Minimize Surprise MAP-Elites: A Task-Independent MAP-Elites Variant for Swarms

Tanja Katharina Kaiser and Heiko Hamann
University of Lübeck, Germany
kaiser@iti.uni-luebeck.de, hamann@iti.uni-luebeck.de

April 14, 2022

Abstract

Swarm robotics controllers are often automatically generated using methods of evolutionary computation with a task-specific fitness function to guide the optimization process. By contrast, our minimize surprise approach uses a task-independent fitness function to generate diverse behaviors over several independent evolutionary runs. Alternatives are divergent search algorithms rewarding behavioral novelty, such as novelty search, and quality-diversity algorithms generating diverse high-quality solutions, such as MAP-Elites. These approaches usually rely on task-dependent measures. We propose Minimize Surprise MAP-Elites, a task-independent MAP-Elites variant that combines MAP-Elites with our minimize surprise approach. Our first experiments result in high-quality solutions that lead to behavioral diversity across tasks and within tasks.

1 Introduction

Methods of evolutionary computation are frequently used to generate swarm robotics controllers. In standard approaches, a targeted behavior is quantified by a task-specific fitness function that is optimized over generations. By contrast, divergent algorithms push for behavioral diversity (e.g., novelty search [4]), and quality diversity algorithms lead to diverse high-quality solutions (e.g., MAP-Elites [5]). These approaches, however, usually still use task-specific measures and find diverse solutions for fulfilling the same task. For example, self-assembly behaviors were evolved with MAP-Elites rewarding how well the swarm assembled into one particular structure [1]. Multi-task MAP-Elites [6] solves many tasks of the same family simultaneously using task-dependent fitness functions and a distance measure between those tasks. Gravina et al. [2] use a task-dependent performance measure to fill the behavior-performance map, but task-independent measures, such as novelty or surprise,

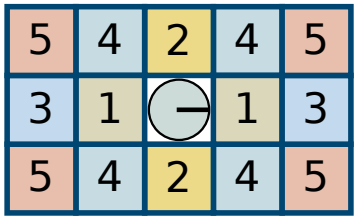


Figure 1: Sensor model. Each colored cell gives one sensor, the circle represents the agent, and the line gives its heading. Grid cells in the same color and with the same index i form dimension X_i in Minimize Surprise MAP-Elites.

for selection in MAP-Elites. By contrast, our standard minimize surprise approach [3] combines simple evolutionary algorithms with a task-independent reward for prediction accuracy that leads to a variety of high-quality behaviors for different tasks over several independent evolutionary runs. These tasks are defined posteriorly to measure the emergent behavioral diversity, since the approach itself is task-independent. We propose Minimize Surprise MAP-Elites, a task-independent MAP-Elites variant, that uses our task-independent minimize surprise reward as the performance measure in MAP-Elites. This approach allows for behavioral diversity within tasks and across tasks. We exemplify our approach using our self-assembly scenario, which permits comparison of Minimize Surprise MAP-Elites with our previous results using standard minimize surprise and novelty search with a task-independent behavioral characteristic [3]. A comparison of Minimize Surprise MAP-Elites and MAP-Elites with a task-dependent performance measure is future work.

2 Setup

We introduce the self-assembly scenario, the three compared evolutionary approaches (i.e., standard minimize surprise [3], Minimize Surprise MAP-Elites, novelty search with a task-independent behavioral characteristic [3]), and our evaluation metrics.

2.1 Self-Assembly Scenario

We use our self-assembly scenario from previous work [3] to allow easy comparison with previous results. In this scenario, a swarm of $N = 100$ homogeneous agents lives on a square 2D torus grid that allows for simplified positioning and sensing. Swarm density is varied by keeping the swarm size fixed and changing grid side length $L \in \{15, 20\}$. Each agent has $R = 14$ binary sensors covering its surrounding grid cells (see Fig. 1) and can execute three actions: turning by $\pm 90°$ or moving one grid cell forward. A move forward is blocked if the targeted grid cell is already occupied. In each time step, a feedforward artificial neural network (ANN) determines the agent's next action based on its current

sensor values and its last action (actor). In standard minimize surprise and Minimize Surprise MAP-Elites, agents are additionally equipped with a recurrent ANN predicting the agent's $R = 14$ sensor values of the next time step based on its current sensor values and its next action (predictor). In all three evolutionary approaches, individual solutions encode the synaptic weights of the actor and, if used, the predictor.

2.2 Evolutionary Approaches

We introduce the three compared evolutionary approaches: standard minimize surprise [3], Minimize Surprise MAP-Elites, and novelty search with a task-independent behavioral characteristic [3].

Standard Minimize Surprise Approach We first introduce our standard minimize surprise approach that we use to build our proposed Minimize Surprise MAP-Elites approach. Standard minimize surprise evolves actor-predictor ANN pairs with a simple evolutionary algorithm and a task-independent fitness function rewarding prediction accuracy. Fitness F is defined as

$$F = \frac{1}{TNR} \sum_{t=0}^{T-1} \sum_{n=0}^{N-1} \sum_{r=0}^{R-1} (1 - |p_r^n(t) - s_r^n(t)|) \ , \tag{1}$$

with evaluation length T (in time steps), swarm size N, number of sensors R, and prediction $p_r^n(t)$ for and actual value $s_r^n(t)$ of agent n's sensor r at time step t. Consequently, fitness only directly rewards the predictor ANN while the actor ANN receives indirect selective pressure by being paired with the predictor. In our self-assembly scenario [3], we run evolution for 100 generations using a population size of 50, an evaluation length of 500 time steps, 10 independent evaluations per solution, elitism of one, no crossover, proportionate selection, and a mutation rate of 0.1. An individual solution's overall fitness is the minimum of its ten independent evaluations. In Sec. 3.2, we compare Minimize Surprise MAP-Elites with 50 independent standard minimize surprise runs [3] per grid size $L \in \{15, 20\}$.

Minimize Surprise MAP-Elites Our proposed Minimize Surprise MAP-Elites algorithm combines the standard MAP-Elites algorithm [5] with standard minimize surprise [3]. Standard MAP-Elites generates diverse high-quality solutions by filling a discrete behavior-performance map over generations with solutions, retaining the highest performing solution per cell. Minimize Surprise MAP-Elites evolves actor-predictor ANN pairs using standard MAP-Elites with the minimize surprise fitness function (Eq. 1) as performance measure, and sensor values as the dimensions of the behavior-performance map. The approach is fully task-independent, potentially allowing for the emergence of solutions for different tasks and diverse solutions for the same task.

For the self-assembly scenario, we define a behavior-performance map with five dimensions based on the swarm's discretized mean sensor values in the last

time step of an evaluation. Mean sensor values greater than 0.5 are mapped to 1, otherwise to 0. For each dimension in the behavior-performance map, we aggregate two or four sensors (cf. Fig. 1) resulting in $3^3 \cdot 5^2 = 675$ map cells. The aggregation of sensor values is based on symmetries found in the emergent self-assembly behaviors in minimize surprise (cf. Sec. 2.3) and keeps the solution set concise. We initialize Minimize Surprise MAP-Elites by evaluating 2,500 solutions and placing them into their respective map cells. Per solution, we do three independent evaluations of 500 time steps each, assigning the minimum fitness reached in these three evaluations as the overall fitness. After initialization, a random cell of the map is chosen, offspring created by mutation with a mutation rate of 0.1, and evaluated. In total, we do 250,000 evaluations, equaling the number of evaluations in 50 independent evolutionary runs with our standard minimize surprise approach. We do 10 independent MAP-Elites runs per grid size $L \in \{15, 20\}$.

Novelty Search We compare Minimize Surprise MAP-Elites with our novelty search approach with a task-independent behavioral characteristic [3]. In this approach, we implement novelty search as described by Lehman and Stanley [4], and define a potentially task-independent R-dimensional behavioral characteristic vector

$$b = \frac{1}{N} \left[\sum_{n=0}^{N-1} s_0^n(T), \ldots, \sum_{n=0}^{N-1} s_{R-1}^n(T) \right] , \tag{2}$$

that is, the mean sensor values in the last time step T of an evaluation. Novelty, which is rewarded here, is the Euclidean distance between the behavioral characteristic of a solution and up to its tenth nearest neighbors in the current population and an archive of past solutions. We base our novelty search implementation on the same evolutionary algorithm as in standard minimize surprise, but do not apply elitism and reward novelty instead of prediction accuracy.

2.3 Metrics

We use three metrics defined by Mouret and Clune [5] to quantify the quality of Minimize Surprise MAP-Elites: coverage C_{map}, precision $\bar{F}$, and global performance $F_{\max}$. Coverage C_{map} is the percentage of behavior-performance map cells filled with solutions. Precision $\bar{F}$ is the mean performance (i.e., prediction accuracy; Eq. 1) of the filled map cells, and global performance $F_{\max}$ is the maximum performance.

In addition, we measure behavioral diversity, solution quality Q, and pattern coverage C_{pat}. Behavioral diversity is measured by classifying formed agent structures using our previously defined nine patterns [3]: aggregation, clustering, loose grouping, pairs, lines, triangular lattices, swirls, squares, and random dispersion. Solution quality Q gives the percentage of agents assembled into the pattern. Pattern coverage C_{pat} is the amount of different emergent patterns normalized by the nine defined patterns, that is, it gives the percentage of found patterns.

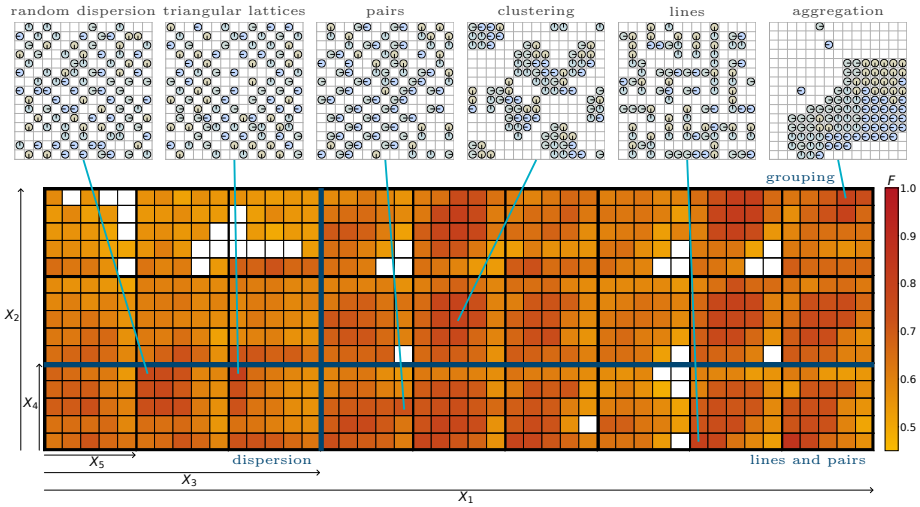
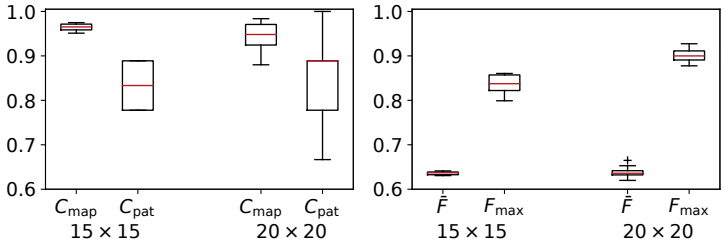


Figure 2: Behavior-performance map of one representative Minimize Surprise MAP-Elites run on the 15×15 grid and examples for several emergent structures. Dimensions $X_1 - X_5$ represent two to four sensors each (cf. Fig. 1) with $X_1 - X_3 \in [0 \,..\, 2]$ and $X_4, X_5 \in [0 \,..\, 4]$. Cell colors give the performance $F \in [0, 1]$, white map cells are empty (i.e., no solution was found). Dark blue lines mark areas where certain pattern types perform well.



(a) coverage and pattern coverage (b) precision and global performance

Figure 3: Coverage C_{map}, pattern coverage C_{pat}, precision $\bar{F}$, and global performance $F_{\max}$ (prediction accuracy, Eq. 1) for ten independent Minimize Surprise MAP-Elites runs per grid side length $L \in \{15, 20\}$.

3 Results

We analyze the performance of our Minimize Surprise MAP-Elites approach individually first and compare it to standard minimize surprise and novelty search afterwards.

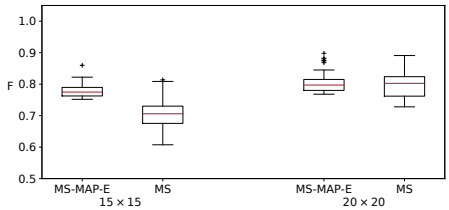
3.1 Minimize Surprise MAP-Elites

We find a median coverage C_{map} of 0.97 on the 15×15 grid and of 0.95 on the 20×20 grid in our Minimize Surprise MAP-Elites runs, see Fig. 3(a). Thus, nearly all cells of the behavior-performance map are filled with solutions. Pattern coverage C_{pat} is also high with a median of 0.83 on the 15×15 grid and of 0.88 on the 20×20 grid. That means, we find a median of 7.5 patterns on the smaller grid and of eight patterns on the larger grid. The difference in pattern coverage C_{pat} is partly caused by swarm density. Squares cannot form on the smaller grid due to the high swarm density of 0.44; the pattern can only perfectly form in swarm densities up to 0.25. However, squares and swirls generally rarely emerge and are probably hard to form. On the larger grid, swarm density is almost too low for the formation of repetitive triangular lattices and thus the pattern does not always emerge. Solutions are high performing with a median precision $\bar{F}$ of 0.64 and a median global performance F_{max} of at least 0.84 on both grid sizes, see Fig. 3(b). In total, we find high-quality behaviors for different tasks.

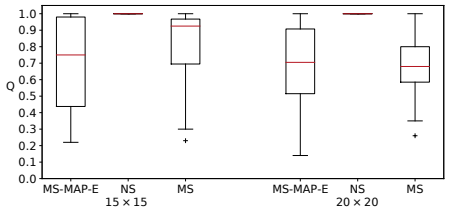
Fig. 2 visualizes the behavior-performance map of one Minimize Surprise MAP-Elites run on the 15×15 grid as representative example. Areas of high performance (i.e., red areas) each match the criteria of one of the nine defined patterns closely. High performance (i.e., prediction accuracy) can still be reached if the structure formation deviates slightly from the defined classification criteria (i.e., $Q < 1.0$). Consequently, several map cells can contain solutions for the formation of the same pattern. We always find more than one solution for a pattern type, if any. Map cells with worse performance usually contain grouping or random dispersion behaviors because these patterns are not based on the exact positioning of the agents unlike, for example, lines and pairs. Overall, this shows that our approach finds high-quality solutions for the self-assembly of different patterns, and diverse solutions to the assembly of each pattern.

3.2 Comparison

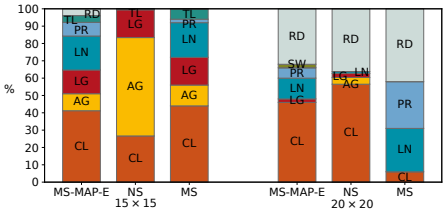
To show the potential of our approach, we compare the results of Minimize Surprise MAP-Elites with standard minimize surprise and novelty search with a task-independent behavioral characteristic. For a fair comparison, we select 50 solutions out of the same number of evaluations per approach. One Minimize Surprise MAP-Elites run has 250,000 evaluations, which is as many as 50 independent minimize surprise and novelty search runs. Thus, we pick the 50 best solutions based on performance (Eq. 1) of one Minimize Surprise MAP-Elites



(a) performance (prediction accuracy, Eq. 1)



(b) solution quality (percentage of correctly assembled agents)



(c) behavior distributions

Figure 4: Performance F (Eq. 1), solution quality Q, and behavior distributions of the 50 best solutions of Minimize Surprise MAP-Elites (`MS-MAP-E`; 1 independent run), minimize surprise (`MS`; 50 independent runs), and novelty search (`NS`; 50 independent runs) per grid size $L \times L, L \in \{15, 20\}$ with clustering (`CL`), aggregation (`AG`), loose grouping (`LG`), swirls (`SW`), lines (`LN`), pairs (`PR`), triangular lattices (`TL`), and random dispersion (`RD`).

run as representative example. From minimize surprise, we take the best evolved individuals of 50 independent evolutionary runs [3]. In contrast, prediction accuracy (Eq. 1) is not and cannot be measured in novelty search. We select the best solutions based on solution quality Q in that case, since high-quality solutions generally would allow for high prediction accuracy.

First, we compare the performance (Eq. 1) of the 50 best solutions of Minimize Surprise MAP-Elites and minimize surprise, see Fig. 4(a). We find statistically significantly better performance for Minimize Surprise MAP-Elites on the 15×15 grid, but no significant differences on the 20×20 grid (Mann-Whitney U test, $p < 0.001$). Next, we compare solution quality Q of the three approaches, see Fig. 4(b). Novelty search reaches statistically significantly better solution quality than Minimize Surprise MAP-Elites and minimize surprise on both grid sizes (Kruskal–Wallis test, $p < 0.001$; Mann-Whitney U test with Bonferroni correction, $p < 0.001$). However, novelty search has a median solution quality of 1.0 as we select best solutions based on solution quality in that case and thus the comparison is biased. Selecting solutions based on solution quality in Minimize Surprise MAP-Elites would lead to a competitive median solution quality of 0.97. For Minimize Surprise MAP-Elites and minimize surprise, we do not find significant differences in solution quality. Last, we compare behavioral diversity and find statistically significant differences on both grid sizes (Fisher's Exact test, $p < 0.001$), see Fig. 4(c). On the 15×15 grid, novelty search has a significantly different behavior distribution than Minimize Surprise MAP-Elites and minimize surprise. On the 20×20 grid, minimize surprise has a significantly different behavior distribution than the other two approaches. This is intuitive, as we find the least number of different patterns (i.e., four) for the significantly different behavior distributions, but the maximum number of different patterns (i.e., seven) on both grid sizes for Minimize Surprise MAP-Elites.

Overall, we find that Minimize Surprise MAP-Elites leads to high-quality solutions and the greatest number of different patterns on both grid sizes. It also requires less post-evaluation effort than novelty search as it outputs 675 high-quality solutions instead of 250,000 potential solutions of varying quality.

4 Conclusion

In our work, we proposed a new MAP-Elites variant that combines standard MAP-Elites with the task-independent minimize surprise approach to generate behavioral diversity across tasks and within tasks. We exemplified our approach in a swarm robotics setting using a self-assembly scenario, and compared our results on Minimize Surprise MAP-Elites with our previous results for minimize surprise and novelty search [3]. Our results show that combining minimize surprise with MAP-Elites leads to diverse, high-quality solutions. But our study is based on a simple experimental setup. Experiments in more realistic simulations or even with real robots may come with several challenges, especially related to the discretization of the behavior-performance map. Nevertheless, our first results on Minimize Surprise MAP-Elites are promising. To explore the potential

of our approach, we will investigate the combination of minimize surprise with other MAP-Elites variants [5, 2], and more complex swarm scenarios in future work.

References

[1] Leo Cazenille, Nicolas Bredeche, and Nathanael Aubert-Kato. Exploring self-assembling behaviors in a swarm of bio-micro-robots using surrogate-assisted map-elites. In *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 238–246, 2019.

[2] Daniele Gravina, Antonios Liapis, and Georgios N. Yannakakis. Blending notions of diversity for map-elites. In *Proc. of the Genetic and Evolutionary Comput. Conf. Companion*, GECCO '19, page 117–118, New York, NY, USA, 2019. Assoc. for Comput. Machinery.

[3] Tanja Katharina Kaiser and Heiko Hamann. Evolution of diverse swarm behaviors with minimal surprise. *Artificial Life Conf. Proc.*, (32):384–392, 2020.

[4] Joel Lehman and Kenneth O. Stanley. Exploiting open-endedness to solve problems through the search for novelty. In S. Bullock, J. Noble, R. Watson, and M. A. Bedau, editors, *Artificial Life XI: Proc. of the Eleventh Int. Conf. on the Simulation and Synthesis of Living Systems*, pages 329–336. MIT Press, 2008.

[5] Jean-Baptiste Mouret and Jeff Clune. Illuminating search spaces by mapping elites, 2015. arXiv.

[6] Jean-Baptiste Mouret and Glenn Maguire. Quality diversity for multi-task optimization. In *Proc. of the 2020 Genetic and Evolutionary Comput. Conf.*, GECCO '20, page 121–129, New York, NY, USA, 2020. Assoc. for Comput. Machinery.