, C.; Dussoux, J.-M.; Dominey, P. F.; and Oudeyer, P.-Y. 2020. Language as a Cognitive Tool to Imagine Goals in Curiosity Driven

Exploration. In *Advances in Neural Information Processing Systems (NeurIPS 2020)*.

[de Boer and Zuidema 2010] de Boer, B., and Zuidema, W. 2010. Multi-Agent Simulations of the Evolution of Combinatorial Phonology. *Adaptive Behavior* 18(2):141–154.

[de Boer 2000] de Boer, B. 2000. Self-organization in vowel systems. *Journal of Phonetics* 28(4):441–465.

[Foerster et al. 2016] Foerster, J.; Assael, Y. M.; de Freitas, N.; and Whiteson, S. 2016. Learning to communicate with deep multi-agent reinforcement learning. In *Advances in Neural Information Processing Systems*, 2137–2145.

[Freeberg, Dunbar, and Ord 2012] Freeberg, T. M.; Dunbar, R. I. M.; and Ord, T. J. 2012. Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1597):1785–1801.

[Gärdenfors 2002] Gärdenfors, P. 2002. *Cooperation and the evolution of symbolic communication*. Lund University.

[Gauthier and Mordatch 2016] Gauthier, J., and Mordatch, I. 2016. A Paradigm for Situated and Goal-Driven Language Learning. In *NIPS 2016 Machine Intelligence Workshop*.

[Ghazanfar and Takahashi 2014] Ghazanfar, A. A., and Takahashi, D. Y. 2014. The evolution of speech: Vision, rhythm, cooperation. *Trends in Cognitive Sciences* 18(10):543–553.

[Grouchy et al. 2016] Grouchy, P.; D'Eleuterio, G. M. T.; Christiansen, M. H.; and Lipson, H. 2016. On The Evolutionary Origin of Symbolic Communication. *Scientific Reports* 6(1):34615.

[Hernandez-Leal, Kartal, and Taylor 2019] Hernandez-Leal, P.; Kartal, B.; and Taylor, M. E. 2019. A survey and critique of multiagent deep reinforcement learning. *Autonomous Agents and Multi-Agent Systems* 33(6):750–797.

[Iriki and Taoka 2012] Iriki, A., and Taoka, M. 2012. Triadic (ecological, neural, cognitive) niche construction: a scenario of human brain evolution extrapolating tool use and language from the control of reaching actions. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1585):10–23.

[Jaques et al. 2019] Jaques, N.; Lazaridou, A.; Hughes, E.; Gulcehre, C.; Ortega, P. A.; Strouse, D.; Leibo, J. Z.; and de Freitas, N. 2019. Social Influence as Intrinsic Motivation for Multi-Agent Deep Reinforcement Learning. In *Proceedings of the 35 th International Conference on Machine Learning, Stockholm, Sweden*.

[Lazaridou et al. 2018] Lazaridou, A.; Hermann, K. M.; Tuyls, K.; and Clark, S. 2018. Emergence of Linguistic Communication from Referential Games with Symbolic and Pixel Input. In *Sixth International Conference on Learning Representations (ICLR 2018)*.

[Leibo et al. 2017] Leibo, J. Z.; Zambaldi, V.; Lanctot, M.; Marecki, J.; and Graepel, T. 2017. Multi-agent Reinforcement Learning in Sequential Social Dilemmas. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, 464–473. International Foundation for Autonomous Agents and Multiagent Systems.

[Liljencrants and Lindblom 1972] Liljencrants, J., and Lindblom, B. 1972. Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast. *Language* 48(4):839–862.

[Lindblom 1984] Lindblom, B. 1984. Can the models of evolutionary biology be applied to phonetic problems. In *Proceedings of the tenth international congress of phonetic sciences*, 67–81. Foris Pubns USA.

[Littman 1994] Littman, M. L. 1994. Markov games as a framework for multi-agent reinforcement learning. *Machine Learning Proceedings 1994* 157–163.

[MacNeilage 1998] MacNeilage, P. F. 1998. The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences* 21:499–511.

[Maddieson and Precoda 1989] Maddieson, I., and Precoda, K. 1989. Updating UPSID. *The Journal of the Acoustical Society of America* 86(S1):S19.

[Mordatch and Abbeel 2017] Mordatch, I., and Abbeel, P. 2017. Emergence of Grounded Compositional Language in Multi-Agent Populations. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

[Moulin-Frier et al. 2015] Moulin-Frier, C.; Diard, J.; Schwartz, J.-L. J.-L.; and Bessière, P. 2015. COSMO ('Communicating about Objects using Sensory-Motor Operations'): a Bayesian modeling framework for studying speech communication and the emergence of phonological systems. *Journal of Phonetics* 53:5–41.

[Moulin-Frier, Nguyen, and Oudeyer 2014] Moulin-Frier, C.; Nguyen, S. M.; and Oudeyer, P.-Y. 2014. Self-Organization of Early Vocal Development in Infants and Machines: The Role of Intrinsic Motivation. *Frontiers in Psychology* 4(1006).

[Oller 2000] Oller, D. K. 2000. *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum Associates.

[Oudeyer and Kaplan 2006] Oudeyer, P.-Y., and Kaplan, F. 2006. Discovering Communication. *Connection Science* 18(June 2006):189–206.

[Oudeyer and Smith 2015] Oudeyer, P.-Y., and Smith, L. 2015. How Evolution may work through Curiosity-driven Developmental Process. *Topics in Cognitive Science.* in press.

[Oudeyer 2005a] Oudeyer, P.-Y. 2005a. The self-organization of combinatoriality and phonotactics in vocalization systems. *Connection Science* 17(3-4):325–341.

[Oudeyer 2005b] Oudeyer, P.-Y. 2005b. The self-organization of speech sounds. *Journal of Theoretical Biology* 233(3):435–449.

[Oudeyer 2006] Oudeyer, P.-Y. 2006. *Self-Organization in the Evolution of Speech*, volume 6 of *Studies in the Evolution of Language*. Oxford University Press.

[Pathak et al. 2017] Pathak, D.; Agrawal, P.; Efros, A. A.; and Darrell, T. 2017. Curiosity-driven exploration by self-supervised prediction. In *International Conference on Machine Learning (ICML)*, volume 2017.

[Pinker and Bloom 1990] Pinker, S., and Bloom, P. 1990. Natural language and natural selection. *Behavioral and brain sciences* 13(4):707–727.

[Quinn 2001] Quinn, M. 2001. Evolving communication without dedicated communication channels. In *European Conference on Artificial Life*, 357–366. Springer.

[Roy and Arbib 2005] Roy, A. C., and Arbib, M. A. 2005. The syntactic motor system. *Gesture* 5(1):7–37.

[Ruhlen 1996] Ruhlen, M. 1996. *The Origin of Language: Tracing the Evolution of the Mother Tongue*. New York: John Wiley & Sons.

[Schueller and Oudeyer 2015] Schueller, W., and Oudeyer, P.-Y. 2015. Active learning strategies and active control of complexity growth in naming games. In *2015 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 220–227. IEEE.

[Schwartz et al. 2012] Schwartz, J.-L.; Basirat, A.; Ménard, L.; and Sato, M. 2012. The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics* 25(5):336–354.

[Smith 2010] Smith, E. A. 2010. Communication and collective action: language and the evolution of human cooperation. *Evolution and Human Behavior* 31(4):231–245.

[Steels 1997] Steels, L. 1997. The synthetic modeling of language origins. *Evolution of Communication* 1(1):1–34.

[Sukhbaatar, Szlam, and Fergus 2016] Sukhbaatar, S.; Szlam, A.; and Fergus, R. 2016. Learning Multiagent Communication with Backpropagation. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*.

[Tomasello et al. 2012] Tomasello, M.; Melis, A. P.; Tennie, C.; Wyman, E.; and Herrmann, E. 2012. Two Key Steps in the Evolution of Human Cooperation. *Current Anthropology* 53(6):673–692.