Evolving Robot Swarm Behaviors by Minimizing Surprise: Results of Simulations in 2-d on a Torus

Richard Borkowski
Heiko Hamann
ribo@mail.upb.de,hamann@iti.uni-luebeck.de
Heinz Nixdorf Institute, Department of Computer Science, University of Paderborn, Germany
Institute of Computer Engineering, University of Lübeck, Germany

CCS CONCEPTS

• Computing methodologies → Multi-agent systems; • Computer systems organization → Evolutionary robotics;

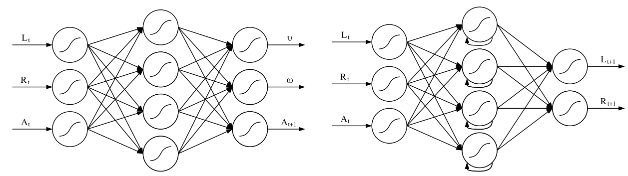
ACM Reference format:
Richard Borkowski and Heiko Hamann. 2017. Evolving Robot Swarm Behaviors by Minimizing Surprise: Results of Simulations in 2-d on a Torus. In *Proceedings of GECCO '17 Companion, Berlin, Germany, July 15-19, 2017,* 2 pages.
DOI: http://dx.doi.org/10.1145/3067695.3082548

1 INTRODUCTION

The application of evolutionary robotics [1] to swarm robotics gives evolutionary swarm robotics [8]. The evolution or learning of multi-agent behaviors is known to be challenging [7]. Hence, new approaches still need to be explored. Examples are innovative methods to explore environment-driven, distributed evolution [2, 4]. Here, we are inspired to evolve collective behaviors following a mathematical framework by Friston et al. [3], which defines an information-theoretic analogon to thermodynamic (Helmholtz) free energy. This free energy is basically an error in the predictions that our brain makes about our environment. Evolution is related by the rationale that minimal prediction errors are achieved by limiting an agent's reactions to sensory input. This results, in turn, in better adapted behaviors: "By sampling [...] the environment selectively they restrict their exchange with it within bounds that preserve their physical integrity and allow them to last longer" [3]. The previously investigated evolution of swarm behaviors by minimizing surprisal [5, 6, 9] is subject to this study. Previous studies were limited to artificial 1-d environments, here, we report first results for 2-d. Although adding one dimension may seem a minor step, there are qualitative changes in the emergent behaviors (e.g., flocking is a collective decision with infinitely many options) and the future transition to real robots will be easier starting from 2-d.

2 IMPLEMENTATION

Our approach to let robots predict their sensory input is realized with 2 artificial neural networks (ANN) for each robot (Fig. 1) with



(a) action network (b) prediction network

Figure 1: Controller for the first sensor model

(a) 2 discrete touch sensors (b) 4 discrete touch sensors

Figure 2: Sensor models 1 and 2

sensor model 1 (Fig. 2a). Sensor model 2 (Fig. 2b) has corresponding ANN. The action network implements the actual controller whereas the prediction network tries to forecast the sensory input of the next time step. Both ANN have the same input: all sensor readings and their last *action value*. The action value implements an action selection that switches between straight motion with speed v or rotation with angular speed w on the spot and increases the probability of observing stationary behaviors. Sensors are discrete touch sensors (1 for contact, 0 for no robot). The robot swarms are homogeneous, that is, we have 2 populations: a population of genomes encoding ANN and a homogeneous population of simulated robots. In an evaluation, all robots share identical genomes. The fitness function

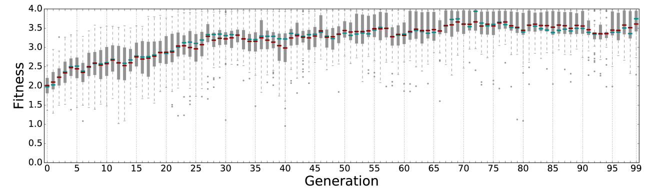
$$f_g = \frac{1}{NT} \sum_{t}^{T} \sum_{n}^{N} \sum_{i}^{S} 1 - |p_{n,i}(t) - s_{n,i}(t)|, \tag{1}$$

encourages to minimize surprise. The fitness f_g of a genome $g \in G$ with $|G| = 50$ is calculated for S sensors, $s_{n,i}(t)$ is sensor value of sensor i of robot n at time step t, $p_{n,i}(t)$ is the prediction for sensor i, $N \in \{25, 50, 75, 100\}$ is swarm size, and $T = 2000$ is the number of time steps. The theoretical best fitness is $f_g^{\max} = S$. We evolve for 100 generations. The environment is implemented as a 2-d torus to emulate an infinite space. This way we avoid wall effects that may complicate behaviors such as flocking. The simulation software is available online[1].

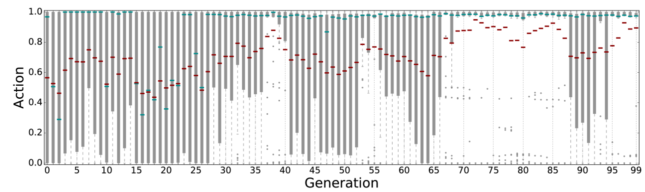
[1] https://github.com/ribork/thesis-code

3 RESULTS

Fig. 3a shows the population's fitness over generations of a selected experiment. The fitness increases about until generation 50 and then saturates. The fitness only rewards correct predictions, hence we need an additional analysis to classify behaviors. We use the action value to detect interesting behaviors in the populations (amount of forward motion vs rotation, Fig. 3b). Genomes producing behaviors with a median action value excluding 0 and 1 are indicating reactive behaviors. Within these genomes we identify basic collective behaviors, such as aggregation (in 2 variations), flocking, and circling dispersion. Aggregation occurs either with stopped robots[2] or with circling robots[3], where robots form rotating clusters (Fig. 4a). Flocking emerged for lower swarm densities where less interactions occur (Fig. 4c). The flocking behaviors show good scalability with an increasing transient time for higher swarm densities[4]. Using sensor model 1, dispersion behaviors are based on motion (Fig. 4b). Robots move in circles and adjust their distances by short stops based on detecting robots[5]. In an early test of whether we can influence the emergence of certain behaviors, we tried to bias towards dispersion with less motion. We introduce that bias by changing the sensor model. We create sensor model 2 by adding 2 sensors at the robot's back (Fig. 2b, the ANN are adapted accordingly). The approach is successful, see trajectories in Fig. 4d and the video[6] (including tests for scalability and robustness by post-evaluations with different swarm sizes and obstacles). In preliminary experiments we implemented continuous distance sensors. Then also the prediction network and the fitness function (eq. 1) operate on continuous values. The hypothesis was that the evolutionary process may profit from the additional information provided by continuous differences. However, continuous sensors combined with the action-value approach resulted in irregular behaviors.

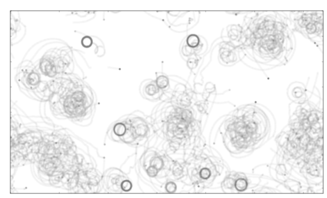


(a) fitness over generations

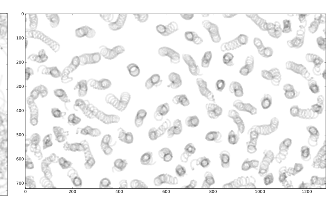


(b) actions over generations

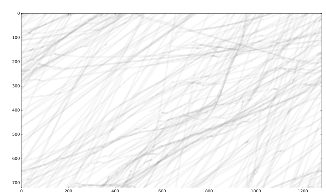
Figure 3: Results, fitness and action values (blue markers indicate median, red markers indicate mean)



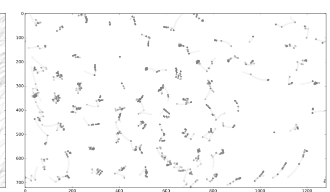
(a) aggregation, sensor model 1



(b) dispersion, sensor model 1



(c) flocking, sensor model 1



(d) dispersion, sensor model 2

Figure 4: Evolved behaviors for sensor models 1 and 2; trajectories of all robots in a 2-d arena

4 CONCLUSION

We have shown that the 'evolution of swarm behaviors by minimizing surprise' approach is successfully applied to 2-d environments and that the evolutionary dynamics can be biased by the sensor models of the robots. In future work, we plan to investigate how to balance exploration during the evolutionary process with bias towards desired behaviors. We also want to make the step to experiments with real robots, such as the Kilobots.

REFERENCES

[1] Josh C. Bongard. 2013. Evolutionary robotics. *Commun. ACM* 56, 8 (2013), 74–83. http://doi.acm.org/10.1145/2493883

[2] Nicolas Bredeche, Jean-Marc Montanier, Wenguo Liu, and Alan F.T. Winfield. 2012. Environment-driven distributed evolutionary adaptation in a population of autonomous robotic agents. *Mathematical and Computer Modelling of Dynamical Systems* 18, 1 (2012), 101–129.

[3] Karl Friston, James Kilner, and Lee Harrison. 2006. A free energy principle for the brain. *Journal of Physiology Paris* 100, 1 (2006), 70–87.

[4] Evert Haasdijk, Berend Weel, and AE Eiben. 2013. Right on the MONEE: combining task-and environment-driven evolution. In *Proceeding of the 15th conference on Genetic and evolutionary computation conference (GECCO 2013)*. ACM, 207–214.

[5] Heiko Hamann. 2014. Evolution of Collective Behaviors by Minimizing Surprise. In *14th Int. Conf. on the Synthesis and Simulation of Living Systems (ALIFE 2014)*, Hiroki Sayama, John Rieffel, Sebastian Risi, René Doursat, and Hod Lipson (Eds.). MIT Press, 344–351. http://dx.doi.org/10.7551/978-0-262-32621-6-ch055

[6] Heiko Hamann. 2014. Evolving Prediction Machines: Collective Behaviors Based on Minimal Surprisal. In *Int. Conf. on Genetic and Evolutionary Computation (GECCO 2014)*. ACM, 31–32. [extended abstract].

[7] Liviu Panait and Sean Luke. 2005. Cooperative Multi-Agent Learning: The State of the Art. *Autonomous Agents and Multi-Agent Systems* 11, 3 (2005), 387–434.

[8] Vito Trianni. 2008. *Evolutionary Swarm Robotics - Evolving Self-Organising Behaviours in Groups of Autonomous Robots*. Studies in Computational Intelligence, Vol. 108. Springer, Berlin, Germany.

[9] Payam Zahadat, Heiko Hamann, and Thomas Schmickl. 2015. Evolving Diverse Collective Behaviors Independent of Swarm Density. In *Workshop Evolving Collective Behaviors in Robotics (GECCO 2015)*. ACM, 1245–1246. [extended abstract].

[2] aggregation: https://youtu.be/OZjOesnZhaY
[3] circling: https://youtu.be/4rAEG7MEORI
[4] flocking: https://youtu.be/OYuiiOgbWkw
[5] moving dispersion: https://youtu.be/uuDwvmM_Z3E
[6] converging dispersion: https://youtu.be/vC3hbAubaMI