

Insights in evolutionary exploration of robot morphology spaces

1st Karine Miras

Dept. of Comp. Sci.

Vrije Universiteit Amsterdam

Amsterdam, The Netherlands

k.s.m.a.dasilvamirasdearaujo@vu.nl

2nd Arwin Gansekoele

Dept. of Comp. Sci.

Vrije Universiteit Amsterdam

Amsterdam, The Netherlands

awgansekoele@gmail.com

3rd Kyrre Glette

Department of Informatics

University of Oslo

Oslo, Norway

kyrrehg@ifi.uio.no

4th A.E. Eiben

Dept. of Comp. Sci.

Vrije Universiteit Amsterdam

Amsterdam, The Netherlands

a.e.eiben@vu.nl

Abstract—In a recent study we have encountered an unexpected result regarding the evolutionary exploration of robot morphology spaces. Specifically, we found that an algorithm driven by selection based on morphological novelty explored *fewer* spots in the space of morphologies than another algorithm based on a combination of morphological novelty and some behavioral criterion (speed of movement). Here we revisit these results, perform new analyses, and obtain new insights. These insights clarify the exploration behavior of these algorithms and provide guidelines for designing selection mechanisms for evolutionary robotics.

Index Terms—evolutionary robotics, morphological evolution, search space exploration, diversity, novelty search

I. INTRODUCTION

In this paper we consider evolutionary robot systems where both the bodies (morphologies) and the brains (controllers) of the robots are evolvable, with a special interest in the evolved bodies. Our long term research goal is to understand how evolutionary search traverses spaces of robot morphologies. This issue is relevant for fundamental reasons as well as for practical purposes. Knowledge about possible search biases or regions of attraction in the morphology space can help design appropriate evolutionary operators for selection and reproduction.

To this end, we have identified several morphological properties, e.g., size, symmetry, branching, that together span a multidimensional search space for evolution in [1]. In a follow-up study we have investigated how these evolved morphological properties change when using different criteria for selection, cf. [2]. In particular, we specified selection mechanisms based on morphological properties (novelty of the body w.r.t. existing robots), robot behavior (speed of movement),

and the combination of these. And, when examining what proportions of the search space were explored under each of these regimes, the results were counter-intuitive. The selection mechanism based purely on morphological diversity exhibited inferior capacity of exploring the space than the one that combined this same criterion (diversity) with a behavioral criterion (speed). Thus, in this paper we investigate this effect further to understand what caused this unexpected system behavior.

This paper is organized as follows. Section II introduces a background for this study. Section III elucidates the design space of the robots, regarding their morphologies and controllers. Section IV explains the encoding, evolutionary operators, evolutionary algorithm, and fitness function used in the experiments. Section V describes how the data was obtained and how the analysis was conducted. Section VI exposes and discusses the results. Finally, Section VII presents our conclusions.

II. RELATED WORK

Novelty search was introduced in [3], [4], as a method for overcoming the deception of local optimum. They compared the use of fitness-based search (selecting for behavioral quality) with behavioral novelty search, and obtained much higher behavioral quality in a maze navigation task when using novelty. Furthermore, they utilized behavioral novelty in a bipedal locomotion task, one more time achieving better results for behavioral quality [5].

In [6] they applied and compared morphological novelty search for evolving virtual creatures that locomote. The performance of novelty search alone at the task was worse than the

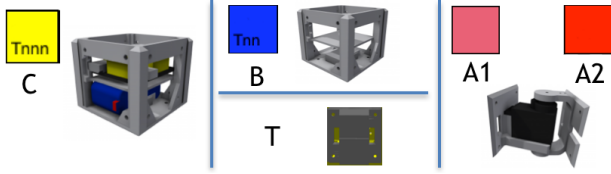


Fig. 1: Robot modules: Core-component with controller board (C); Structural brick (B); Active hinges with servo motor joints in the vertical (A1) and horizontal (A2) axes; Touch sensor (T). C and B have attachment slots on their four lateral faces, and A1 and A2 have slots on their two opposite lateral faces; T has a single slot which can be attached to any slot of C or B. The sequence of letters (T or n) in C and B indicate if there is a sensor on the laterals left, front, right and back (for C only), in this order.

fitness-based search, and this is reasonable, once the concept of novelty was purely morphological, not taking the behavior into consideration. The combination of morphological novelty search with a fitness-based search was also experimented as multiobjectivization, achieving a better performance at the task than the individual searches. Nevertheless, the best levels of diversity were, as expected, achieved by novelty alone.

Other studies have discussed the benefits and also the challenges of using novelty search, alone, or in combination with other fitness functions [7]–[9], and there still are open questions concerning when the use of novelty search is worth it. For instance, in [10] it was argued that “novelty search alone does not scale to large search spaces, but, when combined with fitness-based selection, it can be a useful diversity sustaining mechanism.”

III. DESIGN SPACE

A. Robot Morphology

The phenotypes of the morphologies (“bodies”) are composed of modules [11] as shown in Fig. 1. The morphologies (Fig. 2) are flat, i.e., the modules do not allow attachment on the top or bottom slots, but only the lateral ones. And, any module can be attached to any other through its attachable slots, except for the sensors, which can not be attached to joints. Each module type is represented by a distinct letter in the genotype and by a different color in the phenotype.

For analyzing the morphological properties of the robots, we used a framework proposed in our previous work [1], [2] containing nine morphological descriptors. The descriptors, for which a more detailed explanation can be found in our previous work, range from 0 to 1 and are named: Branching,

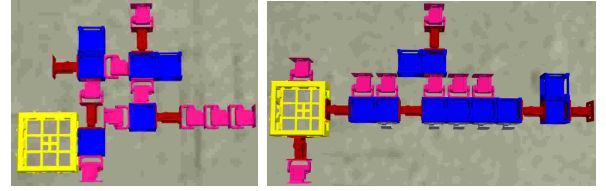


Fig. 2: Examples of evolved robots.

Number of Limbs, Length of Limbs, Coverage, Joints, Proportion, Symmetry, Size, and Sensors.

B. Robot Controller

The phenotype of the controller (“brain”) is a multilayer Artificial Neural Network, not necessarily fully connected (Fig. 3, right). For every joint in the morphology, there exists an equivalent Oscillator neuron in the network, and every sensor is reflected as an input of the network. The intermediate topology may vary for each robot and is composed of neurons which might have Linear or Sigmoid transfer functions. The weights of the connections range from -1 to 1 . In this paper, the neurons are also referred to as nodes and the connections as edges.

IV. EVOLUTION

A. Generative Encoding

The generative encoding to represent the genotypes of the robots is a grammatical parallel rewriting system called Lindenmayer-System [12], and conjointly includes elements relative to both morphology and controller, as in [13]. The grammar of an L-System is defined as a tuple $G = (V, w, P)$, where

- V , the alphabet, is a set of symbols containing replaceable and non-replaceable elements.
- w , the axiom, is a symbol from which the system starts.
- P is a set of production-rules for the replaceable symbols.

The following didactic example depicts the process of iterative-rewriting of an L-System. For a determined number of iterations each replaceable symbol is simultaneously replaced by the elements of its production-rule. Given $w = C$, $V = \{C, B, T\}$ and $P = \{C : \{C, B\}, B : \{T\}, T : \{C, T\}\}$, the rewriting follows as:

Iteration 0: C

Iteration 1: C B

Iteration 2: C B T

Iteration 3: C B T C T

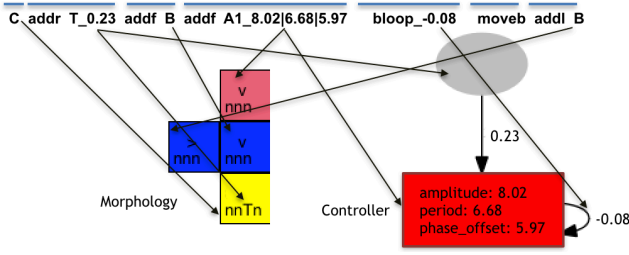


Fig. 3: Process of decoding an early-developed phenotype into a late-developed phenotype with morphology and controller.

Each genotype is a distinct grammar, making use of the same alphabet (Tab. I), and the alphabet is formed by types of morphology modules and commands to attach them together, as well as commands for defining the structure of the controller. To construct a robot, firstly (early development), the axiom of the grammar is rewritten into a more complex string of elements, according to the production-rules of the grammar (the parameter of number of iterations was set 3). Secondly (late development), this string is decoded into a phenotype. The decoding process of the phenotype of morphology and controller is illustrated in Fig. 3. During this construction phase, two references are always maintained in the phenotype, one for the morphology (pointing to the current module) and one for the controller (pointing to the current edge). The application of the commands to the phenotype happens in the current module for the morphology and in the current edge for the controller.

B. Evolutionary Operators

For all the experiments, the same evolutionary operators and parameters were applied. The population size was $\mu = 100$, being evolved for 100 generations. In each generation: an offspring $\lambda = 50$ was created by producing 1 individual from each of 50 binary tournaments for parent selection and mutating the new individual. From the resulting population of μ parents and λ offspring, 100 individuals were selected for survival, also using binary tournaments. For each fitness function, the experiments were repeated 10 times. The genotypes were initialized by adding four elements to each production-rule, one of each of the categories, Controller-moving, Morphology-mounting, Morphology-moving and Modules, and, each element of the categories was chosen randomly. The maximum amount m of modules allowed in a morphology was 100. So, during the phase of decoding the genotype into the phenotype, after reaching the maximum, the succeeding modules were

Modules

C	core-component
B	brick
A1	vertical joint
A2	horizontal joint
T	touch sensor

Morphology-mounting commands

addr	add the next module to the right
addf	add the next module to the front
addl	add the next module to the left

Morphology-moving commands

moveb	move reference to the module at the back
mover	move reference to the module to the right
movef	move reference to the module to the front
movel	move reference to the module to the left

Controller-change commands

bnode	add a new node to the neural network
bedge	add a new edge to the neural network
bperturb	perturb the weight of an edge
bloop	add a self-connection edge to the network

Controller-moving commands

bmVFTC	move current edge origin-reference to a child
bmVFTP	move cur. edge origin-reference to a parent
bmVFTS	move cur. edge origin-reference to a sibling
bmVTTC	move cur. edge dest.-reference to a child
bmVTTP	move cur. edge dest.-reference to a parent
bmVTTS	move cur. edge dest.-reference to a sibling

TABLE I: Alphabet of the grammars. (cur.=current and dest.=destination)

ignored. Additionally, morphologies without at least one joint or with intersecting parts were considered invalid, and though having been kept in the population, were not evaluated, receiving a value of zero to the measurement of speed. The crossovers were performed by taking the production-rules randomly from the parents, and all individuals underwent mutation by adding/deleting/swapping one random element from a random production-rule/position. The crossover probability was 100%, as it is possible that during the rewriting of the L-System, only the rules of one single parent end up being expressed. And, as it is not rare that one mutation happens for non-expressed genes, thus, to minimize this effect the mutation rate was 100%.

C. Fitness Functions

The fitness function for the diversity-oriented search is defined as $N = n$, where n is a measure of novelty which is calculated as the average distance to the k -nearest neighbors of an individual, for which $k = 1$ and the distance is the

Euclidean distance regarding the nine morphological descriptors. The set of neighbors for the comparison is formed by the current population, plus an archive, to which every new individual has a 10% probability of being added, with the individuals added to the archive remaining in it until the end. This fitness has a purely relative morphological property.

The behavior-oriented search is calculated as $S = s$, where s is a measure of the speed (m/s) of the displacement of the robot's head from its initial position to its final position during the evaluation time, having a purely behavioral character.

Finally, the combined search (SN) is calculated as $SN = s * n$, combining behavioral and relative morphological characters.

V. METHODS

The data analyzed in this paper was generated in a previous study [2], on which distinct experiments were conducted evolving robots for the behavior of undirected locomotion on a plain terrain, using a different fitness function in each experiment. The first was a diversity-oriented search, i.e., Novelty Search [3], the second was a behavior-oriented search, and the third was a combination of the other ones. The study left an open question regarding a counterintuitive observation that, a combination of a search for behavior with a search for morphological diversity had a greater capacity of exploring the search space than a search for sole morphological diversity. Therefore, grounded on these previous results, in this paper we deepened this analysis to understand this phenomenon.

A search space can be seen as different layers which build-up over each other. Those are, the design space, the representation, and the reproduction operators. These layers result in a set of phenotypes, to which we refer here as morphological phenotypic search space.

In the previous study, the exploration of the search space was examined by quantifying the sampling of points in the space, which was defined through a measure called Number of Sampled Cubes (NSC). The NSC is accumulated along the full evolutionary run and accounts for the number of cubes in the morphological multidimensional space discovered, that is, the number of distinct morphologies (within a determined granularity) ever found along the full course of evolution. The grid of cubes has its dimensions composed of a number x of morphological descriptors, having each one divided into y bins of equal size. This way, the number of cubes in the space is calculated as y^x . Every new morphology is attributed to a

proper cube given its descriptors, and if the cube does not contain any morphology yet, one more unit is summed to the NSC.

In this study, additionally to the NSC, we propose and use a measure called Average Spread (AS), targeting to quantify how distant from each other the points in space are. The calculation is done by, firstly, averaging the Euclidean distance (regarding x morphological descriptors - dimensions) from each individual to every other, and secondly, averaging these values across all individuals. The measure is calculated in four different versions, using both median and mean in the first and second level of aggregation, to make sure that from all averaging-perspectives possible differences can be observed.

VI. RESULTS AND DISCUSSION

In the previous study, the granularity used to observe the sampling of the search space was high, with $x = 9$ descriptors and $y = 100$ bins. We consider this level of granularity as high because it uses all of the nine descriptors as dimensions, divided in $y = 100$ bins, which in practice, is equivalent to almost all¹ of the robots set individually into one bin. In the current study, we investigated how the NSC changes when reducing this granularity. Thus, when considering fewer dimensions (descriptors) and/or fewer bins.

Our first step was reducing the bins-granularity to the lowest possible² value, i.e., $y = 3$. Figure 4 shows the progression of the average NSC with $x = 9$ (all dimensions) of all fitness functions for two different bins-granularity scenarios, i.e., $y = 100$ bins and $y = 3$ bins. The black vertical bars indicate the generation on which the differences in average among all functions become significant (Wilcoxon $p < 0.05$). As expected given the selection pressure of S being purely towards a behavior, in both scenarios its NSC is much lower than for the functions that include diversity preferences. Additionally, as it was known from the previous study, when having $y = 100$ bins, N , that searches only for diversity, curiously presents a lower NSC than SN , which combines diversity with behavior. This is counterintuitive, as one would initially expect that the later one, for including behavior as part of its preferences, would suffer the effect of the selection pressure for certain traits that are suitable for the task, being

¹Among thousands of individuals of all experiments, only two individuals of one same bin did not present precisely the same values for their descriptors.

²The granularity 2⁹, that is, $y = 2$ bins when $x = 9$ (all dimensions), is lower than the total number of individuals of one full evolution (5.050), and it would make no sense to use a granularity level that leaves any individuals out of the calculation.

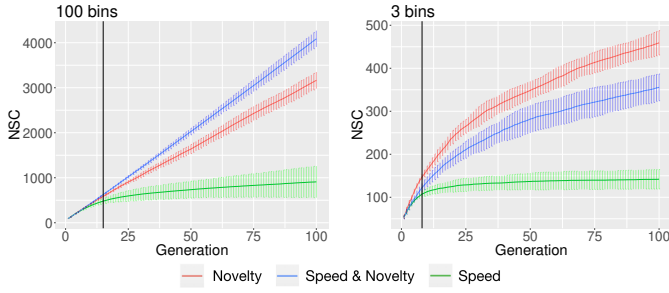


Fig. 4: Progression of NSC with $x = 9$ (all dimensions), throughout the generations of all fitness functions for two bins-granularity scenarios, i.e., $y = 100$ bins and $y = 3$ bins. Values are the mean of all runs.

less diverse than when searching purely for diversity. However, when having $y = 3$ bins, the situation is inverted, with N surpassing SN .

Our second step was reducing the dimensions-granularity (descriptors), measuring the progression of the NSC for all the groups of descriptors, considering both low and high bins-granularity. The groups were formed by making all possible combinations of descriptors, for every size (number of descriptors included) of a group, going from size 1 to size 9. Figure 5 shows the proportions of groups for which each fitness function presented a significantly (Wilcoxon $p < 0.05$) higher (superior) NSC than the other functions³. We see that by having $y = 100$ bins (Fig. 5, left), N is always superior when the number of dimensions is $x = 1$. And, the frequency of cases on which N is superior decreases as the number of dimensions increases. Furthermore, by having $y = 3$ (Fig. 5, right), N is superior in all cases. This shows that N is superior to SN when the granularity is low, and inferior when it is high.

Having these new insights in mind, we drew down into the distribution of the discovered points throughout the search space. And assessing how spread the points in space are, we obtained a much clearer understanding of the phenomenon. Figure 6 shows the AS (with $x = 9$, i.e., all dimensions) for each fitness function⁴. In all cases we see an evident difference in the average distances. As expected, the points of S present the lowest average distances. But when comparing SN with N , the later presents higher distances amongst the points.

After all, although SN has a greater capacity for sampling the search space, the points sampled by N are much more different from each other. Therefore, while N is superior in

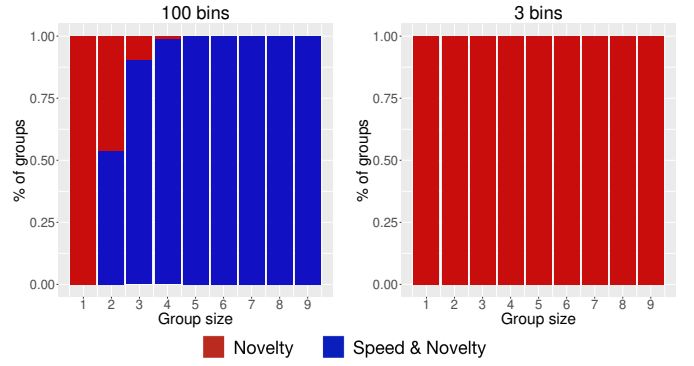


Fig. 5: For all sizes of group of descriptors, proportions of cases on which each fitness function was significantly (Wilcoxon $p < 0.05$) superior than the others in sampling the search space.

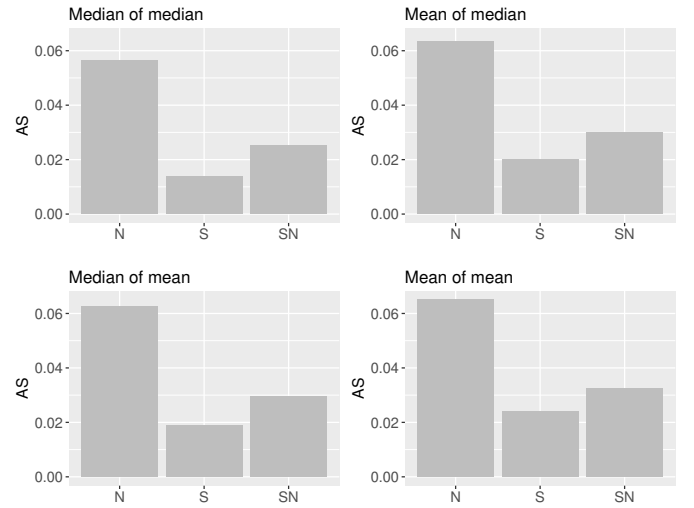


Fig. 6: Comparison of AS among fitness functions. Differences among outer means are significant (Wilcoxon $p < 0.001$). N = Novelty, S = Speed, SN = Speed & Novelty

discovering novelty for a global search, SN is superior in finding novelty for a local search.

To illustrate this difference in the spreading of the points in space, we plotted the points for the pairs of dimensions (2 descriptors) with which any fitness function significantly (Wilcoxon $p < 0.05$) surpassed all the others in sampling the search space (NSC). A higher number of dimensions would not be feasible to visualize, but two dimensions elucidate enough the effect of the differences in distance verified previously. Figure 8 depicts the pairs for which N is superior regarding NSC, while figure 9 depicts the cases for which SN is superior (S was not superior in any case). For all displayed pairs, the plots show two distinct perspectives, the centroids; and all

³Cases of insignificant differences were not included.

⁴This includes individuals of all evolutionary runs.

the individuals. For both cases, the results of the 10 repeated experiments were merged in each generation. The centroids were calculated as the mean values of the descriptors among all individuals of each generation within the 9 morphological descriptors (dimensions). For all cases, regardless of the fitness function that is superior in sampling the search space (NSC), N always finds centroids more spread in space than SN . When considering individuals instead of centroids, this difference is visually less obvious, but it still exists.

Through these observations we conclude that N is able to find more novelty in terms of distance among points (AS), when SN finds more novelty in the sense of distinct points. This corroborates with the previous observation that when using a lower granularity, N surpasses SN regarding NSC. If SN discovers a superior number of points in space, but these points are very close to each other, and N discovers fewer points in space, but these points are much further from each other, this difference could only be observable by reducing the size of the cubes when measuring the space.

The reason for N discovering fewer cubes in space than SN is that N revisits the same cubes more often than SN . Figure 7 shows the frequency of visitation of the cubes in the space for all fitness functions. The cause of this revisitation is still unclear, but it could perhaps be related to parameters for the archive of the Novelty Search, or even the metric of distance used, given that other studies have argued the Euclidean distance is not always the best option [14].

In summary, the observed anomalous effect of SN (combination of a preference for morphological diversity with a preference for a given behavior), makes sense, once it included two contradictory pressures. On one side, the pressure for diversity, and on the other side, the pressure to the traits that better fit the task. This resulted in a level of exploration that surpasses N from one perspective (quantity) but falls behind N from another perspective (magnitude).

VII. CONCLUSION AND FUTURE WORK

In this paper we performed further analyses to understand an unexpected system behavior encountered in a previous study. We examined the effect of reducing the granularity with which one divides a search space into small cubes, and learned that it has a great impact on observed levels of space exploration. Through this we have gained a better understanding of how different fitness functions explore a robot morphological search space. The results showed that

search driven solely by morphological novelty discovers fewer points in space than search that combines preferences for morphological novelty and behavior. However, the former discovers points which are further from each other. This illustrates two possible definitions of diversity throughout the evolutionary process: exploring more points or exploring more different points. If the users priority is to encounter solutions as different from each other as possible (global search), then the novelty-driven search is appropriate. If encountering as many solutions as possible, even though they might be more similar (local search), then the combination of novelty with behavior is more suitable.

As future work, the reason why N revisits more cubes in the space than SN could be investigated. Interesting experiments would be the use of different parameters to build the archive, and the test of distinct metrics of distance.

VIII. ACKNOWLEDGEMENTS

This work has received funding from the European Union's PHOENIX project under grant agreement No 665347.

REFERENCES

- [1] K. Miras, E. Haasdijk, K. Glette, and A. E. Eiben, "Search space analysis of evolvable robot morphologies," in *Applications of Evolutionary Computation - 21st International Conference, EvoApplications 2018*, ser. Lecture Notes in Computer Science. Springer, 2018, vol. 10784, pp. 703–718.
- [2] —, "Effects of Selection Preferences on Evolved Robot Morphologies and Behaviors," in *Proceedings of the Artificial Life Conference 2018 (ALIFE 2018)*, T. Ikegami, N. Virgo, O. Witkowski, R. Suzuki, M. Oka, and H. Iizuka, Eds. Tokyo: MIT Press, 2018.
- [3] J. Lehman and K. O. Stanley, "Exploiting open-endedness to solve problems through the search for novelty," in *ALIFE*, 2008, pp. 329–336.
- [4] —, "Efficiently evolving programs through the search for novelty," in *Proceedings of the 12th annual conference on Genetic and evolutionary computation*. ACM, 2010, pp. 837–844.
- [5] —, "Abandoning objectives: Evolution through the search for novelty alone," *Evolutionary computation*, vol. 19, no. 2, pp. 189–223, 2011.
- [6] —, "Evolving a diversity of virtual creatures through novelty search and local competition," in *Proceedings of the 13th annual conference on Genetic and evolutionary computation*. ACM, 2011, pp. 211–218.
- [7] P. Krcak and D. Toropila, "Combination of novelty search and fitness-based search applied to robot body-brain co-evolution," in *Czech-Japan Seminar on Data Analysis and Decision Making in Service Science*, 2010, pp. 1–6.
- [8] S. Kistemaker and S. Whiteson, "Critical factors in the performance of novelty search," in *Proceedings of the 13th annual conference on Genetic and evolutionary computation*. ACM, 2011, pp. 965–972.
- [9] J.-B. Mouret, "Novelty-based multiobjectivization," in *New horizons in evolutionary robotics*. Springer, 2011, pp. 139–154.
- [10] G. Cuccu and F. Gomez, "When novelty is not enough," in *European Conference on the Applications of Evolutionary Computation*. Springer, 2011, pp. 234–243.

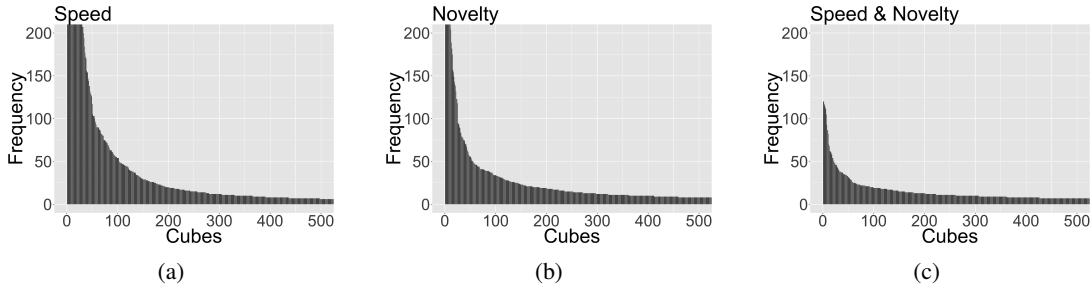


Fig. 7: Concentration of morphologies in each sampled cube of the morphological multidimensional space considering all runs. Including all morphologies explored, each bar represents the frequency of morphologies in one specific cube, and the bars are sorted in descending order. All charts are scaled on the y-axis limited to 200 and in the x-axis limited to 500.

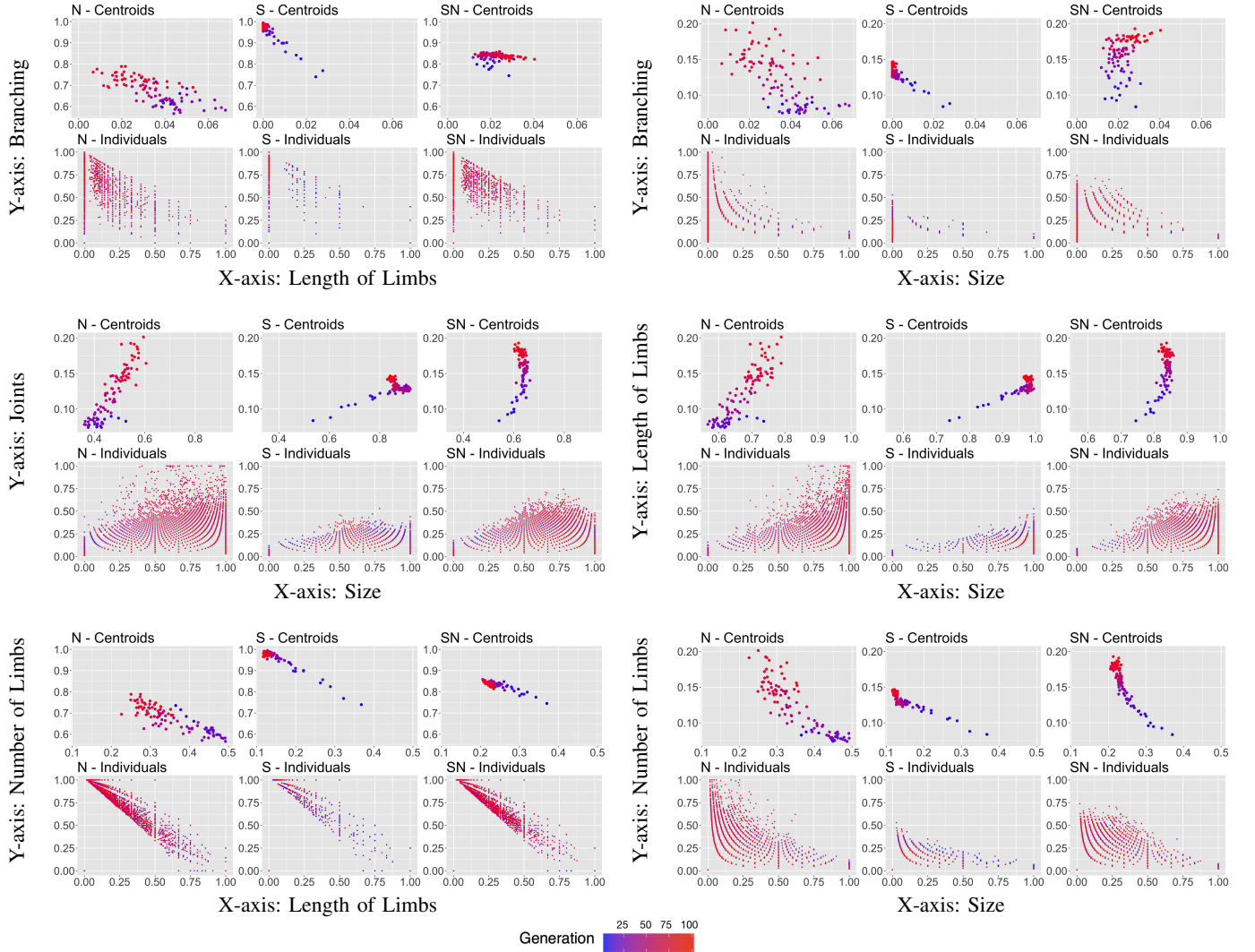


Fig. 8: Points in space of pairs of descriptors for which N had a significantly greater NSC than the other fitness functions. N = Novelty, S = Speed, and SN = Speed & Novelty.

[11] J. Auerbach, D. Aydin, A. Maesani, P. Kornatowski, T. Cieslewski, G. Heitz, P. Fernando, I. Loshchilov, L. Daler, and D. Floreano, "Robogen: Robot generation through artificial evolution," in *Artificial Life 14: Proceedings of the Fourteenth International Conference on the*

Synthesis and Simulation of Living Systems. The MIT Press, 2014, pp. 136–137.

[12] C. Jacob, "Genetic I-system programming," *Parallel Problem Solving from NaturePPSN III*, pp. 333–343, 1994.

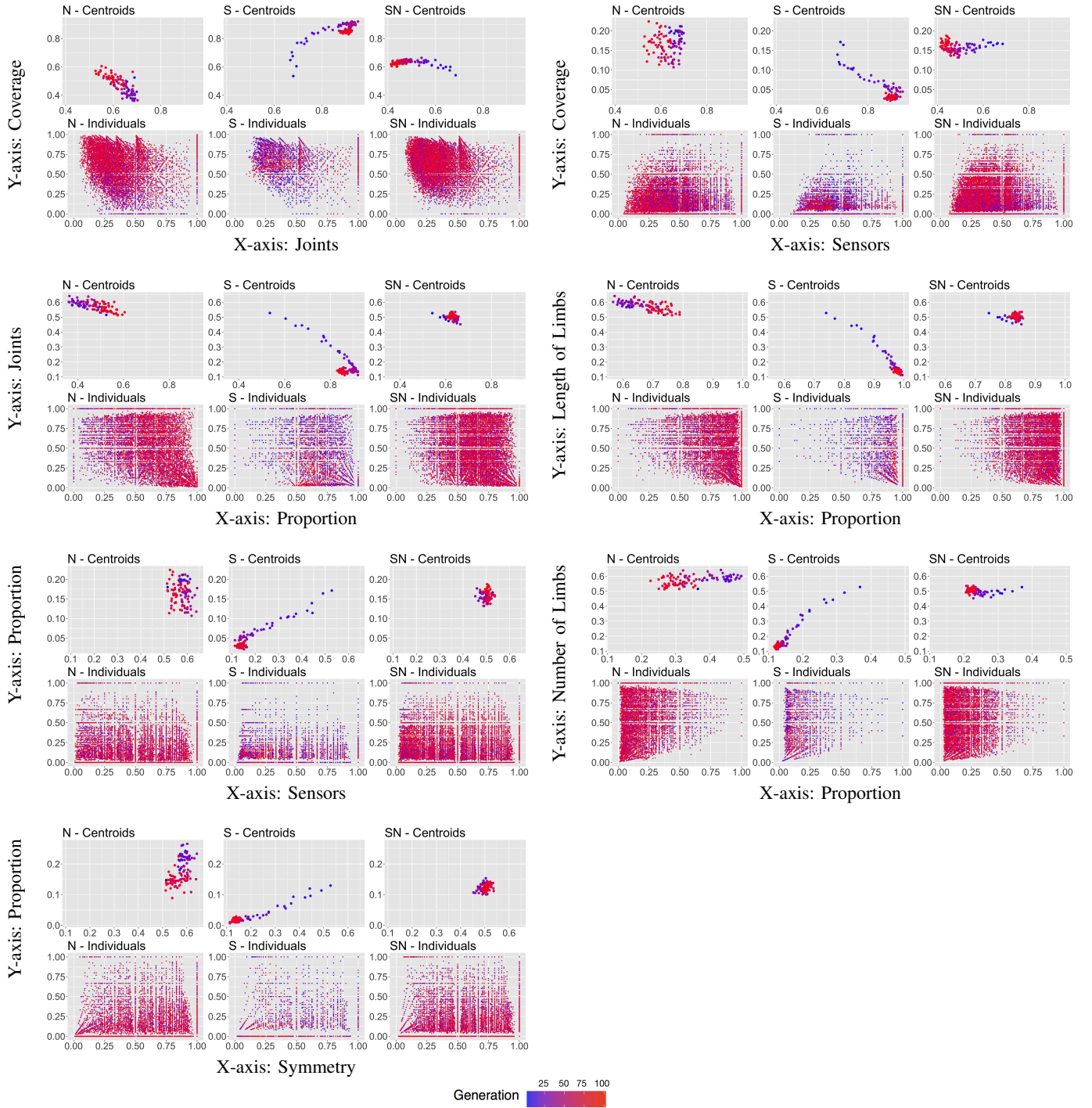


Fig. 9: Points in space of pairs of descriptors for which *SN* had a significantly greater NSC than the other fitness functions. N = Novelty, S = Speed, and SN = Speed & Novelty.

- [13] G. S. Hornby and J. B. Pollack, "Body-brain co-evolution using l-systems as a generative encoding," in *Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation*. Morgan Kaufmann Publishers, 2001, pp. 868–875.
- [14] C. C. Aggarwal, A. Hinneburg, and D. A. Keim, "On the surprising behavior of distance metrics in high dimensional space," in *International conference on database theory*. Springer, 2001, pp. 420–434.