Achieving emergent behaviour in a heterogeneous underwater robot society using multiagent reinforcement learning and embodied evolution

Anja Babić
Department of Control and Computer Engineering
Faculty of Electrical Engineering and Computing, University of Zagreb
Zagreb, Croatia
anja.babic@fer.hr

Abstract— **Behavioural approaches to robotics are becoming increasingly relevant, especially in the field of artificial life. Viewing behaviour as emerging from the robot-environment interaction and being greatly influenced by the robot's physical presence and existence offers a novel perspective on the approach to understanding, design, and control of multiagent systems. This paper gives an overview of robot societies and several methods of their classification. An introduction to the concept of embodied evolution and its application in multiagent robotic systems is presented. An overview of the general basics and interactions between the concepts of embodied evolution, distributed intelligence, and multiagent reinforcement learning is given, with some particulars on their application in the study of a heterogeneous underwater robot society. The issue of emergent behaviour is addressed, with the primary concern being understanding how emergence happens, and how it can be prompted and encouraged.**

Index Terms— **artificial life, embodied evolution, emergent behaviour, multiagent reinforcement learning, robot society**

I. INTRODUCTION

Starting in the mid-1980s, the scientific approach to the field of artificial life has been to study living systems using a synthetic "bridge": or, building life, as hardware or software or even wetware, in order to reach a better understanding of it. While intensely interdisciplinary, artificial life research can be classified into fourteen main themes: origins of life, autonomy, self-organization, adaptation (including evolution, development, and learning), ecology, artificial societies, behaviour, computational biology, artificial chemistries, information, living technology, art, and philosophy.

Reinforcement learning is a common form of machine learning inspired by behaviourist psychology, and includes adaptation occurring through agents interacting with their environment. The increasing acceptance that the contributions and interactions between the "body" of a studied agent and its environment cannot be ignored - hence, embodied evolution and cognition - is leading to a rise of artificial life, as opposed to traditional artificial intelligence, as an approach within cognitive science [1].

Early steps towards an increasingly behavioural approach to robotics and viewing behaviour as emerging from the robot-environment interaction, while relying on a grade of autonomy that makes micromanaging all aspects of a system unnecessary (and undesirable), were made in [2]. Explicitly encouraging behavioural diversity using behavioural distance/diversity mechanisms has been empirically shown to lead to substantial performance improvements in several typical evolutionary robotics experiments [3].

Several novel frameworks for studying cognitive systems have arisen from devising and implementing distributed approaches to the problem. Swarm cognition is one such framework, and it promotes the study of cognition as an emergent collective phenomenon, and as something that can be recognised in the behaviour of entire collectives, including complex societies [4].

The evolution of cooperation, frequently based on game theory, is an increasingly popular research topic [5].

This paper gives an overview of several important concepts of multiagent reinforcement learning and embodied evolution, with the proposed goal of achieving beneficial emergent behaviour in a heterogeneous underwater robot society. Section II gives some details on the classification and study of robot societies. Section III provides an overview of the intertwining concepts of embodied evolution and multiagent reinforcement learning, while Section IV goes into the problem of emergent behaviour. Section V briefly describes a sample simulated scenario featuring an underwater robot society. Finally, Section VI presents some conclusions and plans for further work.

II. ROBOT SOCIETY

A useful way of classifying robot societies is according to modes of interaction between their members, as shown in Figure 1. The most common and relevant forms these interactions take are:

- *collective*, where the robot entities are not aware of

each other, although they share goals and their actions benefit each other and the "team" as a whole. An example of this type of interaction in multi-robot systems is a robotic swarm. An individual robot in this kind of society likely has a very simple controller and structure, but by interacting with others a global goal (some sort of flocking, formation-keeping, or foraging and similar) is achieved, often as an emergent property of these many smaller local interactions.

- *cooperative*, where all members of a team of robots that is trying to accomplish a common goal are aware of each other, and their actions are mutually beneficial. Robots in a society that functions this way may work together and reason about each other's capability to contribute to the completion of a given task. At times, the individual robots may be working on different parts of the higher goal, and so need to ensure that they don't interfere with each other's work. Optimal task allocation and coordination are very important, even while the majority of the robots' operation is focused on working together towards the common goal.

- *collaborative*, where robots have individual goals, but are aware of the other robots on their team, and their actions advance both their own and others' goals. Each member of a group of robots of this type has its own agenda, all agendas are compatible with each other, and the robotic agents might be heterogeneous with regards to sensor/actuator setups and related properties. In this kind of society, focus is on how the unique expertise of the individual can contribute to another individual achieving its goal and thus the ultimate goal of the team as a whole, by bringing together agents with complementary skills and assets. Coalition formation is a notable type of collaboration.

- *coordinative*, where all robotic entities are aware of each other, but do not share a common goal, and their actions are not helpful to the rest of the team. Typically this implies several robots sharing a common workspace, and thus needing to coordinate among themselves to avoid interference and mitigate any potential negative outcomes of intra-robot interaction. Multi-robot path planning techniques and collision avoidance are a good example of this.

Additionally, the interaction between robots can be *adversarial*, i.e. the goals of some agents may have a negative effect on others. Two rival teams of robots attempting to accomplish the same task while hindering each other's progress would fall under this category, a good example of which is robotic soccer [6].
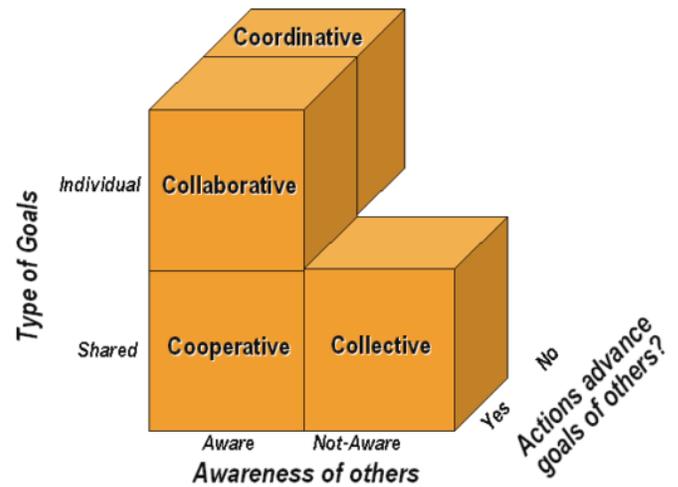

Fig. 1. Types of interactions in a distributed intelligence system, dictating the formation of multi-robot societies [6].

The key to successfully performing advanced tasks in complex modern-day environments is the emergence of cooperative abilities among robots forming a society. Two types of robot societies can be considered when evaluating group cooperation ability:

- *integrating robot societies*, consisting of a smaller number of heterogeneous individuals with highly specialised skills (biological examples include wolf packs and bird colonies)

- *differentiating robot societies*, consisting of a large number of homogenous individuals with relatively limited abilities (biological examples include bees and ants)

Both types of societies require individuals to have well-defined roles and the ability to dynamically modify their behaviour while the group is performing an assigned task [7].

The proposed subject of empirical study and application here is the multi-layer underwater robot society developed in the Horizon 2020 FET project subCULTron, shown in Figure 2. The aim of this project is achieving long-term autonomy in a learning, self-regulating, self-sustaining underwater society and culture of robots in the high-impact application area of the lagoons surrounding Venice, Italy.

The heterogeneous robotic system consists of three separate agent types, shown in Figure 3. Artificial mussels (aMussels) on the seafloor represent the bottom layer and serve as the collective long-term memory of the system, as well as sensor hubs that monitor the environment which represents the natural habitat. On the surface of the water, robotic platforms that represent artificial lily pads (aPads) interface with human society, enabling an exchange of both information and energy. Between these two layers lies the operating area of the artificial fish (aFish), which move, monitor, and explore the environment and exchange information with both the mussels and the lily pads. The sub-populations of robots perform learning algorithms on their local data, within their own layer, thus promoting separate sub-culture development and promoting diversity within the artificial ecosystem [8].

Fig. 2. Sample simulation scenario of the subCULTron collaborative heterogeneous artificial underwater society. Artificial mussels line the seafloor, while active agents, or artificial fish (shown in red) patrol around them. Artificial lily pads supervise from the surface.

Fig. 3. Robot members of the subCULTron underwater society. From left to right: aMussel, aPad, aFish.

III. EMBODIED EVOLUTION AND MULTIAGENT REINFORCEMENT LEARNING

While the field of evolutionary robotics typically studies the use of evolutionary algorithms in an *offline* fashion - i.e. running a set of algorithms and optimising the robots' controllers in simulation software before any transfer to actual robotic hardware, use, and deployment, and without any additional adaptation as the robots operate - the concept of embodied evolution, espoused by the embodied evolutionary robotics subfield, focuses on implementing evolutionary algorithms on the hardware itself and having them run during the robots' operational period - an *online* approach. A marked advantage to this is giving the robotic system a chance to acquire beneficial behaviours as it functions, without a human in the loop, thus giving it an ability to adapt to previously unknown environments or environmental conditions that change through time [9].

The field of robotics has been the most active in the development of embodied evolution algorithms mainly due to the fact that the intrinsically adaptive and self-organising properties of the algorithms make them highly appropriate for use in real-time autonomous systems that a group of robots

presents. There are two main approaches that have arisen in the study of embodied evolution:

- *encapsulated embodied evolution algorithms*, where each individual agent carries an entire population of controllers upon which an independent evolutionary algorithm is run
- *distributed embodied evolution algorithms*, where each individual agent in the population carries only its own genotype, necessitating extensive interaction between agents

Distributed embodied evolution algorithms offer a much bigger potential when it comes to the emergence of a variety of self-adaptive and cooperative behaviours, due to the complex interactions that can occur within the population. Their drawback is that they converge less easily and require larger populations of robots to do so at all. They are also very sensitive to their configuration parameters, requiring close task-dependent regulation to ensure that a valid solution is reached [10].

Embodied evolution can also be described as taking place in a population of robots where selection, evaluation, and reproduction, the main mechanisms of evolutionary algorithms, are carried out by and between the robots, all in a distributed, asynchronous, and autonomous manner [11].

Thus, this approach seems naturally suited for work with multi-agent systems attempting to solve complex but structured problems via decentralised collaboration, as there exists a need in multiagent systems to coordinate local policies of each agent with their restricted capabilities to achieve a system-wide goal. The presence of innate uncertainty in this type of system adds to the complexity of this task, as agents need to learn unknown environment parameters while forming these local policies in an online fashion [12].

Although the agents in a system can be pre-programmed with a set of behaviours designed in advance, the learning of new behaviours online is often necessary to ensure the gradual improvement of the entire system's performance [13]. Once again, the main issues that arise when considering a more offline approach are a dynamically changing, complex, and not entirely known environment, making it a very real possibility that a hardwired behaviour may at one point become inappropriate or even outright negatively affect performance. The benefits of multiagent reinforcement learning arise primarily from the distributed nature of the multiagent system, and include a certain computational speedup made possible by parallel computation, and positive effects of sharing experience between multiple agents taking the form of communication, teaching, and imitation.

Beyond challenges that are inherited from classical single-agent reinforcement learning, such as the problem of dimensionality and the need to strike a balance between exploration and exploitation (or task completion) by careful reward selection, a multiagent approach also engenders difficulties in the form of a need for coordination and the challenge of specifying a learning goal [14]. The presence of multiple agents in the same environment, learning in parallel, complicates matters considerably, as the action space scales exponentially with the number of agents. A notable consequence of this is that some standard learning techniques that store a reward value for every possible state-action combination become unfeasible. Another issue is that the behaviour of one agent influences the outcomes of other agents' individually selected actions, thus incurring change in the environment and possibly compromising convergence [15].

The design of the reward function is a critical part of achieving a real-world task by using reinforcement learning and, although it may appear straightforward, it can demand careful tuning when multiple rewards are present [16]. An evolutionary approach to multiagent reinforcement learning includes the so-called intrinsically motivated framework, based upon several psychological theories of motivation, in which agents themselves find appropriate intrinsic rewards that implicitly help task success or affect the fitness of the agent (while striking a balance between rewarding exploration and task completion, the previously mentioned common problem in reinforcement learning design) [17]. Evolutionary algorithms also optimise meta-properties in reinforcement learning by influencing the selection of behaviours, modulating the efficiency of learning by affecting learning meta-parameters, or changing and forming entirely new reward signals that guide the reinforcement learning process [18][19].

Designing an online evolutionary algorithm distributed among a fixed population of autonomous robots for the purposes of long-term survival and operation in a given environment greatly depends on the chosen fitness function for the environment-driven distributed evolutionary adaptation. The implicit nature of this fitness function is the consequence of two complementary or conflicting motivations that are present in agents:

- *extrinsic motivation*, which arises purely from the interaction between an agent and its environment (with or without other agents) and states that an agent must maximise its chances of survival while working within environmental constraints and overcoming environmental challenges
- *intrinsic motivation*, which arises from the agent itself, and which states that, in keeping with the evolutionary process, a certain set of parameters (or genomes, as they are understood here) must spread across the agent population in order to survive, meaning the genomes promote maximising the number of agents met and interacted with (thus maximising their own opportunities for proliferation).

In a situation where the motivations are well-balanced and the environment-driven distributed evolutionary adaptation algorithm is an efficient one, a state of equilibrium should be achieved in which the genome that can be considered optimal reaches maximum spread while preserving maximum possible survival efficiency. While there would appear to be inherent conflict between these two motivations, it is possible for them to correlate enough so that they can be treated as one motivation, greatly reducing the complexity of the problem. The application of these motivations and a focus on the effect

of the environment on the evolutionary adaptation process in embodied evolution is an important concern, since embodied evolution concerns itself primarily with a fixed number of physical agents meant to operate in real-world environments - meaning their most basic operation can be severely impacted by obstacles or energy constraints [20].

An important thing to note is that the study of the development of robot controllers or robot morphologies via the use of artificial evolution does imply a certain need to rely on simulation environments during, at the very least, early development, since the duration of a single experiment can be impractically long, and the necessity of trial-and-error interactions with the environment during attempts at reinforcement learning can lead to damage of actual physical robotic platforms [21].

However, caution needs to be exercised while working within simulated environments, as the generation of simulators introduces a strong bias into the system by way of using symbols and models to represent real phenomena the way they are perceived by the person developing the simulation platform. Additionally, simulating the interaction between the robots and their environment can be very computationally demanding, especially in the case of multiagent systems, thus necessitating considerable simplification to start with. This clashes with some of the fundamental attributes of the evolutionary approach, specifically that it encourages the autonomous generation of robotic systems while leaving out a considerable amount of designer bias, and presents, theoretically, an open-ended approach, while allowing the generation of controllers that are optimally suited for a given operational environment. In any case, the reality gap, or the substantial difference between simulation and reality, persists, and needs to be taken into account when considering robotic behaviour generation.

Closing the reality gap would mean that there exists no difference between simulation and reality, from the viewpoint both of the internal agent - the robot controller - and the external observer - the experimenter evaluating the generated behaviour. This is not possible to completely achieve when doing real-world experiments, so care must be taken to ensure algorithms and desired behaviours are robust with regards to the reality gap, for example by modifying experimental setups to compensate for the reduction in stability that occurs with a decrease in population (as a population of physical robots will almost always be smaller than a simulated one) [22].

IV. EMERGENT BEHAVIOUR

Since it is hard to arrive at a satisfactorily encompassing definition, the term *emergent* is frequently described as something that is "more than the sum of its parts". In a robotic context, emergence is taken to mean that a robot's behaviour has become something not explicitly defined in its controllers, but something that has arisen as a consequence of its interaction with its environment.

The term reactive control refers to the coupling of perception and action in a way that ensures robots respond in a timely manner to moving and working in dynamic, unstructured, and at least partially unknown environments. It also represents a behavioural approach to robotic applications, which is more suited to the complex online context than the classical hierarchical and decision-based approach [23].

Some points in the design of an autonomous robot system that are deemed favourable in encouraging emergence are [24]:

- Any emergent behaviour should arise out of a large number of parallel and not too closely coupled processes
- The system must be redundant
- The system needs to employ principles of self-organisation
- The system should employ cheap and simple design as much as possible, and exploit its environment
- Intelligence should be regarded in the context of sensory-motor coordination

It is clear that multiagent systems provide excellent ground for studying emergent behaviour, especially easily monitored and highly convenient multiagent simulations, which constitute a very good place for the study of emergence, as they allow for easy highlighting and discovery of phenomena that would signify emergence has occurred [25].

A group of robots can be seen as embodying the "more than the sum of its parts" concept when it becomes a robotic team: once it shows some degree of specialised aptitude of performing a task cooperatively, in the sense that the group provides better performance than its individual components would, by taking advantage of its distributed sensing and acting capability to carry out complex tasks while also taking into consideration increased fault tolerance thanks to agent redundancy and group cohesion obtained from formation-keeping algorithms and related trajectory calculations and motion planning [26].

Another emergent phenomenon is collective intelligence, the result of two or more agents engaged in global behaviours - meaning an intelligent multi-robot system arises from a group of mobile robots that cooperate, communicate, and dynamically reconfigure their group during their attempts to solve a complex task [7].
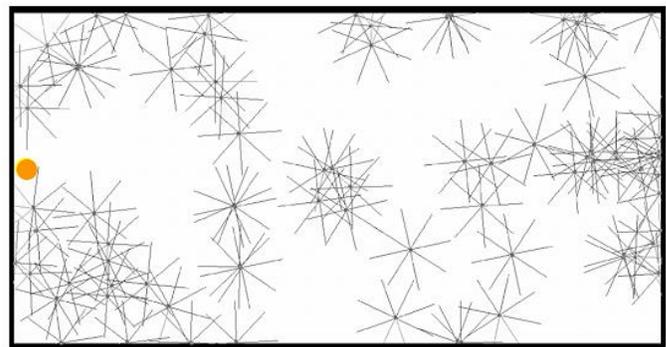

Fig. 4. A bird's eye view of the arena during the "two suns" experiment. The current sun is shown in orange. Variants of agent grouping are clearly visible [20].

It is good to note that emergence can have both a positive or a negative effect on system performance and task completion. An example of emergence in a multiagent system is the emergence of consensus, which is not always necessarily

explicit. To illustrate this, the so-called "two suns" setup experiment, shown in Figure 4., is appropriate: a population of robotic agents with limited communication ranges and with no task but the drive to "mate" is placed in an empty environment - an arena without any obstacles - where a "sun" is present. The sun itself offers no advantage to any of the agents, and it doesn't affect their energy levels (their main means of survival) in any way. In essence, there is no explicit and immediate benefit to heading towards or being close to the sun. Each agent gets information about its own distance and relative angle to the sun. Every 50 generations (or iterations of the evolutionary algorithm that the agents are running), the sun disappears and appears at the opposite end of the testing arena.

In this setup, the sun could be considered a distraction or could be made into a useful feature of the environment. Indeed, the outcome of several experimental runs supports this: the majority of agents stand very close to the sun, while others form groups at the opposite end of the arena, furthest from the sun - the implication is that two kinds of consensus have emerged, a desire to go towards the sun, and a desire to move away from the sun. A general agreement to go towards the sun makes sense when looked at through the lens of mating strategy (and thus genome proliferation and, indirectly, conservation) - adopting this strategy ensures that several agents will meet at a defined

and robust location in the environment, easily exchanging information and increasing the chance that their genome will survive [20].

V. VEHICLE ROUTING SCENARIO EXAMPLE

As a collaboration between the LABUST and LARICS laboratories of the University of Zagreb, Faculty of Electrical Engineering and Computing, a sample simulated scenario using the underwater robot society described in Section II was implemented during the early stages of the subCULTron project, using a coalition-based approach described in [27]. The proposed proof-of-concept scenario includes two aPads and ten aMussels operating in parallel. The aMussels work as sensor hubs, anchored to the seafloor, until their battery level falls under a pre-set threshold. When this occurs, they surface and immediately send out requests for charging by broadcasting their current coordinates, determined via GPS. The supervising aPads collect charging requests and run an optimisation algorithm for task allocation, with the goal of determining which aMussel should be charged by which aPad and in what order. Once this decision is made, both aPads will commence operation and visit each of their newly assigned aMussels in turn.

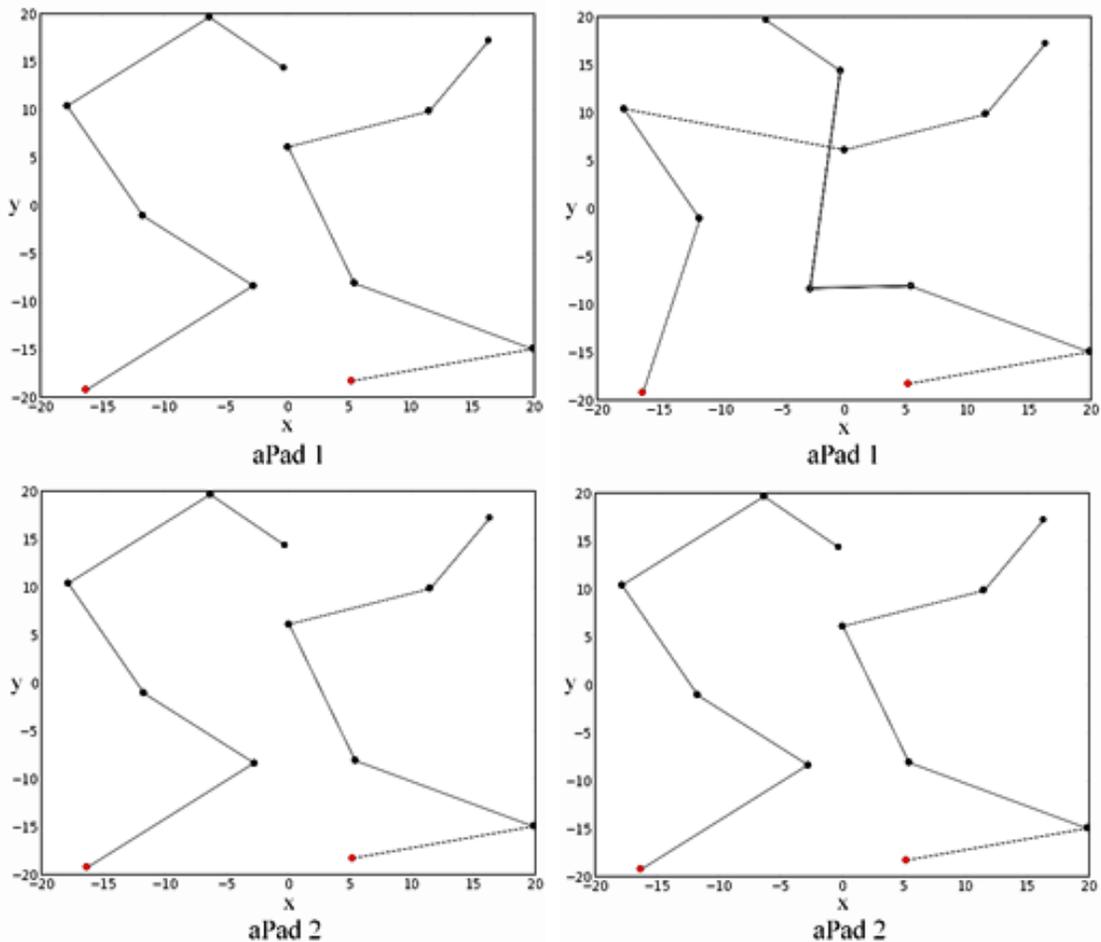The optimisation goal is minimising the distance travelled by

Fig. 5. Vehicle routing solutions reached by the two aPads: initial agreement (left) and disagreement (right). Red markers represent starting aPad positions, and black markers show the reported aMussel positions. No movement will occur until all aPads have reached a consensus about the route (plotted using black lines), avoiding conflict and ensuring that each of the aMussels is visited exactly once.

the aPads (in the interest of conserving energy) and maximising the number of aMussels served by each individual aPad. This optimisation based on a blend of genetic algorithms and reinforcement learning is bolstered by the agents (in this case, aPads) sharing acquired knowledge and individually reached solutions *before* committing to a final task allocation, ensuring that the chosen solution is the best one currently present in the society (with the ultimate goal being, of course, achieving a globally optimal solution). An example of the final decision is given in Figure 5.

The coalition-based metaheuristic combines principles of evolutionary algorithms, distributed artificial intelligence, and reinforcement learning, and adopts a fully decentralised approach to control. Members of the coalition are agents with a capacity to make decisions, learn, and cooperate directly on a peer-to-peer basis. Each agent manages a current solution by itself, and performs a search by applying standard intensification and diversification operators. Operator selection is done by roulette wheel selection, and the agent's behaviour during this search is constantly adapted using a reinforcement learning mechanism. Solution sharing, including the final validation, is an important aspect of the cooperative approach. This type of system structure also ensures a greater degree of robustness, since an agent being removed or added does not affect the global performance of the system.

In the current implementation, all aPads have access to the same input data. Various alternations, such as diversifying agents by way of giving them different communication ranges or manually inserting noise into some or all of the communication channels, will likely lead to an emergence of several interesting behaviours.

VI. CONCLUSION

The concepts of embodied evolution, multiagent reinforcement learning, and distributed control interact and even overlap in a myriad of interesting ways. The study of emergent behaviour is a relatively new and promising approach to understanding cognitive systems, with one of the primary methods of exploring the field being angling towards an understanding of how emergence can be encouraged, prompted, or even outright caused.

Future research will focus on the underwater robot society described earlier, first by modifying the representation of the environment or the representations of the agents themselves in order to achieve greater behavioural diversity and open-endedness. Situations in which one or more of the agents experience technical failure of one or even several of their systems (whether core or peripheral) are of particular interest, since they would provide an excellent test of adaptability and robustness granted by a distributed and behavioural approach to system control and robot controller design, as well as creating an opportunity for interesting new avenues of cooperation between similar or diverse robots. The heterogeneity of the system has yet to be truly taken advantage of in a behavioural context, opening up a wide array of novel and potentially unpredictable interplay.

Furthermore, an empirical approach using the tenets of embodied evolution is considered highly promising, as testing on actual physical systems will undoubtedly provide valuable insight into the simulation-reality differences present in the proposed society, as well as an opportunity to study novel behaviours that might potentially emerge from a combination of agent-agent and agent-environment interactions.

REFERENCES

[1] W. Aguilar, G. Santamaria-Bonfil, T. Froese and C. Gershenson, "The past, present, and future of artificial life", *Frontiers in Robotics and AI*, vol. 1, 2014.

[2] R. Brooks, "Intelligence without representation", *Artificial Intelligence*, vol. 47, no. 1-3, pp. 139-159, 1991.

[3] J. Mouret and S. Doncieux, "Encouraging behavioral diversity in evolutionary robotics: an empirical study", *Evolutionary Computation*, vol. 20, no. 1, pp. 91-133, 2012.

[4] V. Trianni, E. Tuci, K. Passino and J. Marshall, "Swarm Cognition: an interdisciplinary approach toÂ theÂ study of self-organising biological collectives", *Swarm Intell*, vol. 5, no. 1, pp. 3-18, 2010.

[5] M. Burtsev and P. Turchin, "Evolution of cooperative strategies from first principles", *Nature*, vol. 440, no. 7087, pp. 1041-1044, 2006.

[6] L. Parker, "Distributed intelligence: Overview of the field and its application in multi-robot systems", *Journal of Physical Agents*, vol. 2, no. 1, pp. 5-14, 2008.

[7] E. Pagello, A. D'Angelo, C. Ferrari, R. Polesel, R. Rosati and A. Speranzon, "Emergent behaviors of a robot team performing cooperative tasks", *Advanced Robotics*, vol. 17, no. 1, pp. 3-19, 2003.

[8] Subcultron.eu, "subCULTron", 2015. [Online]. Available: http://www.subcultron.eu/. [Accessed: 05- Feb- 2016].

[9] E. Haasdijk, N. Bredeche and A. Eiben, "Combining environment-driven adaptation and task-driven optimisation in evolutionary robotics", *PLoS ONE*, vol. 9, no. 6, p. e98466, 2014.

[10] P. Trueba, A. Prieto and F. Bellas, "Distributed embodied evolution for collective tasks", Proceeding of the fifteenth annual conference companion on Genetic and evolutionary computation conference companion - GECCO '13 Companion, 2013.

[11] R. Watson, S. Ficiei and J. Pollack, "Embodied evolution: embodying an evolutionary algorithm in a population of robots", *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406)*.

[12] W. Teacy, G. Chalkiadakis, A. Farinelli, A. Rogers, N. Jennings, S. McClean and G. Parr, "Decentralized Bayesian reinforcement learning for online agent collaboration.", in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, 2012, pp. 417-424.

[13] G. Weiss, *Multiagent Systems*. Cambridge, Mass.: MIT Press, 1999.

[14] L. Busoniu, R. Babuska and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning", *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 2, pp. 156-172, 2008.

[15] J. Kok and N. Vlassis, "Collaborative multiagent reinforcement learning by payoff propagation.", *The Journal of Machine Learning Research*, no. 7, pp. 1789-1828, 2006.

[16] E. Uchibe and K. Doya, "Finding intrinsic rewards by embodied evolution and constrained reinforcement learning", *Neural Networks*, vol. 21, no. 10, pp. 1447-1455, 2008.

[17] M. Schembri, M. Mirolli and G. Baldassarre, "Evolving internal reinforcers for an intrinsically motivated reinforcement-learning robot.", in *Development and Learning, 2007. ICDL 2007. IEEE 6th International Conference on*, 2007, pp. 282-287.

[18] S. Singh, R. Lewis, A. Barto and J. Sorg, "Intrinsically motivated reinforcement learning: an evolutionary perspective", *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 2, pp. 70-82, 2010.

[19] S. Elfwing, "Embodied Evolution of Learning Ability", Doctoral thesis, KTH, School of Computer Science and Communication (CSC), Computer Vision and Active Perception, CVAP., 2007.

[20] N. Bredeche, J. Montanier, W. Liu and A. Winfield, "Environment-driven distributed evolutionary adaptation in a population of autonomous robotic agents", *Mathematical and Computer Modelling of Dynamical Systems*, vol. 18, no. 1, pp. 101-129, 2012.

[21] N. Jakobi, P. Husbands and I. Harvey, "Noise and the reality gap: The use of simulation in evolutionary robotics", *Advances in Artificial Life*, pp. 704-720, 1995.

[22] J. Zagal, J. Ruiz-del-Solar and P. Vallejos, "Back to reality: Crossing the reality gap in evolutionary robotics.", in *IAV 2004 the 5th IFAC Symposium on Intelligent Autonomous Vehicles*, Lisbon, Portugal, 2004.

[23] E. Billing, "Cognitive perspectives on robot behavior", in Proceedings of the 2nd International Conference on Agents and Artificial Intelligence: Special Session on Computing Languages with Multi-Agent Systems and Bio-Inspired Devices, Portugal, 2010, pp. 373-382.

[24] R. Pfeifer and C. Scheier, *Understanding Intelligence*. Cambridge, Mass.: MIT Press, 1999.

[25] D. David and R. Courdier, "See emergence as a metaknowledge. a way to reify emergent phenomena in multiagent simulations", in *International Conference on Agents and Artificial Intelligence (ICAART'09)*, Porto, Portugal, 2009, pp. 564-569.

[26] D. Kurabayashi, "Toward realization of collective intelligence and emergent robotics", IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.99CH37028).

[27] D. Meignan, A. Koukam and J. Créput, "Coalition-based metaheuristic: a self-adaptive metaheuristic using reinforcement learning and mimetism", *Journal of Heuristics*, vol. 16, no. 6, pp. 859-879, 2009.