EMERGE - Emergent Awareness from Minimal Collectives

Davide Bacciu[1(✉)], Vincenzo Ambriola[1], Bahador Bahrami[3], Andrea Ceni[1], Andrea Cossu[1], Valerio De Caro[1], Cosimo Della Santina[2], Ophelia Deroy[3], Claudio Gallicchio[1], Riccardo Guidotti[1], Sabine Hauert[5], Simon Jones[5], Jurgis Karpus[3], Suet Lee[5], Jingyue Liu[2], Vincenzo Lomonaco[1], Nadine Meertens[3], Emma Milner[5], Anna Monreale[1], Renan Picoreti Nakahara[4], Mariano Ramirez[2], Ebrahim Shahabi[2], and Maximilian Stölzle[2]

[1] University of Pisa, Pisa, Italy
davide.bacciu@unipi.it
[2] Delft University of Technology, Delft, Netherlands
[3] Ludwig Maximilian University, Munich, Germany
[4] Da Vinci Labs, Tours, France
[5] University of Bristol, Bristol, UK

Abstract. We introduce the concept of collaborative awareness as a means to enhance interoperability, resilience and self regulation in synthetic agent collectives. We discuss the theoretical, computational and engineering framework of collaborative awareness built by the EU project EMERGE, and its application to distributed robotic systems.

Keywords: awareness · distributed robotic systems · robot swarms · neural networks · dynamical systems

1 Introduction

Awareness is a fundamental aspect of any agent capable of information integration into complex and rich internal representations, across a multitude of dimensions including perceptual, introspective and social facets. In this respect, it characterizes as a fundamental concept to engineer efficient and effective distributed loosely-coupled systems of synthetic agents [6], as it caters for the ability to communicate and interoperate by exchanging bits of semantically rich information, which are the awareness states possessed by the single agents. Awareness in biological agents has converging definitions when considering local states describing content-related awareness from an agent-specific perspective. However, it becomes highly debated when it comes to global states. The issue magnifies when considering collectives of artificial agents, such as in robotic swarms. Several frameworks exist, all unsatisfactory in the limitations posed to agents' heterogeneity and disappearance of the local self into an integrated state. Ultimately, existing awareness frameworks are ineffective in explaining, facilitating,

C. Secchi and L. Marconi (Eds.): ERF 2024, SPAR 32, pp. 87–91, 2024.
https://doi.org/10.1007/978-3-031-76424-0_16

and supporting cooperative behaviours in artificial agents. The lack of a compelling theory of global awareness in AI is currently a significant barrier to the effective deployment of artificial agents in the real world. The EIC Pathfinder Challenge project EMERGE (grant n. 101070918,2022–2026)[1] is a recent 4-years research endeavour tackling this grand challenge by introducing and characterizing a novel concept of global awareness that we call collaborative awareness. EMERGE builds on a scenario comprising collectives of artificial beings with no shared language and constrained individual capabilities, which nevertheless leads to high-complexity behaviours at the collective level. Collaborative awareness becomes an emergent process supporting complex, distributed, and loosely coupled systems capable of high degrees of collaboration, self-regulation, and interoperability without pre-defined protocols. EMERGE designs a philosophical, mathematical, and technological framework that enables to know how and where to allocate awareness to optimally achieve a goal through the collective, and whose characterizing aspects are overviewed in Sect. 2. Collaborative awareness can have a transformative effect on those application domains involving delivering a service over a loosely coupled collective of virtual/physical entities, by providing enhanced resilience to high heterogeneity and compelling constraints on computation and energy. As such, we envision IoT, smart cities and transportation, microservice-based ICT systems, biomedical nanodevices, and robotics as immediate impact areas. Among them we believe that robotics provides the perfect setting to test the fundamental concepts of our framework and, in Sect. 3, we provide a brief account of how EMERGE is feeding into robotic applications involving soft robots, robotic swarms and cobots.

2 A Theoretical, Computational and Engineering Framework of Collaborative Awareness

Theoretical Framework. EMERGE defines collaborative awareness using an emergent approach where locally aware artificial agents, of possibly heterogeneous capabilities and without a shared language, develop a form of collective awareness through interacting with each other and the environment. We build on the idea that this form of awareness emerges when coordination or cooperation across a collective of task-aware agents is needed. We formalise collaborative awareness following a dimensional framework [1] representing the variations across individuals on a multidimensional space of awareness capabilities and corresponding scales. Our framework acknowledges that there are different domains where awareness needs to be defined and measured, e.g. bodily, spatial, temporal, situational, etc. It acknowledges for possible dependencies between dimensions, without imposing a unified model throughout the collective: single agents can be comparable in terms of quantity of awareness while differing on the dimensions on which they fare higher. This framework allows us to consider collaborative awareness as an addition of new dimensions and changes the way awareness is measured across dimensions, when confronted to purely local awareness states.

[1] EMERGE project web page: https://eic-emerge.eu/.

Ethical Framework. EMERGE assesses the awareness framework also from an ethics perspective, evaluating the attribution of agentive and moral responsibility to collectives having collective awareness, and characterising transparency, trust, and risks of such systems. In particular, we are exploring key ethical challenges and opportunities that could arise with the introduction of collective aware systems powered by AI into human society. Some of our early results in this sense, highlight that humans are keen to exploit benevolent algorithms with significant differential treatment of humans and machines under the same experimental scenario. Another direction of study [10] explores the key differences between human attitude about responsibility against other humans, AI agents (single and collective), and conventional, non intelligent, tools. The third dimension of our ethical framework studies transparency of AI systems and, in particular, defining and measuring awareness through explainability approaches [3]. To this end, we are building a framework that relates the different awareness types of our dimensional model with specific explainability techniques and methodologies.

Computational Framework. Collaborative awareness emerges as goal-promoted interactions unfold following transformations of perceptual states into richer awareness states. To this end, it is important to be able to identify and characterize template dynamical transformations, and their compositions, that favour emergence of collaborative awareness from local perceptions. We use nonlinear dynamical system theory, for its ability to integrate rich temporal information across scales, to establish, analyse, implement, and test a new AI framework for collaborative awareness. The underlying mathematical framework comprises a dictionary of low-dimensional template dynamical systems and ways of connecting them (*Archetypes*) so that they can be combined (in *Archetypes Networks*) to generate awareness. We complement this by mechanisms capable of inferring a map from a real-world system to its archetype representation, and viceversa. The composed Archetype Networks are not abstract and static items. Rather, they are meant to be executed and deployed in the physical world, i.e. to provide agents with the ability of developing awareness about specific aspects, from their interaction with the environment and other agents. This calls for a computational engine (the *Archetype Computing System* (ACS)) with the level of adaptivity needed to process and represent dynamically changing information. To this end, we provide the ACS with life-long learning mechanisms operating on Archetype units parameters and evolutionary mechanisms to adapt structural aspects of the Archetype Networks. The ACS provides a computing engine to perform dynamic distributed information processing by executing the Archetype Networks, seeking emergence of awareness in the attractors of their non-linear systems. Our computational framework is general enough to be able to describe both the information transformation *computed* by the body of a robot, as well as neural layer transformations. The ultimate goal of EMERGE, in this sense, is going towards the realization of a universal dynamical systems computing engine [7], strengthening the bridge between physics and computation, creating a substrate for the execution on non-conventional architectures, e.g. neuromor-

phic and morphological computing [11]. In our early results, we have shown how the Archetype framework can be used to define new recurrent neural networks designs based on oscillator-like neural units [2]. From a mapping perspective, we have looked into how to learn physical-consistent dynamics of Lagrangian systems from data [9] and how to compress high dimensional mechanical systems akin to soft robots into low dimensional neural representations [8].

Engineering Framework. To realize an actionable implementation of collaborative awareness for collectives of intelligent synthetic agents, we need an additional ingredient, that is an engineering framework that can control emergence of (the different dimensions of) awareness. Understanding and engineering awareness requires to have impartial metrics that can measure changes in behaviour based on different dimensions of awareness. However, awareness cannot be measured directly. Hence, our engineering framework builds first and foremost a scheme that will allow to evaluate each dimension of awareness in a given context for any system. We put forward a framework in which an awareness dimension is measured by evaluating the performance of the systems in *proxy tasks*, associated with specific *capacities* for the given awareness dimension. Performance is measured with respect to defined *metrics* (e.g. robustness, adaptability, speed) and tasks are chosen to allow differential evaluation with respect to capacities. Under such a setting we measure awareness through the change of task performance between systems with and without awareness capacities. Such awareness metrics will then serve two purposes: (i) they will be used by evolutionary methods to guide the evolution of agent collectives with emergent collaborative awareness; (ii) they will serve to provide human-readable insight and control over the resulting collective.

3 Collaborative Awareness in Distributed Robotics

EMERGE results find perfect match in distributed robotic systems. We consider three scenarios each corresponding to a representative areas of robotics: (i) modular soft robots, as an example of a physically distributed collective where the body needs to self-organize to account for the dynamic addition of components; (ii) robotic swarms, as an example of large scale minimal collective where agents need coordination to achieve a collaborative goal; and (iii) cobots, as a closer-to-market use case where interoperability is currently a significant barrier. *Soft robots* [4] are characterized by continuously deformable bodies inspired by invertebrates. We target a local minimal collective under the form of a modular soft squid-like robot entirely made of silicon. Its main body will contain some central intelligence, where any number of soft tentacles can be added in a modular fashion (along with their local intelligence). Minimal or no explicit information is shared between the various parts, which will autonomously learn to swim in an unknown environment. This will allow us to investigate the capability of physically distributed systems to create awareness in a scenario of contained complexity. Our *robot swarm* deployment [5] is based on an autonomous

warehouse with users taking robots out-of-the-box and needing the collective to gain collaborative awareness in order to achieve the task and communicate their awareness to the user. Swarm members will be artificially made minimal, their capabilities varied by evolutionary learning to understand the impact of complexity and heterogeneity on the emergence of collaborative awareness. *Cobots* have potential for large-scale industrial use and to allow cooperation with human workers. However, they still suffer from two significant issues: (i) limited individual intelligence and inability to autonomously adapt to situations and, (ii) lack of interoperability between systems by different providers. We will investigate if emergent awareness can solve these two challenges by putting collaborative robots in a realistic retail environment, and comprising cobots from different companies. These will interact without direct information transfer, assessing whether collaborative awareness can emerge to a sufficient extent to perform standard tasks like filling shelves with products.

Acknowledgements. This work was funded by the EU under GA 101070918, EMERGE project. UK participants in EMERGE are supported by UKRI grant number 10038942.

References

1. Birch, J., Schnell, A.K., Clayton, N.S.: Dimensions of animal consciousness. Trends Cogn. Sci. **24**(10), 789–801 (2020)
2. Ceni, A., et al.: Randomly coupled oscillators for time series processing. In: ICML Workshop on New Frontiers in Learning, Control, and Dynamical Systems (2023)
3. Cossu, A., Spinnato, F., Guidotti, R., Bacciu, D.: A protocol for continual explanation of shap. In: Verleysen, M. (ed.) Proceedings of ESANN 2023 (2023)
4. Della Santina, C., Duriez, C., Rus, D.: Model-based control of soft robots: a survey of the state of the art and open challenges. IEEE Control Syst. Mag. **43**(3), 30–65 (2023)
5. Jones, S., Hauert, S.: Frappe: fast fiducial detection on low cost hardware. J. Real-Time Image Process. **20**(6), 119
6. Jones, S., Milner, E., Sooriyabandara, M., Hauert, S.: Distributed situational awareness in robot swarms. Adv. Intell. Syst. **2**(11) (2020)
7. Kia, B., Lindner, J.F., Ditto, W.L.: Nonlinear dynamics as an engine of computation. Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. **375** (2017). https://api.semanticscholar.org/CorpusID:3711548
8. Lepri, M., Bacciu, D., Della Santina, C.: Neural autoencoder-based structure-preserving model order reduction and control design for high-dimensional physical systems. IEEE Control Syst. Lett. (2023)
9. Liu, J., Borja, P., Santina, C.D.: Physics-informed neural networks to model and control robots: a theoretical and experimental investigation (2023)
10. Longin, L., Bahrami, B., Deroy, O.: Intelligence brings responsibility - even smart AI assistants are held responsible. iScience **26**(8) (2023)
11. Marković, D., Mizrahi, A., Querlioz, D., Grollier, J.: Physics for neuromorphic computing. Nat. Rev. Phys. **2**(9) (2020)