



Effects of excessive elitism on the evolution of artificial creatures with NEAT

Siti Aisyah Binti Jaafar¹ · Reiji Suzuki¹ · Satoru Komori¹ · Takaya Arita¹

Received: 5 July 2023 / Accepted: 6 March 2024 / Published online: 3 April 2024
© International Society of Artificial Life and Robotics (ISAROB) 2024

Abstract

This paper proposes a simple method based on a novel use of elitism to increase the population size of artificial creatures while minimizing evaluation cost. This can contribute to preventing premature convergence of the population. We propose the “Excessive Elitism (EE)” method by modifying elitism in HyperNEAT (Hypercube-based NeuroEvolution of Augmenting Topologies), which is an evolutionary algorithm commonly used to evolve genotype [i.e., Compositional Pattern Producing Network (CPPN)] of artificial creatures. In EE, the evaluated fitness of best-fit individuals will be succeeded and reused instead of being re-evaluated during subsequent fitness evaluations, thereby reducing the evaluation cost if the elite size is excessive. Notably, EE also disables speciation and fitness sharing, serving to simplify the population structure and reduce complexity. In a 3D multi-agent environment, we evolved the morphology and behavior of artificial creatures with a simple target approach task. We assumed a baseline case (EE (2, 20)) in which a small population size was used due to the strong limitation of the evaluation cost and adopted a normal small elite size. This often led to premature convergence of the population to suboptimal individuals who could not reach the target. However, with the application of EE, the population was capable of evolving to reach the target, maintaining an evaluation cost comparable to EE (2, 20). We demonstrate that EE method serves as a simpler alternative to speciation for diversity preservation, capable of enhancing both the average and optimal fitness of a population, thus preventing premature convergence at a minimal evaluation cost. Further research in complex environments is required to fully uncover the potential and limitations of this method.

Keywords Artificial creatures · Artificial life · Elitism · Evaluation cost · 3D multi-agent simulation

1 Introduction

An approach to evolving artificial creatures has gained significant interest in scientific (e.g., evo-devo [1, 2]) and engineering research (e.g., design of soft robots [1–3]). In particular, this approach has proven useful for understanding the interactions between ecological, developmental, and evolutionary processes known as eco-evo-devo. Furthermore, it has provided clarity on the principles that govern

the relationship between environmental complexity and the evolution of morphological and behavioral traits. We aim to discuss the evolution of development and niche construction within a 3D multi-agent environment. To achieve this, we will utilize the framework by Chiba et al. [3] which enables the evolution of creatures' behaviors toward environmental modification.

However, incorporating various factors related to the above topics into the model has significantly increased the evaluation cost for the physical simulation. As a result, evolutionary experimentation grown in complexity and computational expense. This is largely due to the substantial cost associated with fitness evaluation, which relies on a physical simulation of complex creatures' behaviors, complexity that emerges from the interactions among the creatures. This has been a significant challenge in the evolution of artificial creatures, thoroughly explored in several previous studies [4, 5]. Consequently, a small population size is often used as an

This work was presented in part at the joint symposium of the 28th International Symposium on Artificial Life and Robotics, the 8th International Symposium on BioComplexity, and the 6th International Symposium on Swarm Behavior and Bio-Inspired Robotics (Beppu, Oita and Online, January 25–27, 2023).

✉ Siti Aisyah Binti Jaafar
binti.jaafar.siti.aisyah.v4@s.mail.nagoya-u.ac.jp

¹ Graduate School of Informatics, Nagoya University, Nagoya, Japan

ad hoc way to resolve the problem, although it may lead to the premature convergence of the population to local optima.

To overcome the problem, a number of studies have expressed confidence that the implementation of elitism can lead to performance improvements. For example, in the field of soft robot evolution, La Cava and Moore proposed search driver algorithms aimed at finding optimal solutions [6]. Elitism, a methodology that involves either preserving the individual with the best fitness for the next generation or replacing the least fit individual with a previous high-fitness individual, was implemented in some of these algorithms. The study's comparative analysis of the algorithms demonstrated that simulations introducing some form of elitism significantly contributed to the discovery of locomoting soft robot morphologies.

Methenitis et al. [7] also suggested that implementing the fitness-elitism technique using NeuroEvolution of Augmenting Topologies (NEAT) [8], which is a popular evolutionary algorithm for evolving complex artificial neural networks, improved the performance in evolving soft robots. The study began by experimenting with two exclusive methods. First, it involved a novelty-search method, which promotes diversity by rewarding novel individuals regardless of their fitness. Second, it implemented fitness elitism, where the best individuals are selected based on their fitness, thereby ensuring the most optimal genetic material is passed from one generation to the next. The study later revealed that by integrating both methods, the fit and unique individual was carried over to subsequent generations, potentially resulting in the production of offspring that are both novel and fit. Their findings emphasized that the combination of methodologies played a critical part in advancing the evolution of soft robots. Moreover, these findings suggest that the presence of elitism significantly enhances the evolution performance of artificial creatures. Beyond simply improving fitness through elitism, it is also crucial to consider the potential benefits it provides in reducing evaluation costs linked to complex components inherent in evolutionary simulations of creatures.

Elitism, as mentioned, has exhibited significant potential in enhancing the performance of artificial creatures' evolution. In coherence with this, an interesting aspect to consider is the role speciation plays in evolution when using NEAT. Speciation based on the affinity of phenotypes and fitness sharing among conspecific individuals serve as key mechanisms for NEAT to maintain the diversity of individuals within a small population. However, the implementation of speciation requires a metric to measure the phenotypic distance between individuals. Consequently, the population structure may become complicated, subsequently making the population dynamics more complex compared to the case with a single species. While the effective combination of methods and maintaining diversity in species presents

complexities, it makes the evolutionary process an intriguing field to explore further.

Given these complexities, this paper introduces a novel approach called “Excessive Elitism (EE)” derived from the evolutionary algorithm Hypercube-based NeuroEvolution of Augmenting Topologies (HyperNEAT) [9], frequently used to evolve the genotype (i.e., Compositional Pattern Producing Network (CPPN) [10]) of artificial creatures. The principal mechanism of EE is that the evaluated fitness of best-fit individuals will be succeeded and reused instead of being re-evaluated during subsequent fitness evaluation. This contributes to a reduced evaluation cost within a large population, particularly when applying an excessively large elite size. Consequently, the EE method presents a straightforward alternative to speciation and fitness sharing to prevent premature convergence [11, 12] of the population. This enables an increase in the population size while maintaining relatively low evaluation costs.

In this study, our focus is on evolution of the genotypes of a rigid-bodied creatures in a 3D-multi-agent environment, a situation typically resulted in the occurrence of premature convergence to local optima. We conducted experiments using varying population sizes whilst maintaining a consistent number of evaluated individuals. The outcome of our experiments revealed that employing EE in a large, single population proved more effective than in smaller populations, even when maintaining an equivalent number of evaluated individuals for each generation. Moreover, the findings pointed towards a potential optimal configuration for population size. Notably, EE showcased a morphological diversity comparable to standard NEAT-evolved populations that involved speciation. This suggests that EE is capable of evolving a population with an adaptability level similar to those evolved by NEAT, while concurrently preserving a less complex population structure.

2 Model

We use a framework for evolving artificial creatures where the morphology of individuals develops within a 3D-multi-agent environment. This is based on the Python module-based physics engine, PyBullet, employed to analyze eco-evo-devo in evolving artificial creatures (Fig. 1). The HyperNEAT (Hypercube-based NeuroEvolution of Augmenting Topologies) is adopted as the evolutionary algorithm to evolve the genotype (i.e., the Compositional Pattern Producing Network or CPPN) of artificial creatures. HyperNEAT is an evolutionary algorithm used for evolving complex neural networks, represented by the CPPN in the model [10]. We utilize this algorithm to evolve the CPPN as the genotype of each artificial creature, used in their morphological development and behavioral generation during the

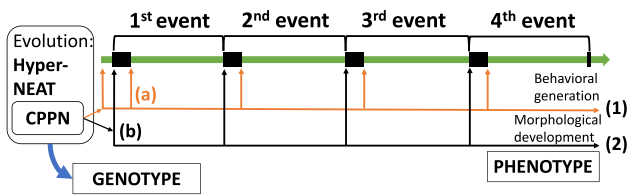


Fig. 1 The lifetime of a creature illustrated through a repeating process that involves two key steps: (1) the generation of behavior through connection weight settings (a), and (2) morphological development that takes place via new block additions (b), with four occurrence events

evolutionary process. We assume that a single large CPPN determined by the genotype of a creature represents both CPPNs as its sub-networks. Accordingly, the genotype of a creature generates two neural networks: one for determining

the morphological development and another for determining the behavioral generation process. These components are depicted in Fig. 3 and will be discussed in greater detail in the sections that follow.

2.1 Morphological development

Figure 1 illustrates the lifetime of a creature within a population. Each rigid-bodied creature is comprised of rectangular blocks interconnected via hinges (Fig. 2(ii)–(iii)). The initial body, composed of a single block, is positioned at the origin of the substrate space—the space where developmental events occur (Fig. 2(i)–(iii)). The creature’s morphology develops through the addition of new blocks to pre-existing body blocks, in several fixed time steps throughout its lifetime.

Fig. 2 The morphological development of an artificial creature, which proceeds from (i) the initial morphology, to (ii) the growth of a new block and results in (iii) the resultant morphology showcasing a fully developed new block, all occurring within a single developmental event in a substrate space

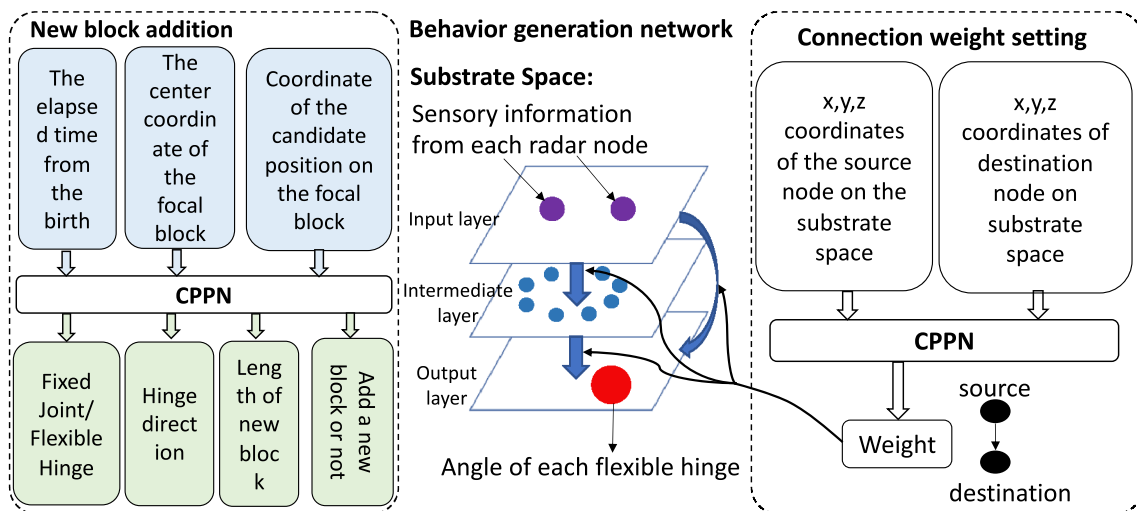
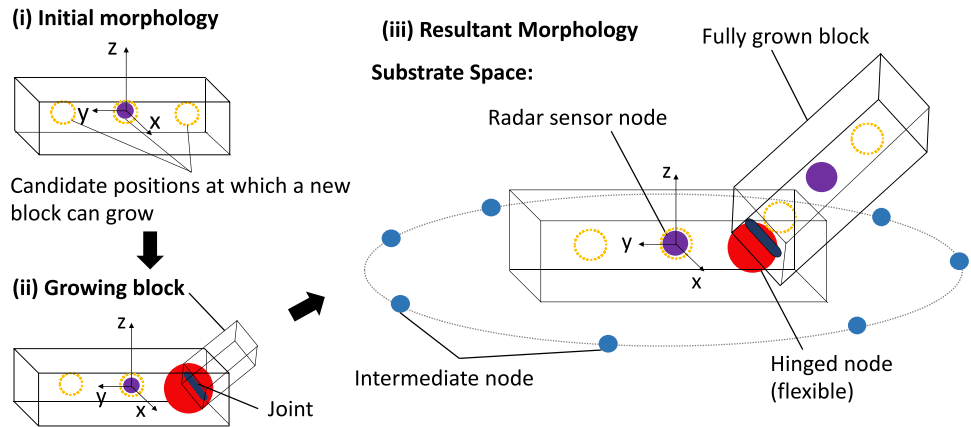


Fig. 3 The neural network architecture of the genotype of an artificial creature, as represented by the Compositional Pattern Producing Network (CPPN). (Left) the CPPN is utilized for morphological develop-

ment via the addition of a new block. (Right) the CPPN is employed to calculate connection weights for nodes in the three-layer behavioral generation network in substrate space

Each developmental event involves the addition of a new block to each candidate position (yellow circle) on the existing body blocks. Every block has three candidate positions. The genotype, CPPN (Fig. 3b left), receives three inputs: the elapsed time from the birth, the central coordinate of the focal block on the substrate space, and the coordinate of a candidate position of the focal block. This generates four outputs: determining if a new block is added or not, the length of the long sides of the new block, hinge direction, and deciding between a fixed hinge and flexible joint. The size of each new block incrementally increases from 0 to a specified length, up until the end of the developmental event. This process is spread over a total of 500-time steps.

An example illustrated in Fig. 2(i)–(iii) demonstrates how the process begins with a single block serving as the initial morphology (i). Subsequently, a new block grows at a right-candidate position (ii), creating candidate positions for additional blocks to grow during subsequent developmental events (iii).

2.2 Behavioral generation

After the completion of each morphological developmental event, the behavioral generation network (Fig. 3 (right)) is constructed by the genotype or updated in accordance with the creature’s morphology changes. This network structure comprises three layers: input, intermediate, and output. Each block has a centrally located radar sensor (Fig. 4) shown in purple, and a corresponding radar node that receive the sensory information from the substrate space (Fig. 3 (middle)). The radar sensor calculates the distance from the creature’s focal block to the target, utilizing the block’s angle and orientation relative to the target object in the evaluation field as depicted in Fig. 4. Consequently, the radar node receives the distance to the target in that specific direction and provides the

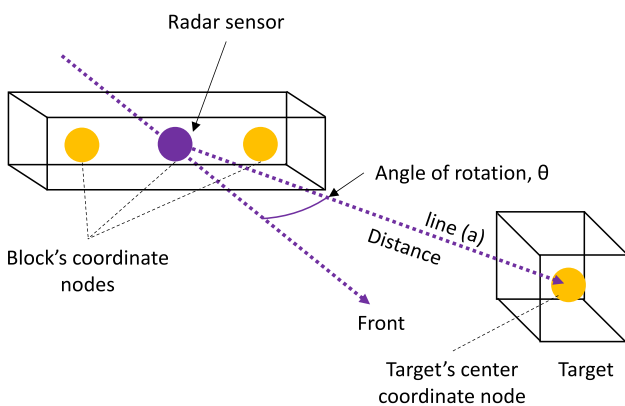


Fig. 4 Illustration of the radar sensor method. Here, we consider the longer side of the rectangular block as the ‘front side’. The angle of rotation derived from line (a) illustrates the current orientation of the block

input to the behavioral generation network, allowing the neural network to attempt to minimize this distance by manipulating the flexible hinges of the creature. Each hinge between blocks has a corresponding hinge node (red), and there are several intermediate nodes scattered around the creature on the substrate space (Fig. 2iii).

We assume that the radar, intermediate, and hinge nodes are located on the input, intermediate, and output layers of the behavioral generation network, respectively. For instance, in the network depicted in Fig. 3 (middle), there are 2 radar sensor nodes (input layer), 7 intermediate nodes (intermediate layer), and 1 hinge node (output layer). Connection exists among the nodes between the input-intermediate, intermediate-output, and input–output layers. Each linkage possesses its own connection weight, which is determined by the CPPN receiving the coordinates of the source and destination nodes pair on the substrate space. Each node within the output layer conveys the angle of the corresponding hinge, adjusting the actual angle of the physical hinge to match this output value. Consequently, the behavioral generation network can produce diverse behavioral patterns in the creature, according to the sensory input received from the radars.

2.3 Task

Figure 5 illustrates a field for fitness evaluation. We conducted experiments involving a locomotion task to examine the implications of excessive elitism on evolution. The artificial creatures were positioned in a circular manner, surrounding a central target (cube) placed on a flat surface. Fitness was defined as the distance traveled by creatures from the initial position towards the target.

Specifically, each creature’s fitness is calculated using the equation provided in Eq. (1)

$$\text{Fitness} = \begin{cases} d_i - d_{io}, & (d_{io} \leq d_i) \\ 0, & (d_{io} > d_i) \end{cases} \quad (1)$$

In this equation, d_i denotes the initial distance between the creature and the target, a value constant in all experiments,

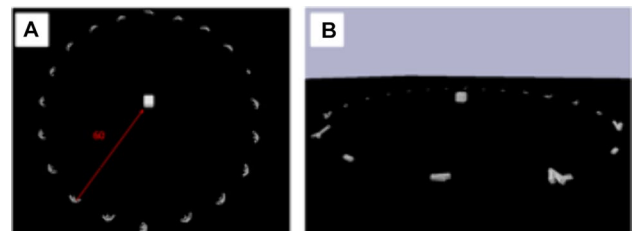


Fig. 5 Fitness evaluation field. **A** presents a top-down perspective, featuring the initial placement of creatures, who are arranged in a circular pattern, with an initial distance of 60 separating them from the target. **B** provides a lateral view of the evaluation field

which is 60. d_{io} is the current distance between the creature and the target object. From Eq. (1), we assume that if the distance between the creature and the target reaches or falls below 60, indicating the creature is either stationary or approaching the target, the fitness corresponds to the distance covered. However, if the creature moves away from its initial position, resulting in a negative total distance, the fitness is considered zero.

2.4 Evolution by HyperNEAT

In this section, we delve into how the evolution of the artificial creatures in this study was conducted using HyperNEAT (Hypercube-based NeuroEvolution of Augmenting Topologies) [9], an evolutionary algorithm frequently used to evolve the genotype of artificial creatures, namely the Compositional Pattern Producing Network (CPPN) [14]. HyperNEAT evolves the genome, a connective CPPN that generates a substrate by computing connection weights to solve a problem.

The selection method employed in HyperNEAT is derived from its underlying evolutionary algorithm, NEAT. According to NEAT, a solution to a problem evolves through the fitness function, which computes a singular number indicative of the quality of each individual genome (60); a higher score signifies enhanced ability. The algorithm proceeds through a pre-determined number of generations (400), each of which is generated through the reproduction and mutation of the fittest individuals from the preceding generation. The processes of reproduction and mutation may introduce additional nodes and/or connections to the genomes.

The variation operators in HyperNEAT play a crucial role in introducing genetic diversity into the population of genomes, which represents ANNs, during the evolution

process. The two primary variation operators employed in HyperNEAT are crossover and mutation.

As the algorithm progresses, genomes and the neural networks they generate may increase in complexity. The discovery of symmetries and regularities begins, though they do not always pertain fundamentally to the task. For instance, in earlier generations, the substrate generated by CPPN produces an output that results in poor creature behavior, inhibiting movements towards the target. However, as the evolutionary process continues, it uncovers the essential regularity necessary for genomes to achieve optimum fitness for the task.

2.5 Excessive elitism method

In the evolutionary process in NEAT, various mechanisms have been incorporated to strike a balance between fitness and population diversity. These mechanisms include gene tracking, speciation, and incremental complexification of the network. NEAT defines two levels of elitism: species elitism and reproduction elitism. Our focus is on modifying the latter, which proposes that the most-fit individuals in each species will be preserved as-is from one generation to the next.

As seen in Fig. 6, we modified the standard elitism in NEAT, originally defined as the process where the best-fit M individuals in each species are preserved as-is from one generation to the next (Fig. 6 left). In the case of excessive elitism, the best-fit M individuals become elite regardless of their species within the whole population of N (Fig. 6 right), under the assumption of an excessive elitism, M . Their genotypes are carried forward to the succeeding generation, with the evaluated fitness values also passed on and used as their fitness value instead of re-evaluating the best-fit individuals in subsequent generations. As such, only the non-elite individuals ($N-M$) generated through selection

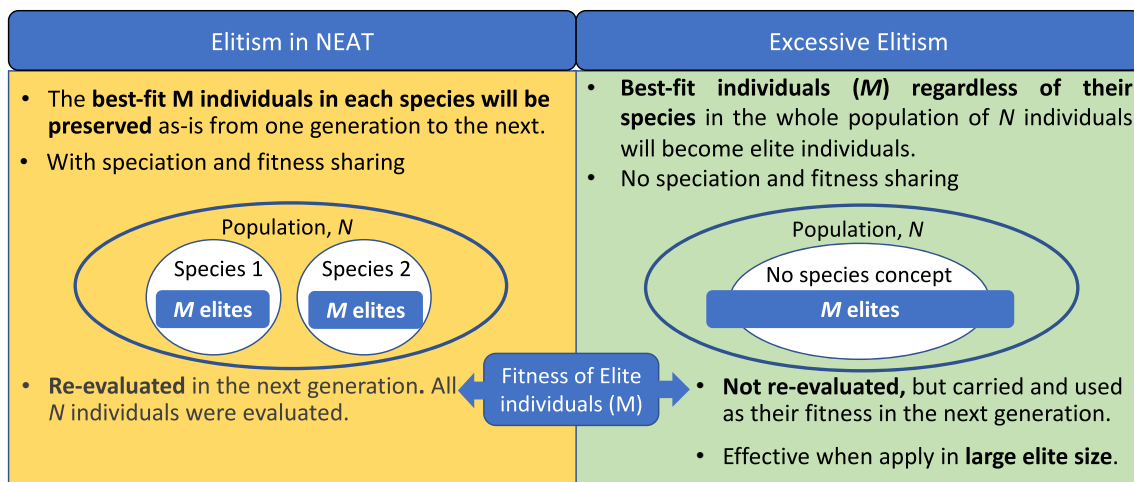


Fig. 6 A comparative analysis between the excessive elitism (EE) method and the standard elitism in NEAT

and genetic operations require evaluation at each generation. In the EE method, concepts of speciation and fitness sharing are turned off, thereby eliminating the species concept during evolution. We implement a roulette wheel selection method to stochastically select parents for reproduction, with being proportional to the fitness within the population. This approach offers a simpler alternative that leads to a simple population structure. More specifically, the excessive elitism (EE) method could affect the algorithm by giving priority to preserving the best-fit individuals within a large portion of the population for the purpose of exploration. The large population capacity (N) in EE enhances the diversity, significantly improving the population’s fitness. Additionally, replacing only a few individuals with offspring in each generation contributes to lower computational costs.

This substantially reduces evaluation costs for each generation, especially if the number of elite individuals (M) is large. Furthermore, we anticipate that maintaining high diversity in a large population can prevent the population from being stuck in local optima, though the rate of evolution may decelerate due to fewer individuals evolving ($N-M$). We adopted the following process: All N individuals were placed in a 3D simulation field; however, only the non-elite individuals ($N-M$) underwent development and movement through their lifetime for fitness evaluation. In contrast, the elite individuals (M) remained static, undergoing no development or movement. Each individual was evaluated in the initial generation. Related codes for the model can be found in the supplementary materials.

3 Experiments and results

We conducted the evolutionary experiments of artificial creatures by implementing HyperNEAT algorithm, utilizing the PUREPLES pure python library. Several key parameters from the default configuration were modified to improve our model’s efficiency in evolving the artificial creatures adapted to our specific evolutionary model. The contributing parameters are presented in Table 1.

We have set the fitness criterion to not exceed a maximum threshold of 10,000. This precaution ensures that the evolutionary process does not terminate before completing a generation's time step. It also serves to maintain the continuity of evolution, irrespective of whether certain individuals have already attained the fitness of 60. The ‘reset on extinction’ is set to false to avoid creating a completely random population should total extinction occur. The number of species that will be protected from stagnation is set to 1 (species elitism) to prevent total extinction. For more details see reference [8].

We conducted four evolutionary experiments, each with distinct yet comparable parameters. We applied the

Table 1 Modified parameters in NEAT’s configuration

Configuration	Parameters
Fitness criterion	Max
Fitness threshold	10,000
Reset on extinction	False
Species elitism	1
Elitism (default)	2

Table 2 Configuration of N and M for NEAT (2, 20), EE (2, 20), EE (82, 100), and EE (18, 20)

Case	Population size, (N)	Elite size, (M)	Evaluating individuals, ($N - M$)
NEAT (2, 20)	20	2	20
EE (2, 20)	20	2	18
EE (82, 100)	100	82	18
EE (18, 20)	20	18	2

Excessive Elite method in the cases of EE (2, 20), EE (82, 100), and EE (18, 20). Case NEAT (2, 20) was conducted to demonstrate the evolutionary process using the existing NEAT selection method.

Table 2 presents the parameters of population (N), elite size (M), and the number of evaluating individuals ($N-M$) in each specific case. For each case, we conducted 20 trials to obtain the most precise average results for fitness and block numbers. As indicated in Table 2, to compare the advantages of EE and NEAT, we executed NEAT (2, 20) with speciation and fitness sharing enabled as in Fig. 6. We assumed EE (2, 20) as a baseline case of EE which featured a smaller population size ($N=20$) and smaller elite size ($M=2$). This configuration is typically used in evolutionary experiments when evaluation costs need to be minimized, and it resulted in average fitness to reach local optima (explained later). To overcome the problem, we increased the population size to $N=100$ and added an excessive elite size of $M=82$ in the case of EE (82, 100), keeping the number of evaluating individuals consistent at $N-M=18$. We expect that there might be an improvement in the fitness increase. Finally, we conducted EE (18, 20) to demonstrate the use of excessive elitism, where $M=18$, while maintaining the smaller population size as in EE (2, 20) case. We expected that this case would also lead to an increased average fitness.

We conducted each experiment over 400 generations to standardize our findings and utilize them for further comparisons in the subsequent sections. We ran the simulation using AMD Ryzen 9 5950X 16-Core Processor and 128 GB memory PC.

3.1 Overall results

The overall results are summarized in Fig. 7. The graphs in Fig. 7(i) showed the evolution of the average fitness for each trial (gray), and their average (blue) and best (orange) for each case. Figure 7(ii) represents the corresponding evolution of the number of blocks within individuals at the end of their lifetime. The standard deviation was obtained as the objective measure for the diversity of morphology in the population (Fig. 7(ii)). For each trial, we determined the standard deviation in block numbers among the individuals within the population. We then computed and plotted the average of these standard deviations across 20 trials. Figure 7(iii) shows the trajectory of evaluated individuals in the last generation of the best trial. For a balanced comparison of fitness increase among the cases, we adjusted the *x*-axis scale based on the evaluation cost, calculated as computation time per generation.

The data presented demonstrate a significant increase in fitness when EE was implemented, almost reaching the maximum value and successfully reaching the target ((EE (82, 100) and EE (18, 20)). Conversely, fitness tended to converge towards an intermediate value in cases with a smaller elite size such as EE (2, 20). The total number of blocks showed a tendency to first increase more than around 10 before evolving to smaller values. In the following sections, we will conduct a comprehensive comparison of these cases.

3.1.1 Evolution of artificial creatures with the original NEAT algorithm (NEAT (2, 20))

We implemented the NEAT algorithm with speciation and fitness sharing enabled in the case of NEAT (2,20) by setting the minimum species size to 2 and the compatibility threshold to 0.03.

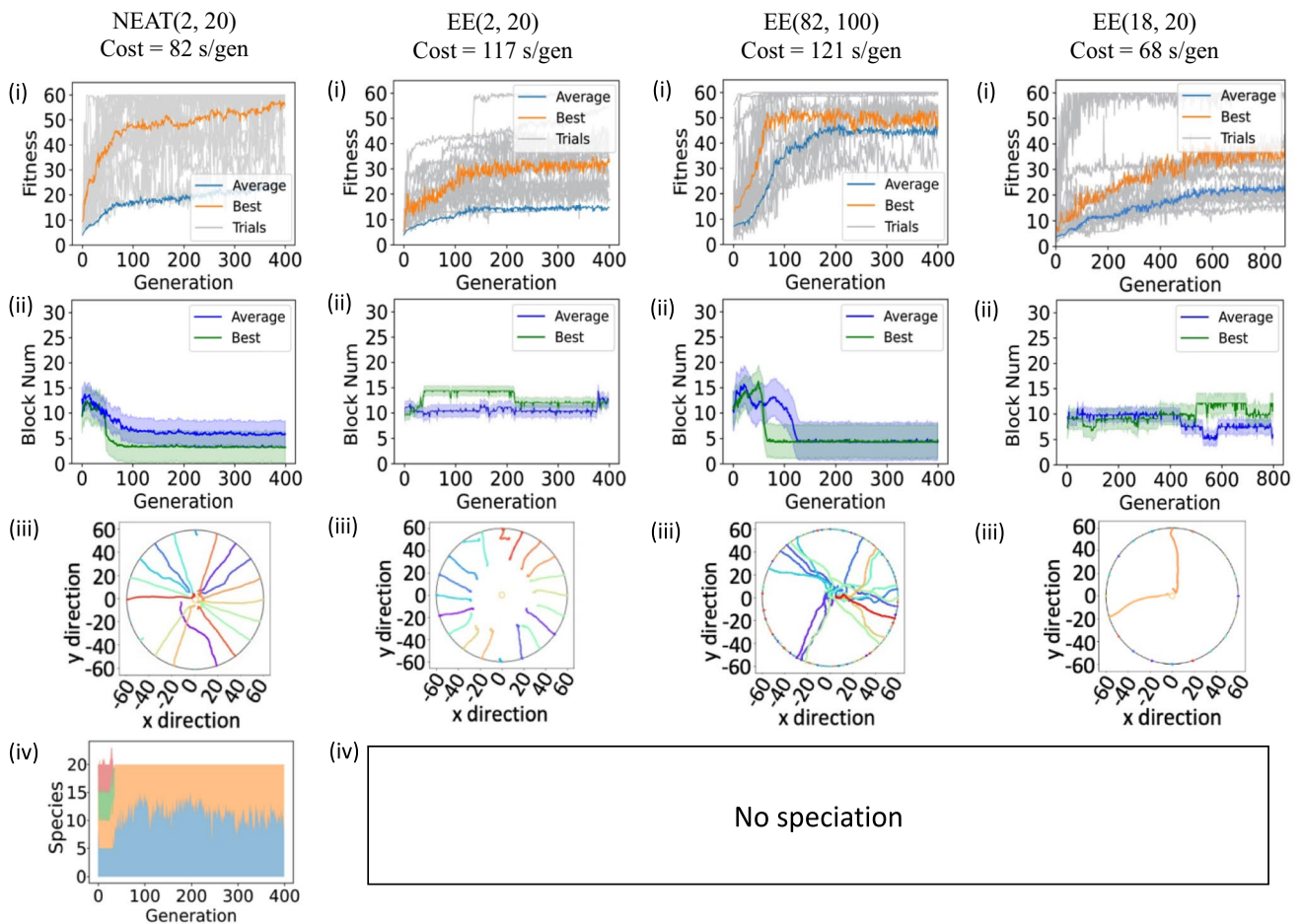


Fig. 7 Comparative results for NEAT (2, 20), EE (2, 20), EE (82, 100), and EE (18, 20). Sub-figure (i) portrays the evolution of fitness across 20 trials (with gray indicating the best fitness in each trial, blue signifying the average of average fitness in each trial, and orange representing the average of best fitness in each trial). Sub-figure (ii)

presents the number of blocks (blue for average, green for best). Sub-figure (iii) illustrates the trajectory of individuals in the last generation of the best trial. Sub-figure (iv) highlights the speciation process during evolution with NEAT, with each color symbolic of a unique species within a population

As shown in Fig. 7 (NEAT (2, 20)-i), most trials showed successful evolution of the population and most creatures were able to reach the target. The increase in fitness was comparable to the evolution with EE (82, 100) depicted in (Fig. 7 (EE (82, 100)—i)). Both cases demonstrated high morphological diversity, as can be seen in Fig. 7 (NEAT (2, 20)—ii) and Fig. 7 (EE (82, 100)—ii), with the best morphology adapting to approximately 3–5 blocks. The evaluation costs for EE (82, 100) and NEAT (2, 20) were comparable, at 121 s/gen and 82 s/gen respectively.

We can conclude that NEAT enabled the population to keep diversity with enabling speciation and fitness sharing, which is expected to contribute to the performance. While EE method had realized equivalent diversity in a simple single population.

3.1.2 Effects of the excessive elite size in increased population size (EE (2, 20) and EE (82, 100))

We conducted a comparative analysis between EE (2, 20) with EE (82, 100). In this comparison, EE (2, 20) served as the baseline experiment of EE, which had previously served as the solution to reduce the evaluation cost by proposing a smaller population size (N), yet this approach caused premature convergence. To overcome this limitation, we increased the population size to $N=100$ and applied an excessive elite size of $M=82$ into EE (82, 100). In this case, we maintained the same evaluating individual size ($N-M=18$), allowing us to compare the changes in the performance of the evolved population.

As depicted in (Fig. 7 (EE (2, 20)—i)), numerous trials showed that the fitness faced premature convergence and

stuck in the local optima of around 20. The evaluation cost was 117 s/gen. In contrast, EE (82, 100) showed an increase in fitness (Fig. 7 (EE—i)), with fitness reaching the maximum value (60) in many cases. The evaluation cost was comparable at 121 s/gen with that in EE (2, 20). This slight difference in evaluation cost between 121 s/gen and 117 s/gen might be due to the computational cost for simulating the whole physical field regardless of the number of evaluating individuals.

Despite the size of the evaluating individuals being the same for both cases ($N-M=18$), as seen in Fig. 7 (EE (82, 100)—ii), case EE (82, 100) led to increased morphological diversity as demonstrated by the standard deviation. The increased diversity significantly contributed to a greater increase in the average fitness. It further affected the course of morphological evolution, resulting in creatures in EE (2, 20) that tended to obtain around 10–15 blocks, as opposed to the fewer average of 3–5 blocks found in creatures in EE (82, 100), as highlighted in Fig. 7 (EE (2, 20)—ii) and (EE)—ii). Hence, the larger diversity seen in EE (82, 100) facilitated better morphological and behavioral evolution in the creatures, encouraging them to travel towards the target. Conversely, the lower diversity in EE (2, 20) led to a poor morphological development, hindering creatures from generating forward movements as shown in (Fig. 8 (EE (2, 20))) and (Fig. 8 (EE (82, 100))).

3.1.3 Effects of the excessive elite size in the small population (EE (2, 20) and EE (18, 20))

We conduct a comparison between EE (2, 20) and EE (18, 20) to assess the positive impact of the proposed elitism

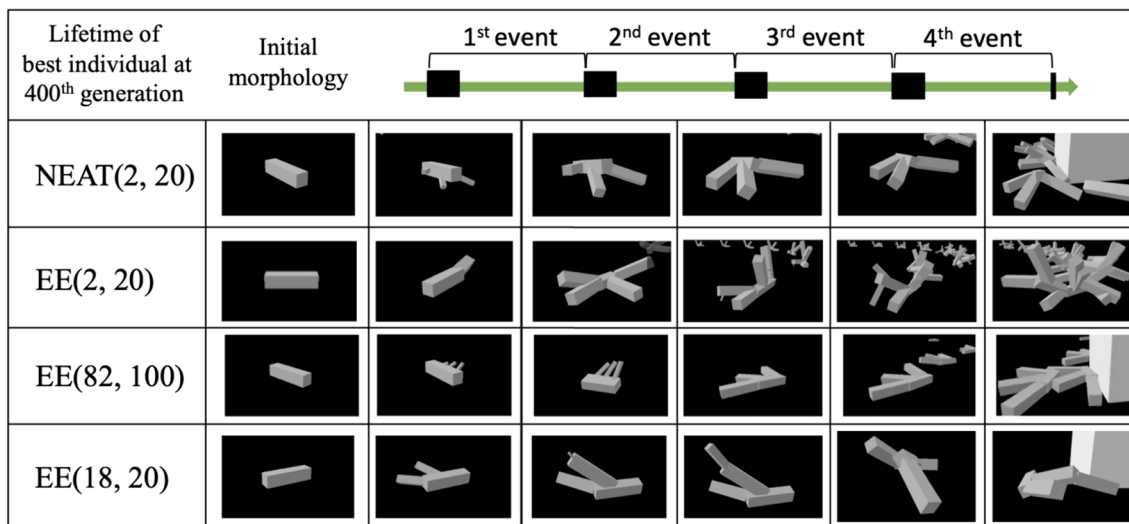


Fig. 8 Illustration of the lifetime of successful individuals during the final (400th) generation, featuring examples of morphology throughout each of four recurring events. These examples highlight the mor-

phological development and behavioral cycles in NEAT (2, 20), EE (2, 20), EE (82, 100), and EE (18, 20)

when the population size was relatively small ($N=20$). The elite size was increased from $M=2$ in EE (2, 20) to $M=18$ in EE (18, 20). As depicted in Fig. 7 (EE (18, 20)—(i)), the average fitness in EE (18, 20) witnessed a substantial improvement (around 30), which were slightly better than in EE (2, 20) (around 20). Many trials in EE (18, 20) are able to reach maximum fitness, successfully approaching the target object. In addition to that, due to the small number of evaluating individuals ($N-M=2$) in EE (18, 20), the average evaluation cost was significantly reduced to 68 s/gen.

In conclusion, EE demonstrated efficiency in evolution within both large population (EE (82, 100)) and small population size (EE (18, 20)). However, its effectiveness is more prevalent in larger populations as has been discussed in the previous sections. This certainly signifies that EE represents a flexible evolutionary algorithm in the field of evolutionary computation.

3.2 Analysis of evolved morphology

Figure 9 shows the morphological development of best individual, evolving from the first generation (on the left) to the final generation (on the right). The figures were collected from snapshots of best individuals during the last generation in each case. The distinct forms of the creatures from each case demonstrated that varying parameters conditions can induce evolutionary processes resulting in differing morphological developments, which leads to diverse

behavioral patterns. Videos illustrating the behaviors and morphological developments of the creatures in each case during the final generation, as demonstrated in Fig. 8, are available in the supplementary materials.

In conclusion, the population tended to obtain lower performance when individuals were composed of a larger average number of blocks, which might have hindered the movements. Conversely, higher performance was achieved when individuals were comprised of fewer number of blocks, potentially encouraging better freedom of movement to reach the target.

It should be noted that the average number of blocks once increased (around 10–15) in early generations across all cases. This is attributable to the NEAT algorithm, where the initial genes were randomly generated, creating a uniform population of simple networks. This likely resulted in a larger average number of blocks in these early generations. We expect that, at first, individuals with more blocks had dominated the population due to an evolutionary tendency. However, these large-sized individuals were too complicated to evolve further and obtain more adaptable morphological structures, leading to premature convergence in cases with smaller elite size. On the other hand, excessive elite sizes maintained diversity within the population which enabled individuals with fewer blocks to survive. This also offered them an opportunity to evolve their morphology and behavior to become simpler and more adaptable.

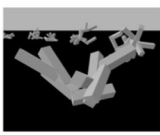



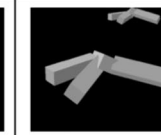






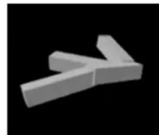
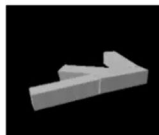
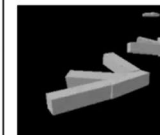

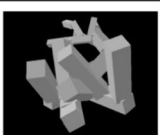
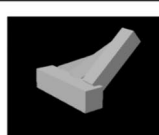
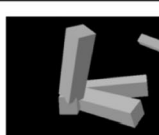


Generations	0th gen	→ 100th gen	→ 200th gen	→ 300th gen	→ 400th gen
NEAT(2, 20)					
EE(2, 20)					
EE(82, 100)					
EE(18, 20)					

Fig. 9 Observations of morphological transformations and behavioral patterns that the most successful individuals adopt throughout various generations in NEAT (2, 20), EE (2, 20), EE (82, 100), and EE (18, 20)

Table 3 Configuration of N and M for all cases

Cases EE (M, N)	Popula- tion size, N	Elite size, M	Evaluating individuals, ($N - M$)
EE (2, 20)	20	2	18
EE (22, 40)	40	22	
EE (42, 60)	60	42	
EE (62, 80)	80	62	
EE (82, 100)	100	82	
EE (102, 120)	120	102	

3.3 Finding optimal values of elite size (M) and population size (N) to successfully facilitate the evolution

As deduced from the results of the cases conducted in 3.1, it has been determined that the evolution of creatures is most efficient in case EE (82, 100) for larger population sizes and EE (18, 20) for smaller population sizes. However, it is still uncertain what the optimal values for the elite size, M , and the population size, N , would be for the best adaptation to the current environment. Therefore, we proceeded to conduct experiments using various configurations of N and M , as shown in Table 3.

We maintained the value of evaluating individual ($N - M$) at 18 for all cases, based on the best case in 3.1 and 3.2 (EE (82, 100)). The values that changes were the population size (N) and elite size (M) to see how it affects the evolution of creatures. The experiments were carried out for 200 generations for all cases and repeated to 20 trials for each case. The results of average best fitness and average computational time were calculated from the average of 20 trials in each case. These were then plotted against the population size (N), as shown in Fig. 10. The error bars present on the average best fitness depict the maximum standard deviation of the average number of blocks in the population. This serves as a representation of the morphological diversity present in each case. This arrangement allows for a clearer observation of the optimal configurations that most effectively evolved the creatures.

Analyzing the results illustrated in Fig. 10, we observed that the average computational time ranged from 117 to 125 s/gen when the evaluating individuals ($N - M$) remained consistently at 18. The average best fitness portrayed a slight improvement, from approximately 25 in (EE (2, 20)) to roughly 30 to 35 in (EE (22, 40)), and this value then remained constant for EE (42, 60). On the contrary, the average time experienced a slight increase from 117 s/gen to approximately 118 s/gen in the case of EE (42,60). Starting from EE (42, 60), we observed a vague increase in the average time, between 120 and 125 s/gen in the case

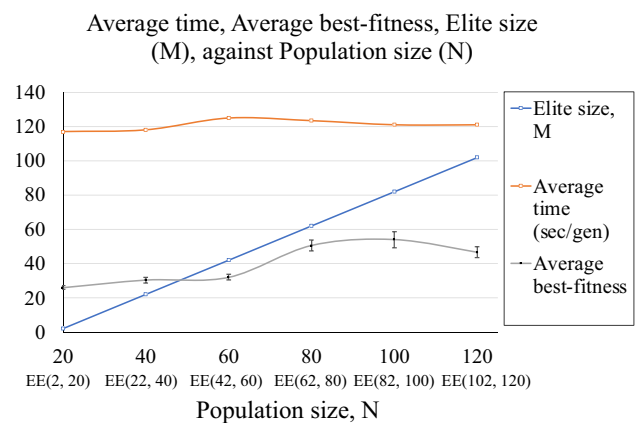


Fig. 10 A depiction of the correlation between the average evaluation time, average best fitness, and elite size (M) with respect to population size (N). The error bars present on the average best fitness are demonstrative of the standard deviation, representing the morphological diversity in each case

of EE (82, 100). However, a significant improvement was noticeable in the average best fitness in the EE (62,80) case, with a slight increase also seen for EE (82,100).

Referring to results in Fig. 10, larger population size cases often result in larger diversity of morphology as represented by the standard deviation on the average fitness in Fig. 10 (EE (82, 100)), where it shows the largest diversity obtained among other cases. We can see a slight decrease in diversity for smaller population case of EE (62, 80), followed by EE (42, 60), EE (22, 40), and EE (2, 20). Hence, larger morphology diversity was achieved in cases with a larger population, N . However, surprisingly, morphological diversity did not continue to improve for case EE (102, 120), although this case has the largest population size among other cases. Due to this, the average fitness slightly decreases to around 40–50. We strongly expect that EE (102, 120) may have an extremely large ratio of elite to evaluating individual size ($M:N$), within the population which potentially leads to a decrease in diversity, thus causing premature convergence. This might be due to the evolution of only the lowest-fitness individuals, when the larger portion of the population (102:18) is exempted from evaluation. We believe that the evolution of this case is too slow to improve that it converges to the local solution.

Therefore, we observed that the best average fitness of the evolved population, which was around 50 to 60, was obtained at EE (82, 100) as demonstrated in 3.1. Consequently, the morphological development is as in Fig. 9 (EE (82, 100)), while the behavior of creatures during final generation resembles those in Fig. 8 (EE (82, 100)). A video showcasing the behavior of creature in EE (82, 100) is also available in the supplementary videos.

4 Conclusions

We proposed the Excessive Elitism (EE) method, based on a novel use of elitism in NEAT. EE method served as a simpler alternative to speciation and fitness sharing in the existing NEAT algorithm, yielding equivalent morphological diversity that prevents premature convergence. We evolved the morphology and behavior of artificial creatures in a 3D multi-agent environment, with a simple target-approaching task. In the proposed EE method, we assumed that the evaluated fitness of the best-fit individuals would be succeeded and reused, instead of being re-evaluated during subsequent fitness evaluations. This in turn contributes to a lower evaluation cost.

As has been demonstrated, the excessive elite size surprisingly improved the average fitness of the population in both small and large population cases. Moreover, the effect of excessive elitism in large population cases significantly reduced the evaluation cost due to the reduced number of individuals being evaluated. Nonetheless, there are limitations on the ratio of elite to evaluating individuals that can be implemented in a population to obtain the optimum performance for our model, as proven in results in 3.3. We have demonstrated that the optimum size for elite and evaluating individuals would be similar to what has been shown in EE (82, 100).

Thus, by applying the EE method, we shall use the advantage of a large population size and enhanced diversity to achieve optimum fitness whilst keeping the evaluation cost minimal. Nevertheless, it is worth noting that certain non-adaptive individuals with initial conditions may accidentally achieve high fitness. These types of problems have been discussed in the evolutionary context [13] and in relation to evolutionary algorithms (EAs) at large [14]. Despite these challenges, the benefits of the EE method significantly outbalance such drawbacks. Additionally, EE method bears similarities to quality-diversity approaches like MAP-Elites, which captures diverse high-performing solutions by partitioning the search space [15]. In this study, the EE method demonstrates the simplest means of securing a niche and preserving diversity within the general framework of a simple genetic algorithm.

In future work, we intend to acknowledge whether the source of the shown performance, derived from the complexity in the search space, is related with the asymmetry between morphology and control [16, 17]. Furthermore, it is crucial to explore to what extent would the excessive elitism method improves the evolution of artificial creatures, while keeping the limitations minimal and maintaining conditions that would facilitate the evolution of a more complex interactions between artificial creatures and their environment through the implementation of niche

constructing behavior. We planned to incorporate various elements into the evolutionary process. Expectations lie in evolving each creature with an exemption from interactions for simplicity, to exclusively analyze multiple levels of adaptive mechanisms—these include the developmental evolution of creatures, the effects of niche construction, and ecological inheritance.

Acknowledgements This work was accomplished under “The Education Program for Data Scientist in Future Value Creation” and was supported in part by JSPS/MEXT KAKENHI 21K12058.

Data and code availability The supplementary materials for this article can be found at the following: Source code: <https://doi.org/10.5281/zenodo.10450560>. Supplementary Videos: Morphological changes and patterns of movements for best individual in EE (2, 20) and EE (82, 100)—<https://figshare.com/s/7f467ac2e8aa7c40f21d>

References

- Schmitt F, Piccin O, Barbé L, Bayle B (2018) Soft robots manufacturing: a review. *Front Robot AI*. <https://doi.org/10.3389/frobt.2018.00084>
- Wehner M, Truby RL, Fitzgerald DJ, Mosadegh B, Whitesides GM, Lewis JA, Wood RJ (2016) An integrated design and fabrication strategy for entirely soft, autonomous robots. *Nature* 536(7617):451–455. <https://doi.org/10.1038/nature19100>
- Chiba N, Suzuki R, Arita T (2020) Evolution of complex niche-constructing behaviors and ecological inheritance of adaptive structures in a physically grounded environment. *Front Robot AI*. <https://doi.org/10.3389/frobt.2020.00045>
- Giráldez R, Díaz-Díaz N, Nepomuceno I, Aguilar-Ruiz JS (2005) An Approach to reduce the cost of evaluation in evolutionary learning. In: *Proceedings of computational intelligence and bio-inspired systems, 8th international work-conference on artificial neural networks*, pp 804–811
- Stoy K (2022) Efficiency through GPU-based co-evolution of control and pose in evolutionary robotics. In: *Proceedings of the ALIFE 2022: the 2022 conference on artificial life*, paper no.: isal_a_00558, 2
- La Cava W, Moore J (2018) Behavioral search drivers and the role of elitism in soft robotics. In: *Proceedings of the ALIFE 2018: the 2018 conference on artificial life*, pp 206–213
- Methenitis G, Hennes D, Izzo D, Visser A (2015) Novelty search for soft robotic space exploration. In: *Proceedings of the 2015 genetic and evolutionary computation conference*, pp 193–200
- Stanley KO, Miikkulainen R (2002) Efficient evolution of neural network topologies. In: *Proceedings of the 2002 congress on evolutionary computation*. CEC'02, pp 1757–1762
- Stanley KO, D'Ambrosio DB, Gauci J (2009) A hypercube-based encoding for evolving large-scale neural networks. *Artif Life* 15(2):185–212. <https://doi.org/10.1162/artl.2009.15.2.15202>
- Stanley KO (2007) Compositional pattern producing networks: a novel abstraction of development. *Genet Program Evol Mach* 8:131–162. <https://doi.org/10.1007/s10710-007-9028-8>
- Jaafar SA, Suzuki R, Komori S, Arita T (2023) How large elite size can facilitate the evolution of artificial creatures with NEAT. In: *Proceedings of the joint symposiums of AROB-ISBC-SWARM 2023*, pp 142–147
- Jaafar SA, Suzuki R, Komori S, Arita T (2023) How excessive elitism can facilitate the evolution of morphology and behavior of artificial creatures with NEAT. In: *Proceedings of the ALIFE*

- 2023: the 2023 conference on artificial life (3-page abstract) (**in press**)
13. Montanier JM, Bredeche N (2011) Embedded evolutionary robotics: the (1+1)-restart-online adaptation algorithm. *Stud Comput Intell* 341:155–169. https://doi.org/10.1007/978-3-642-18272-3_11
 14. Jin Y, Branke J (2005) Evolutionary optimization in uncertain environments—a survey. *IEEE Trans Evol Comput* 9(3):303–317. <https://doi.org/10.1109/TEVC.2005.846356>
 15. Nordmoen J, Veenstra F, Ellefsen KO, Glette K (2021) MAP-Elites enables powerful stepping stones and diversity for modular robotics. *Front Robot AI*. <https://doi.org/10.3389/frobt.2021.639173>
 16. Ito T, Pilat ML, Suzuki R, Arita T (2013) ALife approach for body-behavior predator-prey coevolution: Body first or behavior first? *Artif Life Robot* 18:36–40. <https://doi.org/10.1007/s10015-013-0096-y>
 17. Cheney N, Bongard J, Sunspirai V, Lipson H (2016) On the difficulty of co-optimizing morphology and control in evolved virtual creatures. In: *Proceedings of the ALIFE 2016, the fifteenth international conference on the synthesis and simulation of living systems*, pp 226–233

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.