Decentralised Emergence of Robust and Adaptive Linguistic Conventions in Populations of Autonomous Agents Grounded in Continuous Worlds

Extended Abstract

Jérôme Botoko Ekila
Vrije Universiteit Brussel
Brussels, Belgium
jerome@ai.vub.ac.be

Jens Nevens
Vrije Universiteit Brussel
Brussels, Belgium
jens@ai.vub.ac.be

Lara Verheyen
Vrije Universiteit Brussel
Brussels, Belgium
lara.verheyen@ai.vub.ac.be

Katrien Beuls
Université de Namur
Namur, Belgium
katrien.beuls@unamur.be

Paul Van Eecke
Vrije Universiteit Brussel
Brussels, Belgium
paul@ai.vub.ac.be

ABSTRACT

This paper introduces a methodology for establishing linguistic conventions in populations of autonomous agents in a fully decentralised manner. As agents take part in local communicative interactions, they gradually establish a common conceptual system and vocabulary that enables them to communicate about arbitrary entities in their continuous environment. Apart from introducing the methodology, we also present six experiments that showcase the robustness of the methodology against sensor defects, its ability to handle noisy observations and uncalibrated sensors, and its suitability for continual learning.

KEYWORDS

language emergence; multi-agent systems; autonomous agents; emergent communication; self-organisation; language evolution

ACM Reference Format:
Jérôme Botoko Ekila, Jens Nevens, Lara Verheyen, Katrien Beuls, and Paul Van Eecke. 2024. Decentralised Emergence of Robust and Adaptive Linguistic Conventions in Populations of Autonomous Agents Grounded in Continuous Worlds: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

1 INTRODUCTION

The field of emergent communication investigates how populations of artificial agents can collaboratively solve tasks by developing communication protocols that evolve through processes of interaction and adaptation. In recent years, the multi-agent reinforcement learning (MARL) framework has been adopted to tackle this challenge [3–6, 10–12, 14–17, 19, 22, 24, 33, 37]. While impressive results have been achieved within the MARL framework, the experimental setups in these studies often deviate significantly from how languages emerge and evolve in humans [34]. For example, some experiments involve only two agents [4, 12, 22], restrict agents to either speaking or listening [6, 7, 15, 19], or learning is not decentralised [11, 14]. In contrast, our objective is to facilitate language emergence through self-organisation, drawing inspiration from the language game experimental paradigm [27] and prior research on the emergence of perceptually grounded vocabularies [1, 2, 18, 20, 21, 23, 25–32, 34–36, 38, 39].

This paper introduces a methodology to establish, in a fully decentralised manner, a linguistic convention within a population of autonomous agents that enables the agents to refer to arbitrary entities in their environment. As they take part in local communicative interactions, agents gradually build up a linguistic inventory consisting of word forms associated with conceptual representations. Our contribution surpasses previous efforts in the language game paradigm as the methodology achieves three properties at once: decentralised, communication-based concept learning (i) in continuous feature spaces (as opposed to the discrete setting in [38]), (ii) in multi-agent emergent settings (as opposed to the tutor-learner setting in [21]), and (iii) with direct applicability to any dataset characterising entities in terms of continuously-valued features. An extended version of this paper is available at https://arxiv.org/abs/2401.08461.

2 METHODOLOGY

Autonomous agents are endowed with sensors to perceive their environment. However, due to variations in hardware, noisy observations, uncalibrated sensors, and potential sensor defects, a misalignment can emerge in the information perceived by individual agents. To address this mismatch, agents create abstractions in the form of concepts. Rather than directly transmitting sensor values, agents learn to communicate by associating words with internal learned conceptual representations. These representations, unique to each agent, serve as a bridge between the raw sensor data and an emergent linguistic convention [21]. Agents start out with an empty linguistic inventory, i.e. there are no predefined

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

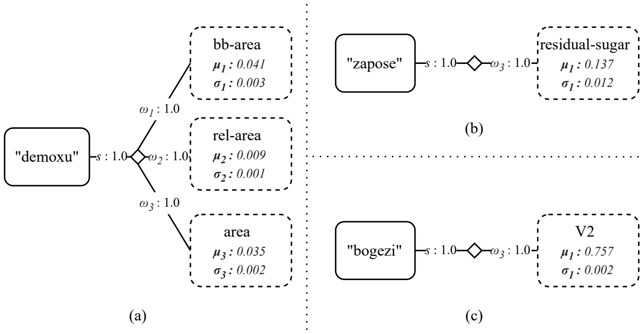

Figure 1: Examples of emerged concepts for the CLEVR (a), WINE (b) and CREDIT (c) datasets.

concepts or words. As they take part in pairwise, local communicative interactions, agents gradually build up this personal linguistic inventory.

In our methodology, the conceptual representation consists of a set of weighted Gaussian distributions, inspired by [21]. The agent's sensors perceive each entity within the environment, producing a set of values. These values are encapsulated by individual Gaussian distributions, each representing the sampled values from a specific sensory channel. The weight associated with each distribution signifies the importance of each sensor to the conceptual representation.

The agents employ a discrimination-based strategy, which focuses on constructing conceptual representations that enable effective differentiation of objects based on their distinct attributes. In each interaction, two randomly selected agents from the population assume the roles of speaker and hearer. The goal of the speaker is to draw the attention of the hearer to a specific object. To achieve this, the strategy involves creating representations that combine attributes that facilitate the discrimination between objects. Critically, the representation is bidirectional as it can be used to both produce and understand utterances in communication. After each interaction, the speaker provides feedback to the hearer by revealing the intended topic. Depending on the outcome of the game, the agents independently update their knowledge by adjusting their conceptual representations. As agents engage in iterative interactions, they progressively refine their representations and communicate more effectively. Over time, this dynamic process gives rise to the emergence of linguistic conventions within the population. Our methodology introduces novel invention, adoption, conceptualisation and alignment mechanisms for constructing such representations in a fully decentralised manner.

3 EXPERIMENTAL VALIDATION

We experimentally validate the effectiveness, flexibility, and robustness of our methodology using three large tabular datasets. These datasets cover three different types of domains. The CLEVR dataset consists of 85k visual scenes of geometric objects [13], the WINE dataset consists of physicochemical analyses of Portuguese wines [8], and finally the CREDIT dataset consists of approximately 250k principal component analyses of credit card transactions [9]. Each row of a dataset represents an entity in the agents' environment.

We conduct six different experiments. The first baseline experiment demonstrates the methodology's effectiveness on the three datasets. After 1 million interactions, the population can successfully communicate ±99.6% of the time on all three datasets. The second experiment tackles compositional generalisability by applying our methodology to a modified version of CLEVR (CoGenT) [13], showcasing the emergent concepts' adaptability to previously unseen attribute combinations. The third experiment extends the evaluation to heteromorphic populations, revealing that even when agents possess varying sensor combinations, a high degree of communicative success is still achievable. The fourth experiment tests the methodology's resilience against sensor defects, demonstrating its robustness in the face of sudden malfunctions and highlighting the lasting benefits of an established linguistic convention. The fifth experiment explores the methodology's robustness against differences in agents' perception. Lastly, the sixth experiment focuses on continual learning, confirming the methodology's adequacy and resistance to catastrophic forgetting.

Figure 1 depicts three concepts that emerged in the first baseline experiment. We only visualise feature channels with positive weights. Through a qualitative analysis, the emerged concept "demoxu" (see Fig. 1a) is found to be primarily used by agents to refer to small objects in the CLEVR environment. Likewise, "zapose" (see Fig. 1b) is used by agents to refer to "medium-sweet" wines in the WINE environment. Finally, "bogezi" (see Fig. 1c) emerged to refer to a particular kind of credit card transaction. While this concept remains transparent, it is not interpretable by humans due to its grounding in principal component analyses. This example demonstrates the usefulness of a self-organising system with emergent conventions. In such a system, agents are not limited to existing words but are able to create and adapt a convention that is tailored to their sensory endowment and their communicative needs.

4 CONCLUSION

In this paper, a novel methodology is presented for the emergence of communicatively effective, robust, and adaptive linguistic conventions among autonomous agents. The methodology facilitates the decentralised emergence of linguistic conventions through local, task-oriented interactions between pairs of agents, resulting in symbolic labels associated with concept representations grounded in a multi-dimensional, continuous feature space. These associations are individually constructed by agents, yet compatible on a communicative level. The methodology is validated through experiments across three diverse datasets, showcasing its effectiveness in various domains. We introduce a model for the emergence and evolution of linguistic conventions in populations of autonomous agents which is applicable to any dataset that characterises entities using continuously-valued features.

ACKNOWLEDGMENTS

This research received funding from EU's H2020 RIA (grant no. 951846), Research Foundation Flanders (FWO) (grant no. 76929), ARIAC (project no. 2010235) and the Flanders AI Research Program. Last authorship is shared between KB and PVE.

REFERENCES

[1] Tony Belpaeme and Joris Bleys. 2005. Colourful language and colour categories. In *Proceedings of the Second International Symposium on the Emergence and Evolution of Linguistic Communication (EELC 2005)*, Chrystopher L. Nehaniv and Angelo Cangelosi (Eds.). AISB, Bath, United Kingdom, 1–8.

[2] Joris Bleys. 2016. *Language strategies for the domain of colour.* Language Science Press, Berlin, Germany.

[3] Ben Bogin, Mor Geva, and Jonathan Berant. 2018. Emergence of communication in an interactive world with consistent speakers. In *Emergent Communication Workshop: NeurIPS 2018*.

[4] Diane Bouchacourt and Marco Baroni. 2018. How agents see things: On visual representations in an emergent language game. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Ellen Riloff, David Chiang, Julia Hockenmaier, and Jun'ichi Tsujii (Eds.). Association for Computational Linguistics, 981–985. https://doi.org/10.18653/v1/D18-1119

[5] Kris Cao, Angeliki Lazaridou, Marc Lanctot, Joel Z. Leibo, Karl Tuyls, and Stephen Clark. 2018. Emergent communication through negotiation. In *6th International Conference on Learning Representations (ICLR 2018)*. 1–15.

[6] Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. 2021. Communicating artificial neural networks develop efficient color-naming systems. *Proceedings of the National Academy of Sciences* 118, 12 (2021), e2016569118. https://doi.org/10.1073/pnas.2016569118

[7] Rahma Chaabouni, Florian Strub, Florent Altché, Eugene Tarassov, Corentin Tallec, Elnaz Davoodi, Kory Wallace Mathewson, Olivier Tieleman, Angeliki Lazaridou, and Bilal Piot. 2022. Emergent communication at scale. In *10th International Conference on Learning Representations (ICLR 2022)*. 1–30.

[8] Paulo Cortez, Antonio Cerdeira, Fernando Almeida, Telmo Matos, and José Reis. 2009. Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems* 47, 4 (2009), 547–553. https://doi.org/10.1016/j.dss.2009.05.016

[9] Andrea Dal Pozzolo, Olivier Caelen, Yann-Aël Le Borgne, Serge Waterschoot, and Gianluca Bontempi. 2014. Learned lessons in credit card fraud detection from a practitioner perspective. *Expert Systems with Applications* 41, 10 (2014), 4915–4928. https://doi.org/10.1016/j.eswa.2014.02.026

[10] Abhishek Das, Satwik Kottur, José M. F. Moura, Stefan Lee, and Dhruv Batra. 2017. Learning cooperative visual dialog agents with deep reinforcement learning. In *2017 IEEE International Conference on Computer Vision (ICCV)*, Rita Cucchiara, Yasuyuki Matsushita, Nicu Sebe, and Stefano Soatto (Eds.). IEEE Computer Society, 2951–2960.

[11] Jakob Foerster, Ioannis Alexandros Assael, Nando de Freitas, and Shimon Whiteson. 2016. Learning to communicate with deep multi-agent reinforcement learning. In *Advances in Neural Information Processing Systems 29 (NIPS 2016)*, Daniel Lee, Masashi Sugiyama, Ulrike Von Luxburg, Isabelle Guyon, and Roman Garnett (Eds.). Curran Associates Inc., Red Hook, NY, USA, 2137–2145.

[12] Serhii Havrylov and Ivan Titov. 2017. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, Isabelle Guyon, Ulrike Von Luxburg, Samy Bengio, Hanna Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett (Eds.). Curran Associates Inc., Red Hook, NY, USA, 2146–2156.

[13] Justin Johnson, Bharath Hariharan, Laurens van der Maaten, Li Fei-Fei, C. Lawrence Zitnick, and Ross Girshick. 2017. CLEVR: A Diagnostic Dataset for Compositional Language and Elementary Visual Reasoning. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE Computer Society, 2901–2910.

[14] Jooyeon Kim and Alice Oh. 2021. Emergent Communication under Varying Sizes and Connectivities. In *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)*, Marc'Aurelio Ranzato, Alina Beygelzimer, Yann Dauphin, Percy S. Liang, and Jennifer W. Vaughan (Eds.). Curran Associates Inc., Red Hook, NY, USA, 17579–17591.

[15] Satwik Kottur, José Moura, Stefan Lee, and Dhruv Batra. 2017. Natural Language Does Not Emerge 'Naturally' in Multi-Agent Dialog. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Martha Palmer, Rebecca Hwa, and Sebastian Riedel (Eds.). Association for Computational Linguistics, 2962–2967. https://doi.org/10.18653/v1/D17-1321

[16] Angeliki Lazaridou, Karl Moritz Hermann, Karl Tuyls, and Stephen Clark. 2018. Emergence of linguistic communication from referential games with symbolic and pixel input. *arXiv preprint arXiv:1804.03984* (2018).

[17] Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. 2017. Multi-agent cooperation and the emergence of (natural) language. In *5th International Conference on Learning Representations (ICLR 2017)*. 1–11.

[18] Martin Loetzsch. 2015. *Lexicon formation in autonomous robots.* Ph.D. Dissertation. Humboldt-Universität zu Berlin, Berlin, Germany.

[19] Igor Mordatch and Pieter Abbeel. 2018. Emergence of grounded compositional language in multi-agent populations. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, Sheila McIlraith and Kilian Q. Weinberger (Eds.). AAAI Press, 1495–1502.

[20] Jens Nevens, Paul Van Eecke, and Katrien Beuls. 2019. Computational construction grammar for visual question answering. *Linguistics Vanguard* 5, 1 (2019), 20180070.

[21] Jens Nevens, Paul Van Eecke, and Katrien Beuls. 2020. From continuous observations to symbolic concepts: A discrimination-based strategy for grounded concept learning. *Frontiers in Robotics and AI* 7, 84 (2020). https://doi.org/10.3389/frobt.2020.00084

[22] Michael Noukhovitch, Travis LaCroix, Angeliki Lazaridou, and Aaron Courville. 2021. Emergent Communication under Competition. In *Proceedings of the 20th International Conference on Autonomous Agents and Multi-Agent Systems*. 974–982.

[23] Pierre-Yves Oudeyer and Frédéric Kaplan. 2007. Language evolution as a Darwinian process: Computational studies. *Cognitive Processing* 8, 1 (2007), 21–35. https://doi.org/10.1007/s10339-006-0158-3

[24] Cinjon Resnick, Ilya Kulikov, Kyunghyun Cho, and Jason Weston. 2017. Vehicle Communication Strategies for Simulated Highway Driving. In *Emergent Communication Workshop: NeurIPS 2017*.

[25] Michael Spranger. 2013. Grounded lexicon acquisition - Case studies in spatial language. In *Proceedings of the 2013 IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL)*. IEEE, 1–6.

[26] Michael Spranger. 2016. *The evolution of grounded spatial language.* Language Science Press, Berlin, Germany. https://doi.org/10.17169/langsci.b53.183

[27] Luc Steels. 1995. A self-organizing spatial vocabulary. *Artificial Life* 2, 3 (1995), 319–332. https://doi.org/10.1162/artl.1995.2.3.319

[28] Luc Steels. 2003. The evolution of communication systems by adaptive agents. In *Symposium on Adaptive Agents and Multi-Agent Systems*, Eduardo Alonso, Daniel Kudenko, and Dimitar Kazakov (Eds.). 125–140. https://doi.org/10.1007/3-540-44826-8_8

[29] Luc Steels. 2015. *The Talking Heads experiment: Origins of words and meanings.* Language Science Press, Berlin, Germany.

[30] Luc Steels and Tony Belpaeme. 2005. Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences* 28, 4 (2005), 469–489. https://doi.org/10.1017/S0140525X05000087

[31] Luc Steels and Martin Loetzsch. 2012. The grounded naming game. In *Experiments in Cultural Language Evolution*, Luc Steels (Ed.). Vol. 3. John Benjamins, Amsterdam, Netherlands, 41–59. https://doi.org/10.1075/ais.3.04ste

[32] Luc Steels, Martin Loetzsch, and Michael Spranger. 2016. A boy named Sue: The semiotic dynamics of naming and identity. *Belgian Journal of Linguistics* 30, 1 (2016), 147–169. https://doi.org/10.1075/bjl.30.07ste

[33] Sainbayar Sukhbaatar, Arthur Szlam, and Rob Fergus. 2016. Learning multiagent communication with backpropagation. In *Advances in Neural Information Processing Systems 29 (NIPS 2016)*, Daniel Lee, Masashi Sugiyama, Ulrike Von Luxburg, Isabelle Guyon, and Roman Garnett (Eds.). Curran Associates Inc., Red Hook, NY, USA, 2244–2252.

[34] Paul Van Eecke, Katrien Beuls, Jérôme Botoko Ekila, and Roxana Rădulescu. 2022. Language games meet multi-agent reinforcement learning: A case study for the naming game. *Journal of Language Evolution* 7, 2 (2022), 213–223. https://doi.org/10.1093/jole/lzad001

[35] Paul Vogt. 2005. The emergence of compositional structures in perceptually grounded language games. *Artificial intelligence* 167, 1–2 (2005), 206–242. https://doi.org/10.1016/j.artint.2005.04.010

[36] Paul Vogt. 2015. *How mobile robots can self-organise a vocabulary.* Language Science Press, Berlin, Germany. https://doi.org/10.17169/langsci.b50.113

[37] Sida I. Wang, Percy Liang, and Christopher D. Manning. 2016. Learning Language Games through Interaction. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Katrin Erk and Noah A. Smith (Eds.). Association for Computational Linguistics, 2368–2378. https://doi.org/10.18653/v1/P16-1224

[38] Pieter Wellens. 2012. *Adaptive Strategies in the Emergence of Lexical Systems.* Ph.D. Dissertation. Vrije Universiteit Brussel, Brussels: VUB Press.

[39] Pieter Wellens, Martin Loetzsch, and Luc Steels. 2008. Flexible word meaning in embodied agents. *Connection Science* 20, 2–3 (2008), 173–191. https://doi.org/10.1080/09540090802091966