# Promoting Originality in Online Swarm Robotics

Amine Boumaza
Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
amine.boumaza@loria.fr

Published in: *Amine Boumaza. 2023. Promoting Originality in Online Swarm Robotics. In Genetic and Evolutionary Computation Conference Companion (GECCO '23 Companion), July 15–19, 2023, Lisbon, Portugal. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3583133.3590563*

**Abstract**

We address the problem of promoting diversity in online embodied evolution of heterogeneous robot swarms. We argue that it is not easy to adapt existing diversity algorithms from traditional evolutionary robotics to this context and describe a method in which selection is based on originality and which allows a swarm of heterogeneous agents to maintain a high degree of diversity in behavioral space. We also describe a behavioral distance measure that compares behaviors in the same conditions to provide reliable measurements in online and distributed contexts. We test the selection scheme on an open-ended survival task and show its effectiveness. Without any other pressure besides that of the environment, the evolved strategies tend toward simplicity, exploiting the existing affordances. An additional external pressure enables the emergence of rich and diverse behaviors.

## 1 Introduction

The goal of embodied evolutionary robotics (EER) is to design collective behaviors for swarms of heterogeneous agents evolving online [7]. In a nutshell, these are algorithms in which evolution is decentralized, where each agent, typically a mobile robot, runs an EA on board and exchanges genetic material with other agents when they meet. Selection and variation are performed locally by the agent, in contrast to traditional evolutionary robotics (ER), where there is a central process that governs evolution. These algorithms have been studied in different contexts and have been shown to be robust in open-ended environments [2].

In the present work, we address the problem of promoting behavioral diversity. Although extremely effective in traditional ER, the use of quality diversity (QD) algorithms [17, 15] in online collective robotics raises some challenges. For instance, they rely on centralized archives that are incrementally filled with newly discovered behaviors, whereas EER algorithms are decentralized and emphasize local information exchange with no history maintenance. Furthermore, EER algorithms operate online in dynamic environments where individual behaviors result from the interactions between agents and other agents, which makes the comparison of behaviors difficult. There have been few attempts

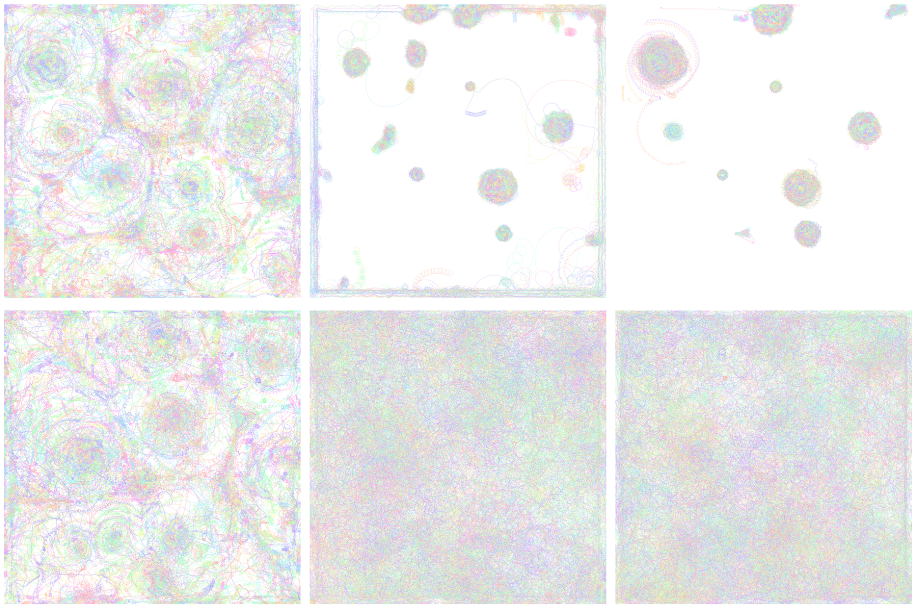

Figure 1: Snapshots of agents trajectories from a typical run with 1000 agents. From left to right beginning, middle and end of evolution. Random selection (top row), behavioral diversity (bottom row). Colors have no meaning except to distinguish agents.

to promote diversity in collective robotics. Some relied on centralized off-line ER settings where all agents are clones [9], or needed task-specific behavioral descriptors [12] and others used *ad hoc* schemes to favor the least prevalent behaviors [11]. Other authors addressed the problem from the perspective of specialization, where subsets of the agents solve different tasks [16, 14].

In novelty-base search [13], the goal is to discover solutions that exhibit new, unseen behaviors. New solutions are compared to ones in an archive, and if they are original enough, they extend and/or replace older solutions in the archive. In EER algorithms, selection is performed locally and independently on each agent, which rules out the use of archives. For that, we propose to slightly redefine the notion of novelty. In traditional algorithms, diversity usually refers to originality throughout the course of evolution, whereas in our context, since we do not keep track of history, diversity should refer to originality at a given point in time. The selection scheme we propose aims to select solutions at the agent level by choosing the ones that have the "most different behavior" from all other solutions.

That said, there are still additional challenges in choosing the most different solution. Comparing behavior in the context of EER is not as straightforward as it is in more traditional ER settings, where it is usually based on behavioral descriptors. It can only be measured online during the lifetime of agents, and since they live specific experiences depending on their encounters, it is difficult to generalize a measurement outside of the situation it was measured in. The selection scheme we propose here promotes diversity by selecting solutions based on their originality in the space of behaviors. It addresses the challenges mentioned above and respects the online and distributed nature of EER algorithms by measuring generic behavioral descriptors locally on agents.

To illustrate the main idea, we choose a simple open-ended survival task where agents must evolve strategies that allow them to spread their genetic material in order to survive and avoid mass extinction. We will show that without additional selection pressure except that of the environment, robots exhibit behaviors that tend to be similar and that seem to be local minima, exploiting certain features of the environment [1]. These behaviors are sufficient to overcome the environmental pressure and survive. However, with the additional selection pressure towards originality, it is no longer the case; agents evolve other behaviors and manage to also survive.

# 2 Methods

We consider the Minimally Environment-driven Distributed EA (mEDEA) [3], which as its name indicates, is an open-ended evolutionary algorithm where the goal is survival. An intrinsic aspect of EER algorithms in general and of mEDEA in particular is the fact that selection pressure acts at different levels. To solve a given task, one can design a fitness function to drive evolution toward that goal. Furthermore, there is a second force that comes into play, which is the pressure linked to the environment and that influences the agents' ability to spread their genomes. Those genomes that do not maximize the number of mating opportunities become extinct, and only those that do so survive.

To overcome the challenges in estimating the behaviors of the agents, we propose to estimate the behavior of the genomes regardless of their agents' experiences. Our goal is to identify genomes with the most different behavior in the agent's selection pools at the time of selection. Therefore, we propose to simulate behaviors, measure their differences on fake sensory data and then use these measurements as a selection criterion.

To be more precise, let us focus on one agent, let it be agent $a$. We note $l^a = \{\mathbf{x}_1 \ldots \mathbf{x}_\mu\}$ the set of genomes it collected during its lifetime. Our goal is to identify $\bar{\mathbf{x}} \in L^a$ that has the most different behavior. This genome is defined as the one that is the farthest away from all the other genomes in the list, according to some metric on the space of behaviors.

We call $\mathcal{S} = [0,1]^s$ and $\mathcal{M} = [-1,1]^m$ respectively the sensory space ($s$ sensors) and the motor space ($m$ actuators). Furthermore, we note $g : \mathrm{I\!R}^N \times \mathcal{S} \rightarrow \mathcal{M}$ such as, $\mathbf{O} = g(\mathbf{x}, I)$ is the function that computes the output $\mathbf{O}$ of a neuro-controller with weights $\mathbf{x}$ on input $\mathbf{I}$. For clarity, we will consider $g$ as taking a subset of $\mathcal{S}$ and generating a subset $\mathcal{M}$, whose size is a parameter.

At the selection step, agent $a$ creates the set $\mathbf{S} \subset \mathcal{S}$ consisting of $K$ samples chosen uniformly from $\mathcal{S}$; the "fake" inputs. It then computes $\mathcal{B}_{\mathbf{x}_i} = g(\mathbf{x}_i, \mathbf{S})$, the "fake outputs", for all the genomes in its list. We consider $\mathcal{B}_{\mathbf{x}_i}$ as the behavior descriptor of genome $\mathbf{x}_i$. Finally, the agent has to select among $\{\mathcal{B}_{\mathbf{x}_1}, \ldots, \mathcal{B}_{\mathbf{x}_\mu}\}$ the behavior that is the farthest away from all the others or

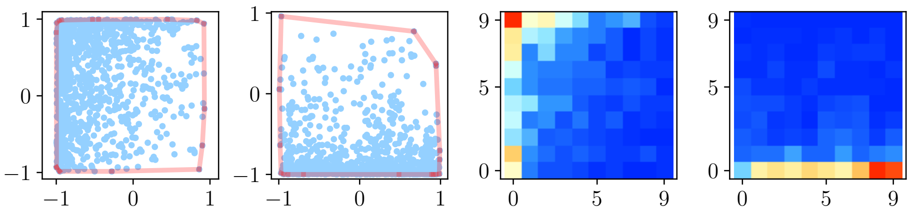


Figure 2: Raw output values (left) and behavioral maps (right)

the one with the largest average distance to all others, given some metric on $\mathcal{M}$:

$$\bar{\mathbf{x}} = \arg\max_{\mathbf{x} \in L^a} \left( \frac{1}{|L^a|} \sum_{\mathbf{y} \in L^a} D\left(\mathcal{B}_{\mathbf{x}} - \mathcal{B}_{\mathbf{y}}\right) \right).$$

We consider behavioral descriptors $\mathcal{B}_{\mathbf{x}_i}$ as samples from a certain distribution on $\mathcal{M}$, which depends on $\mathbf{x}_i$. The distance $D$ above can be any distance between distributions, and in this paper we chose it to be the Jensen-Shannon distance [6]. It is defined as:

$$D_{PQ}^2 = 2H\left(\frac{1}{2}(P+Q)\right) - H(P) - H(Q), \tag{1}$$

where $P$ and $Q$ are the distributions to be compared and $H$ is the Shannon entropy. In practice, we use samples from $P$ and $Q$ which in our case, are histograms of the values of $\mathcal{B}_{\mathbf{x}_i}$ (the behavioral maps). Figure 2 illustrates this process. In there, two neuro-controllers with distinct sets of weights are given the same set of randomly generated input values. The distance between these two behaviors is estimated using equation 1 using the distribution on the right side of the figure.

## 3 Experiments

In the following experiments, we consider the same setting as in [1], where a swarm of agents evolves in an empty square arena. There are 10 randomly placed landmarks that agents can perceive; these however do not provide any advantage or disadvantage to agents and can be considered useless. We do not consider a specific task to solve but rather perform an open-ended evolution and observe the behaviors that emerge in different settings.

The experiments[1] were performed on the Roborobo simulator [5], an environment that allows us to run experiments on large swarms of agents[2]. Agents are e-puck-like mobile robots with limited-range sensors and two differential drive wheels. Sensors are placed at 12 locations around the agent's body to perceive obstacles and other agents. Two additional sensors provide the distance and orientation of the closest landmark with an unlimited range. Finally, agents can move using 2 motors and, the neuro-controller we consider here is a simple

[1]The code can be downloaded at: `https://gitlab.inria.fr/boumaza/public-code`
[2]Roborobo4 at commit `e5319de6488b4b1b6ec593e4c4fb42cf71ec2de4`

Table 1: Simulation parameters.

| | | | |
|---|---|---|---|
| Arena size | $1000^2$ pix. | Nb. agents | $\{400, 1000\}$ |
| Nb. landmarks | 10 | Lifetime | 400 tics |
| Sens./com. range | 24 pix. | Nb. Gens | 1000 gen |
| Robot diam. | 6 pix. | Init. step-size | 0.25 |
| Max trs. vel. | 2 pix. / tic | Adapt. s.s rng | (0.01, 0.5) |
| Max rot. vel. | 30 deg / tic | Behav. maps | $10^2$ |
| Init. weights | $[-1, 1]$ | Fake inputs | 128 |

feed-forward perceptron with a hyperbolic tangent activation function. All the parameters of the experiments are summarized in Table 1 and all experiment where repeated 32 times.

We devise two sets of experiments, each with a different population size: a small one with $\lambda = 400$, and a large one with $\lambda = 1000$. Both population evolve in the same environment, and we consider a long evolution time of 1000 generations. For each experiment we test both selection schemes and compare them on different aspects. The number of active robots indicates how many agents were successful at gathering genomes and managed to survive. To assess if robots were successful at spreading their genomes, we measure the number of transmissions and the number of effective inseminations. Finally, a simple quantity that contributes in assessing agents' behaviors, is the range or the maximum distance traveled from the starting position during one generation.

To observe the evolved behaviors, we record at different times of the evolution, the trajectories of the robots over the course of a few generations. In addition, we estimate their complexity in terms of information (Kolmogorov complexity) using a lossless compression[3] algorithm [10]. Rich and diverse trajectories will tend to have a larger complexity than ones with more regularity.

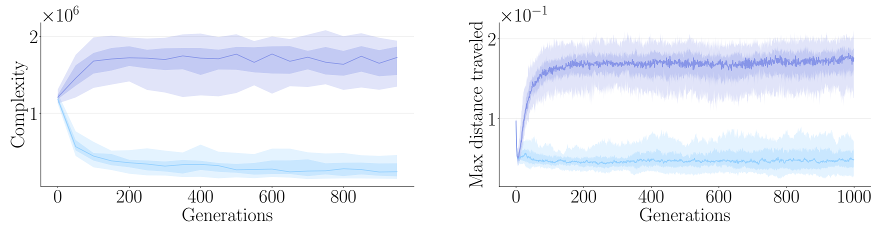


Figure 3: The complexity of the behaviors (left). The maximum traveled distance (right). Median values and the range between the $25^{\text{th}}$ and the $75^{\text{th}}$ percentiles. Light blue corresponds to random selection and purple to diversity selection.

[3] The compression algorithm is a variant of the LZ77 (Lempel-Ziv) implemented in the zlib library

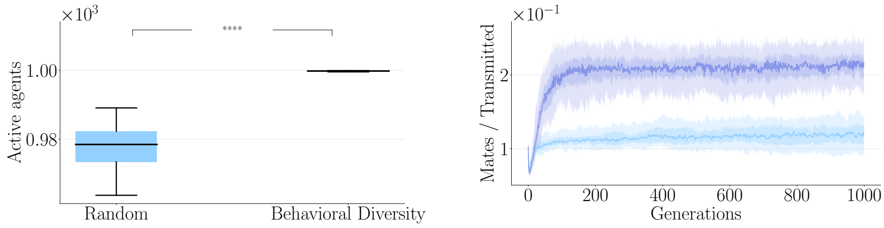


Figure 4: The median number of active robots at the end of their lifetime (left). The ratio of inseminations over transmissions (right).

## 4 Results and discussions

We start by discussing the trajectories realized by the agents. Figure 1 shows snapshots at various times. We clearly see that in the early stages of evolution, both selection schemes evolve similar behaviors. Agents tend to gather at landmarks, and close to walls. They also tend to orbit around landmarks and travel between them. There are what appear to be lines between landmarks, akin of a Voronoi tessellation. These behaviors are consistent with those described by [4, 1]. When we advance in time, things change a bit. With random selection, agents reinforce their gathering behavior to end up clustered and tightly packed towards the end. On the other hand, with diversity selection, agents seem to travel all around the arena. Taking a closer look, we can notice that the robots still exploit the landmark, but differently. They are not used as gathering places but rather as way-points where robots cross each other's paths. In terms of complexity, Figure 3 shows that indeed the information contained in the trajectories generated in the case of diversity selection is higher than that of random selection. In other words, there is more randomness and less regularity.

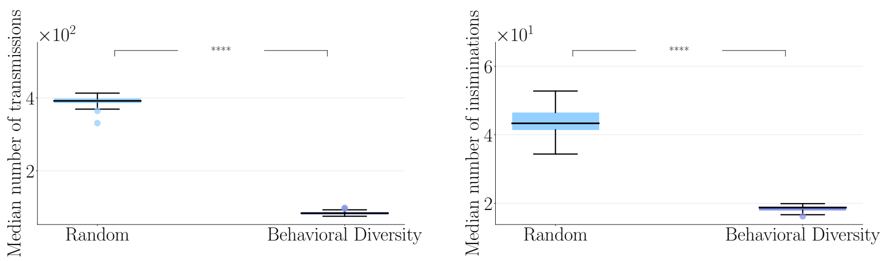


Figure 5: The number of transmissions (left) and the number of inseminations (right) for 1000 agents.

It is interesting to observe, at least in the case of random selection, that agents end up gathering at different places in the environment. In fact, this is an efficient strategy for survival where agents are guaranteed a longtime supply of genetic material, but the choice of meeting points could seem surprising.

A likely explanation could be the following. In the absence of all other information except for what can be perceived (obstacles, agents, landmarks), finding and agreeing on gathering places is not an easy task. The fact that this strategy evolves suggests that agents end up taking advantage of what their environment has to offer, or what J. Gibson named the *affordences* [8]. These features of the environment act as drivers for evolution to select the behavior we observe.

Landmarks are perceived from faraway, and in the absence of any other features (obstacle or robot), they are the only activity recorded on the sensors. Furthermore, landmarks are the only constant stimulus, whereas, obstacles and other agents are intermittent. They provide the only exploitable gradient of information (in the sens of Gibson's optical flow). This may explain, why landmarks or information about them, plays an important role in shaping the agents' behavior. Landmarks do not provide any advantage, but evolution found an effective use for them as meeting places, the same goes for corners and walls. Environmental pressure is overcome by this strategy, and in the absence of other forces, the strategy persists.

The "subjective" visual observations above, are further supported by a few objective quantities. In the case of random selection, the number of active robots indicates that survival is achieved, except for a negligible few. Furthermore, the distance traveled also suggests that agents do not move very far. On the other hand, this is not the case with diversity selection, agents travel longer distances (Figure 3) and somehow manage to survive event better (with less variance) than in the case of random selection (Figure 4 left).

A likely explanation is one that suggests selection pressure (from the selection scheme) prevented the fixation of gathering behaviors and drove evolution to discover other strategies, e.g. travel and search for partners. This is further supported by the number of mates and the number of transmissions (Figure 5). In the case of random selection, agents have higher numbers of transmissions and mates, than with diversity selection, which can be explained by their gathering strategy. But again, the fact that agents survive in the case of diversity, suggests that they evolved strategies that achieve what is necessary, in terms of mating, to survive; they overcame the pressure from the environment. This is consistent with results on energy foraging, where agents only forage what is needed to survive [4]. We may even further argue that they evolved a more efficient mating strategy. Indeed, if we look at the ratio of the number of mates over the number of transmissions (Figure 4 right), we see that they are twice as much more efficient at mating than in the random case.

One could argue that gathering strategies might be simple and easily attainable behaviors, and to some extent, can be considered as local minima in the space of behaviors. In our experiments, both selection schemes started evolving strategies that tended in that direction; however, in the case of diversity selection, these strategies were not retained. The additional force from the selection pressure pushed agents to diverge and breakaway from clusters (which are genetically close) and search for other strategies that were still as efficient in terms of survival.

On a final thought, if in our experiment, evolution stumbled easily on strategies that exploit the affordences of the environment, could that always be the case? And if not, what would be the necessary conditions for that? Could we design environments that provides specific affordences that drive evolution in a certain direction? Investigating these questions seems to be very interesting.

## 5 Conclusions

Promoting behavioral diversity is a major issue in ER, and designing efficient algorithms is a hot topic. In this paper, we address this question in the context of EER, where it is also a major problem, as many unsolved questions rely directly or indirectly on having diversity in the swarm of robots. For example, specialization and division of labor need to have different individuals doing different things at the same time, and these are only the first steps before tackling more interesting problems.

The main message of this paper is that well-established diversity methods from ER are difficult to implement directly in EER algorithms. These algorithms may share some similarities with their older siblings in ER; they have, however, many notable differences. We discussed some of them: the interactions between agents that influence the course of evolution in unpredictable ways, their distributed nature, and the fact that measuring behavior may be erroneous.

In this paper, we present a selection scheme that is able to increase behavioral diversity in EER algorithms. In order to reliably measure behaviors, it estimates behavioral distances on board the agents. It compares the outputs of candidate controllers on the same reference input. We also describe, a metric on the space of behaviors, that allows us to compare behaviors. We present a few experiments we ran to put the selection scheme to the test, and the results were conclusive. The behaviors that evolved in the random selection case tend to be similar, achieving just what is required to survive, while, on the other hand, diversity selection evolves rich and complex behaviors. We discuss and give some insights to explain the resulting behaviors and put forward the idea, which remains to be tested, of exploiting affordences.

## References

[1] Nicolas Bredeche. Embodied evolutionary robotics with large number of robots. In *the 14th international conference on the synthesis and simulation of living systems (ALIFE'14)*, pages 1–2, New York, USA, 2014.

[2] Nicolas Bredeche, Evert Haasdijk, and Abraham Prieto. Embodied evolution in collective robotics: A review. *Front, in Robo. and AI*, 5:12, 2018.

[3] Nicolas Bredeche and Jean-Marc Montanier. Environment-driven Embodied Evolution in a Population of Autonomous Agents. In *PPSN 2010*, pages 290–299, Krakow, 2010.

[4] Nicolas Bredeche, Jean-Marc Montanier, Wenguo Liu, and Alan Winfield. Environment-driven Distributed Evolutionary Adaptation in a Population of Autonomous Robotic Agents. *Mathematical and Computer Modelling of Dynamical Systems*, 18(1):101–129, 2012.

[5] Nicolas Bredeche, Jean-Marc Montanier, Berend Weel, and Evert Haasdijk. Roborobo! a fast robot simulator for swarm and collective robotics. *CoRR*, abs/1304.2888, 2013.

[6] D.M. Endres and J.E. Schindelin. A new metric for probability distributions. *IEEE Transactions on Information Theory*, 49(7):1858–1860, 2003.

[7] Sevan Ficici, Richard Watson, and Jordan Pollack. Embodied evolution: A response to challenges in evolutionary robotics. In *Proc. of the* $8^{th}$ *Europ. Wrks. on Learning Robots*, 1999.

[8] J. James Gibson. *The Ecological Approach to Visual Perception*. Psychology Press, 2015.

[9] Jorge Gomes, Paulo Urbano, and Anders Lyhne Christensen. Introducing novelty search in evolutionary swarm robotics. In *Swarm Intelligence*, pages 85–96, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[10] Faustino J. Gomez. Sustaining diversity using behavioral information distance. In *Proceedings of the 11th Annual Conference on Genetic and Evolutionary Computation*, page 113–120, New York, NY, USA, 2009. Association for Computing Machinery.

[11] Evert Haasdijk, Nicolas Bredeche, and A. E. Eiben. Combining Environment-Driven Adaptation and Task-Driven Optimisation in Evolutionary Robotics. *PLoS ONE*, 9(6):e98466, June 2014.

[12] Emma Hart, Andreas S. W. Steyven, and Ben Paechter. Evolution of a functionally diverse swarm via a novel decentralised quality-diversity algorithm. In *Proc. of the Genetic and Evolutionary Computation Conference, Kyoto, Japan, July 15-19*, pages 101–108. ACM, 2018.

[13] Joel Lehman and Kenneth O. Stanley. Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary Computation*, 19(2):189–223, June 2011.

[14] Jean-Marc Montanier, Simon Carrignon, and Nicolas Bredeche. Behavioral specialization in embodied evolutionary robotics: Why so difficult? *Frontiers in Robotics and AI*, 3:38, 2016.

[15] Jean-Baptiste Mouret and Jeff Clune. Illuminating search spaces by mapping elites. *CoRR*, abs/1504.04909, 2015.

[16] Abraham Prieto, Josè Antonio Becerra, Francisco Bellas, and Richard J. Duro. Open-ended evolution as a means to self-organize heterogeneous multi-robot systems in real time. *Rob. and Auto. Sys.*, 58(12):1282 – 1291, 2010.

[17] Justin K. Pugh, Lisa B. Soros, and Kenneth O. Stanley. Quality diversity: A new frontier for evolutionary computation. *Frontiers in Robotics and AI*, 3:40, 2016.