



Understanding fitness landscapes in morpho-evolution via local optima networks

Sarah L. Thomson
Edinburgh Napier University
Edinburgh, Scotland, UK
s.thomson4@napier.ac.uk

Emma Hart
Edinburgh Napier University
Edinburgh, Scotland, UK
e.hart@napier.ac.uk

Léni Le Goff
Edinburgh Napier University
Edinburgh, Scotland, UK
l.legoff2@napier.ac.uk

Edgar Buchanan
University of York
York, England, UK
edgar.buchanan@york.ac.uk

ABSTRACT

Morpho-Evolution (ME) refers to the simultaneous optimisation of a robot's design *and* controller to maximise performance given a task and environment. Many genetic encodings have been proposed which are capable of representing design and control. Previous research has provided empirical comparisons between encodings in terms of their performance with respect to an objective function and the diversity of designs that are evaluated, however there has been no attempt to *explain* the observed findings. We address this by applying Local Optima Network (LON) analysis to investigate the structure of the fitness landscapes induced by three different encodings when evolving a robot for a locomotion task, shedding new light on the ease by which different fitness landscapes can be traversed by a search process. This is the first time LON analysis has been applied in the field of ME despite its popularity in combinatorial optimisation domains; the findings will facilitate design of new algorithms or operators that are customised to ME landscapes in the future.

KEYWORDS

fitness landscape analysis; evolutionary robotics; local optima networks; indirect representation

ACM Reference Format:

Sarah L. Thomson, Léni Le Goff, Emma Hart, and Edgar Buchanan. 2024. Understanding fitness landscapes in morpho-evolution via local optima networks. In *Genetic and Evolutionary Computation Conference (GECCO '24)*, July 14–18, 2024, Melbourne, VIC, Australia. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3638529.3654059>

1 INTRODUCTION

Ever since the pioneering work of Karl Sims almost 30 years ago [40], evolutionary approaches to simultaneously optimise both the designs (i.e. the body) and controllers of robots have gathered pace. We term this morpho-evolution (ME) in this article. A variety of

approaches now exist ranging from those that apply evolution purely in simulation [19], through hybrid approaches that combine hardware and software evolution [11] to those that evolve purely in hardware [5]. Evolutionary frameworks that facilitate ME must consider the challenging question of how to represent both designs and controllers in a form that can be evolved. Direct encodings such as directed trees [49] provide a one-to-one mapping between a genotype encoding the design and control that is realised as the phenotype, and have been postulated to facilitate the fine-tuning of parameters [50]. On the other hand, indirect encodings such as Compositional Pattern Producing Networks (CPPNs) [42] and L-Systems [18] result in many-to-one genotype-phenotype mappings which have several potential benefits, e.g. in being able to produce repeating patterns. This characteristic is generally useful in robotics as it can facilitate the creation of symmetrical designs, e.g. with equal numbers of actuators on each side of a body. However, when using an indirect encoding, the genotype and the phenotype spaces are not necessarily isometric: small changes in the genotype can lead to major changes in the phenotype, and vice-versa [1, 12]). Several authors have also noted that the evolutionary progress using an indirect encoding is slower than that of a direct one (e.g. [16]). While there have been several studies that have compared direct and indirect encodings for jointly optimising design and control in terms of the level of performance reached (i.e. objective fitness) and the diversity of designs of evolved robots [8, 26, 48–50], to the best of our knowledge there has been no attempt to *explain* these empirical observations. That is, there has been no attempt to study the fitness landscapes induced by a choice of encoding to shed light on the ease by which these landscapes can be traversed during a search process. Understanding these landscapes could help explain empirical observations regarding performance or diversity, but could also facilitate the design of new algorithms or operators that are customised to the type of landscapes found.

The novel contribution of this paper is to apply *Local Optima Network* (LON) analysis [36] to analyse the structure of the fitness landscapes created by different encodings when jointly optimising the body and control of a simulated robot. LONs are a technique first introduced in 2008 [32] which enables the information contained in a search space to be compressed into a single graph. Analysis of the graph enables the calculation of a set of metrics which characterise the structure of a landscape. The technique has proved very popular in the combinatorial optimisation community [28, 30, 33], and more



This work is licensed under a Creative Commons Attribution International 4.0 License. *GECCO '24, July 14–18, 2024, Melbourne, VIC, Australia*
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0494-9/24/07.
<https://doi.org/10.1145/3638529.3654059>

recently with continuous functions [27] but has not been used in evolutionary robotics. We pose the question *"To what extent can applying LON analysis to fitness landscapes induced by different encodings in morpho-evolution provide an explanation for (1) observed performance and (2) the extent to which the search-space is explored"*. To answer this, we apply LON analysis to landscapes induced by three different encodings (a direct encoding producing a tree, L-System, CPPN) for an ME locomotion task, shedding new light on the effect of solution encoding on search-space *navigability*. We define navigability as the ability for the search to a) consistently escape local optima, and b) explore a large number of genotypes. This is the first time that LON analysis has been applied within Evolutionary Robotics (ER). Furthermore, it is also novel in applying LON analysis to representations in which there is no direct equivalence between genotype and phenotype: there is only one other attempt at this that we are aware of in which LON analysis is applied to Linear Genetic Programming [13]. The results indicate that L-System indirect encoding leads to better fitness landscape navigability; this helps to explain its superior performance in terms of fitness in some empirical studies [49, 50]. In contrast, the CPPN encoding is shown to induce lower navigability — in fact, it seems to have landscapes which exhibit large numbers of low-quality local optima which are difficult to escape.

2 RELATED WORK

Recent years have seen a significant increase in efforts focused on the joint optimisation of both design and control of robots using evolutionary methods [2, 5, 11, 17, 19, 24]. Much of this work is inspired by the seminal work of Pfeifer & Bongard [38] ‘How the Body Shapes the Way We Think’ which argues that thought is not independent of the body but is tightly constrained, and at the same time enabled, by it — suggesting the evolution of design and control cannot be considered independently. Several evolutionary frameworks have been proposed to simultaneously evolve design and control [10, 11, 17, 24, 26]. Regardless of the choice of framework, a decision must be made regarding how to encode both design and control on a single genome. A number of encodings have been suggested, which can be categorised as *direct* or *indirect*. In the former, typical examples include the use of directed trees [2, 48, 49] which for example specify how modules are connected to create a design, as well as the control parameters of each module. Indirect encodings require a mapping from genotype to phenotype. The use of CPPN [6, 15, 42] is common here: a particular advantage of this type of encoding is that it can produce repeating patterns, a characteristic which is often beneficial in producing symmetric robots which facilitate locomotion. An alternative indirect encoding is an L-System [18] in which an evolved grammar produces rules that specify the robot [23, 25]. A number of studies have examined the effect of the choice of encoding on performance (i.e. objective fitness) in joint optimisation of design and control [48, 50]. In [49], the effect of the encoding on solution diversity is also examined. Miras *et al.* [22] compare two generative encodings in terms of how the choice of encoding influences the phenotype of evolved robots. In [39], the extent to which the encoding influences learning ability in joint optimisation. However, there have been very few attempts to relate observed empirical performance metrics to the

underlying fitness landscape induced by the choice of encoding, particularly in morpho-evolution. Some early literature in fitness landscape analysis in ER focused solely on the controller of the robot [9]. In [41] a simple form of morpho-evolution was investigated in which the sensor placement of a fixed robot body is evolved along with the controller. The authors investigate ruggedness and neutrality within this landscape. Naya-Varela *et al.* [29] propose an evolutionary algorithm for *growth-based morphological development* in which the robot structure ‘grows’ overtime following a fixed developmental schedule. In this work, the authors employed a graph-based tool called Search Trajectory Networks (STN) [31] which enables the trajectory of the best solutions found by a population-based algorithm to be visualised over time. The STN reveals information about how the algorithm traverses the underlying fitness landscape. However, this algorithm does not fall in the class of ‘morpho-evolution’ algorithms considered in this article as the design changes according to a pre-determined schedule, rather than being evolved. Outside of ER, LONs have been used to bring new insights into the optimisation dynamics of a wide variety of different problems e.g. in combinatorial optimisation [32, 37], neural architecture search (NAS) [34], in algorithm-selection [3] and in continuous landscapes [36]. All of the aforementioned examples apply LON analysis to encodings in which there is a one-to-one mapping between genotype and phenotype. The only example of which we are aware in which LON analysis is applied to a scenario in which there is a many-to-one mapping is a recent analysis of Linear Genetic Programming [13]. In summary, this article adds to the field of ER by using LONs for the first time to understand fitness landscapes, in the context of jointly optimising design and control using different encodings.

3 PRELIMINARIES

This section provides a brief primer on LONs and defines the terms that are used throughout the remainder of the article in this respect, given that this is the first time they have been used in the context of Evolutionary Robotics. LONs [32] are a means to study the global structure of a fitness landscape. In a LON, nodes represent local optima in a fitness landscape, and edges between nodes are search transitions between them. By *search transition*, we mean that two locations in the configuration space were linked during a metaheuristic search: in our study, a search transition between nodes is an application of perturbation followed by hill climbing. It is therefore a sequence of search operations applied to the source node which results in the destination node. LONs can be created by sampling local optima in a fitness landscape using Iterated Local Search (ILS) [20]. We now define some key terminology:

Neighbourhood. The *neighbourhood* of a solution, s_i , are the solutions which are adjacent to s_i according to a neighbourhood function: $N(s)$. In this work, the notion of adjacency depends on the design encoding type: the particulars of these are detailed in Section 4.4.1.

LON nodes. A local optimum has superior or equal fitness to its neighbours according to a fitness function f . In this work, we do not exhaustively search the neighbourhood: this would be computationally infeasible. Instead, we consider that a solution l_o is a local

optimum if it has superior or equal fitness to its *sampled* neighbourhood SN . Formally: $\forall n \in SN(lo_i) : f(lo_i) \geq f(n)$ (assuming maximisation, as is the case for this study) where $SN(lo_i)$ is the sampled neighbourhood, n is a particular neighbour. The nodes in a LON, LO , are the local optima as just defined.

LON edges. There is an edge from local optimum lo_i to local optimum lo_j , if lo_j can be obtained after applying a random perturbation to lo_i followed by local search, and $f(lo_j) \geq f(lo_i)$. In LON terminology, these are called *escape* edges [52]. The edges are called *monotonic* because they record only non-deteriorating, directed connections between local optima. Edges are weighted with the frequency of transition: the number of times during searches that lo_j was reached by applying perturbation then local search to lo_i . The set of edges is denoted by E .

Local optima network (LON). A local optima network, $LON = (LO, E)$, consists of nodes $lo_i \in LO$ which are the local optima, and edges $e_{ij} \in E$ between pairs of nodes lo_i and lo_j with weight w_{ij} iff $w_{ij} > 0$. A LON which only includes neutral or improving transitions between local optima is a *monotonic* LON, or MLON; we construct these for the present work. For the interested reader, detailed descriptions of MLONs can be found in previous literature [35].

4 METHODOLOGY

4.1 Task and Environment

We conduct LON analysis on a joint optimisation task taken from Veenstra *et al.* [50] which focused on empirical performance comparisons of robots evolved using different encodings. The goal is to evolve a robot that needs to traverse a virtual landscape consisting of a set of gentle hills. The objective is to maximise the distance travelled. We use an environment called gym2D which facilitates robot design evolution for OpenAI Gym [4]; its foundation was the bipedal walker environment¹. The beginning section of the environment we use is depicted in Figure 1. A virtual robot begins at the left of a rectangular box containing a virtual 2D landscape and the objective is to move as far to the right of the box as possible; the horizontal axis for the box is associated with a scale between 0 and 100 (0 is no movement; 100 is reaching the end) and the fitness of a robot is how far along this axis it travels. There is also a "kill-switch" implemented to reduce the computational cost: if a robot travels slower than a minimum speed (set at 0.04) then the attempt terminates and the robot is ascribed a default fitness of 5.0. The same environment configuration is used across all the experiments. A robot can be seen attempting to travel across the landscape in Figure 1.

4.2 Robots

A *robot* is defined by a set of *modules* combined to produce a 2D shape. Modules are either circular or rectangular and each has parameters defining its size. The allowable size ranges for modules are as follows: a width and height between 0.5 and 1, and (in the case of circle shapes) a radius between 0.25 and 0.5. Modules also each have a controller, as described shortly. A list of eight modules

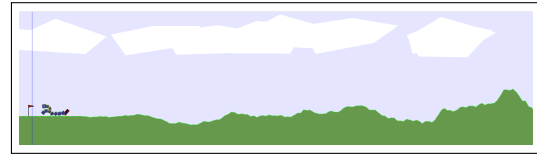


Figure 1: An example robot attempting to move horizontally through the virtual environment. Each shape is a module, and each module has its own controller.

is kept in memory and these are selected to be used as nodes to build the robot. For each individual run, a new initial *module list* is generated of four circles and four rectangles. The radius of the circles is randomly generated between the bounds just mentioned; similarly, the width and height of the rectangles are randomised. The parameters of the modules are mutated during the evolutionary runs. Each virtual robot has one controller per module. A controller is encoded as a sine wave with parameters: amplitude (α , frequency θ , phase δ , and offset ϵ ,

$$y(t) = \alpha \sin(\theta t + \delta) + \epsilon \quad (1)$$

The ranges for the parameters are: α and δ can vary between -1 and +1; θ between -0.1 and +0.1; and ϵ between $-\pi$ to $+\pi$. These parameters are also subject to mutation (details of mutation are found later in Section 4.4.1. Finally, the maximum depth of the tree which represents a robot is seven, while the maximum size (number of allowable modules) is 40. These parameters are taken directly from the literature [50].

4.3 Encodings

The phenotype of every virtual creature in Gym2D is represented as a tree (see the left sub-plot of Figure 2). We evaluate three different encodings to produce this tree. The first encoding is direct, therefore exactly represents the tree. Two encodings are indirect therefore there is a mapping from genotype to phenotype. The same three encodings were used in [49]. For the indirect encoding, the tree starts from an axiom and then expand iteratively until reaching a maximum depth. Each node can have a maximum of three nodes attached. The following section describes how each encoding generates a tree.

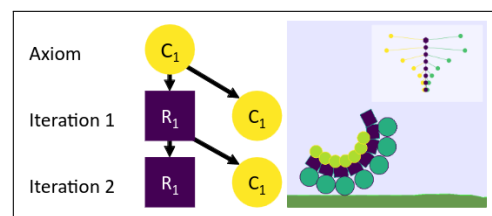


Figure 2: On the left: directed tree created by an L-System encoding after two iterations. The nodes labelled C_1 and R_1 show a circle and rectangle module, respectively. On the right: an example of a virtual 2D robot; its graph is visualised in the top right. Figure is from [50]

¹https://github.com/openai/gym/blob/master/gym/envs/box2d/bipedal_walker.py

Direct encoding. The direct robot design encoding is a directed tree structure which represents the topology of the robot’s design. For each node that makes up a robot, the encoding describes: the node’s index in the tree; the index of the module from the module list it represents; the index of its parent node; and the parent connection coordinates.

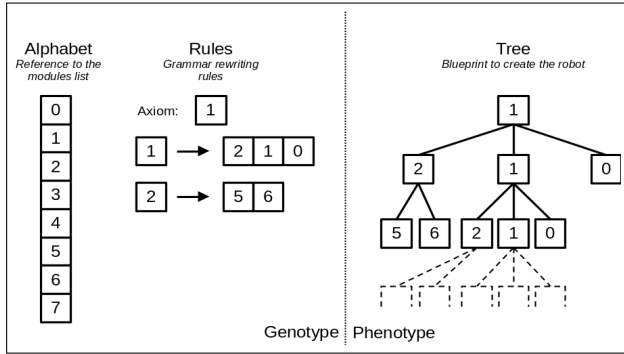


Figure 3: Diagram representing how an L-system works

L-System. For the L-System encoding, a parametric approach is used where symbols and grammar rules are used (see figure 3). A set of symbols is specified via an alphabet. The alphabet is composed of eight symbols, each of which corresponds to a module in the modules list. A set of rules, known as *grammar rewrite rules*, are applied to the symbols in order to construct the tree. A rule is composed of an input symbol and substitution symbols, traditionally, the input symbols are replaced by the substitution symbols. In this system, starting from an axiom symbol, the rules are used to add new leaves (each leaf corresponds to one of the substitution symbols) to the tree from a previously created node (the input symbol).

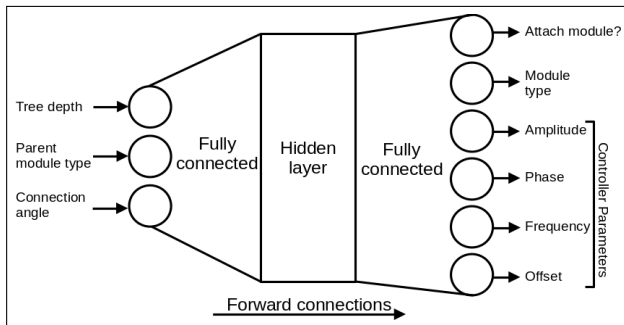


Figure 4: Diagram representing a CPPN network

CPPN. CPPNs are neural networks specifically designed to produce patterns where the inputs are spatial coordinates. Each network node is a function representing a basic pattern such as Gaussian for axial symmetry, sinusoidal for repetition, or sigmoid for central symmetry. As CPPN is similar to a feed-forward neural network it is classically trained using neuro-evolution. In this study, we make use of the Neuro-evolution through Augmenting Topologies

(NEAT) [43]. To create the tree, the CPPN is queried incrementally for each connection site, i.e. where a node can be potentially attached. The CPPN takes three inputs: the depth of the connection site in the tree, the parent index in the module list and the angle of connection. Thus, the CPPN produces 6 outputs: whether a module should be attached or not, the type of module, i.e. the index in the modules list, and four outputs corresponding to the controller parameters (see figure 4).

4.4 LON Construction

Following the convention in the LON literature (see, for example: [37, 46, 47, 51]), we sample local optima by executing runs of Iterated Local Search (ILS). ILS [20] is a single-point meta-heuristic which conducts repeated cycles of applying large perturbations followed by local search. Although algorithms with recombination are typical in evolutionary robotics and neuro-evolution, here we use mutation only – following the majority of work in LONs – but return to the question of whether crossover should be used later. At each iteration during the LON construction process, a local optimum is obtained; an acceptance condition then dictates whether it is accepted as the new incumbent solution, or whether the previous local optimum should be obtained. In this way, a process reminiscent of a Markov chain is followed. We gather data from 30 runs of ILS, each commenced from random starting solutions during the LON construction process. ILS combines local search with random perturbations. The local search employs a first-improvement pivot rule. This means that whenever an improving (or equal) neighbour is found, it immediately becomes the incumbent genotype [44]; this is in contrast to best improvement, where the best possible neighbour is found. This would require too much computation here. In the experiments, the local search process stops when there has been 100 operations applied with no improvement. Each of the runs terminates after 30 iterations with no improvement in local optimum quality or after 100 iterations in total. The local search conducts random mutation on both the robot controller and the robot design according to pre-defined probabilities described in the next section. During the runs, every accepted local optimum is hashed and logged as a network node, and any transition between local optima is logged as network edges. When a local optimum is added, its fitness, an index, and its hashed phenotype are noted too, in addition to the hashed genotype. Edges are recorded as the combination of the source node index and the destination node index, and weight is logged as well. Edge weights are the number of times a transition was followed. The LON is an amalgamation of the nodes and edges from the separate runs.

4.4.1 Mutation operators. We use the mutation settings which were tuned in previous literature [49]; these are provided in Table 1. Perturbation is the mutation operation applied three times. Only improving or equal local optima are accepted during the search. As mentioned, both the controllers and the designs are mutated and each has a separate application rate (see Table 1). For the design, the approach to mutation depends on the encoding type:

Direct encoding. Mutation on the direct robot design encoding has three possible operations: removing a node, adding a node to the tree structure, and modifying a module’s node. For each

possible location for adding a node, a node is added according to the mutation rate; for removals, the mutation rate is halved (owing to the fact that sub-nodes are removed alongside the original node; that is, it is a more dramatic mutation). Finally, the module’s nodes are modified by mutating the shape (width, height, radius) and connection angle using Gaussian mutation.

L-System. Mutation on the L-System robot design encoding takes place in the following way: the rules are changed by adding or removing substitution symbols and the shape (width, height, radius) and angle of the modules in the list are mutated using Gaussian mutation.

CPPN. For the CPPN encoding, the mutation is the default NEAT mutation from the NEAT-PYTHON² library. This has several operations modifying the network structure: adding a connection; deleting a connection; adding a node; deleting a node; and replacing bias. All of these are carried out according to the robot design mutation rate. There are a large number of parameters associated with the CPPN encoding. We use the values provided in a configuration file³ and associated with previous literature [50]. As well as the CPPN mutation, the shape (width, height, radius) and angle of the modules in the list are mutated according to the mutation probability.

Controller mutation. For all three encodings, the same controller representation is used (recall Equation 4.2). For the direct and L-System encodings, controllers’ parameters are mutated according to a mutation rate (controller rate in Table 1) using Gaussian mutation. The standard deviation of the Gaussian distribution is set at 0.2 (the same as in [49]). For the CPPN, the process is different: in that case, each of the four parameters which define a controller is an output from the network (as described in Section 4.3 and with Figure 4); they are therefore affected when the NEAT mutation is carried out on the CPPN.

Table 1: Mutation rates in the LON construction algorithm

parameter	direct	L-System	CPPN
rate (controller)	0.32	0.16	0.02
rate (design)	0.16	0.04	0.02

4.5 Experimental Setup

The majority of our experimental setup matches that of Veenstra *et al.*, who compared morphological encodings [49, 50]. We begin from the associated online repository⁴ as a base for our implementation; from now on, we will refer to this as *veenstra-repository*. All parameters associated with the encodings and environment are from there. For each of the three encodings, we conducted 30 independent ILS runs to log local optima and their edges; these were amalgamated to form a single LON object for each encoding.

Statistics are calculated using the Mann-Whitney u-test [21] to reject the hypothesis that two distributions are statistically the same

²<https://neat-python.readthedocs.io/en/latest/>

³https://github.com/FrankVeenstra/gym_rem2D/blob/master/ModularER_2D/NeuralNetwork/config

⁴https://github.com/FrankVeenstra/gym_rem2D

where * means $p < 0.05$, ** means $p < 0.005$, *** means $p < 0.0005$ and **** means $p < 0.00005$ and p represent the probability.

Analysis of the LONs includes considering their weakly connected components. In directed networks, a *weakly connected component* (WCC) is a group of nodes which are all reachable from each other, even if edges are uni-directional or in different directions; WCCs are identified in the analysis using Tarjan’s algorithm [45] and are referred to as *components* from now on.

5 RESULTS

5.1 Visualisation

Figures 5a-5d display visualisations of the extracted LONs; there is a separate Figure for each of the three robot design encoding types. In the plots, larger size and darker colour of nodes indicate better fitness. Node size is proportional to fitness. For colour, we took the approach of separating nodes into three fitness levels according to the *overall distribution of fitness* across all three LONs: nodes with fitness in the first quartile (below 9.59) are low-fitness and are coloured very pale purple. Nodes within the interquartile range are middle-fitness and are light purple in colour. Finally, nodes with fitness in the upper quartile — which is 24.48 or above — are high-fitness and are dark purple. Self-loops are represented by edges which curve out of and back into the left of a node. The force-directed network layout algorithm *kamada kawai* [14] in R has been used to layout the LONs. Let us first consider the direct encoding LON shown in Figures 5a (with self-loops) and 5b (without self-loops). It is immediately evident from the latter that this is a network containing several isolated components; in fact, there are 30, equal to the number of ILS runs used to construct the LON. This is not a typical phenomenon seen in other domains where LON analysis has been used, e.g. [7, 27, 47] which usually result in more densely connected graphs. This is likely due to the vastness of the search space. In evolutionary robotics, the search space can be essentially infinite (most controllers are encoded as real-valued numbers) and the goal is not necessarily to find a global optimum. Instead, algorithms are typically applied to find a *good enough* robot that functions well.

In Figure 5b we can see that the 30 components are chains — each node has either zero or one outgoing directed edges, and either zero or one incoming directed edges. We note from Figure 5a that the direct encoding LON exhibits several self-loops: these are recorded when an escape attempt from a local optimum is made, but returns to the same local optimum. Self-loops are not present for the LONs of the other two encodings, indicating that it may sometimes be difficult to escape local optima when using the direct encoding (i.e., there are strong attractors present). Additionally, the chains of the direct encoding (Figure 5b) seem to be shorter when compared to the other two encodings. This indicates a lack of *navigability* associated with this encoding: it may struggle to explore genotypes (indeed, we can see from Table 2 that this encoding resulted in far fewer unique designs being considered than the L-System). Despite this, we notice from Figure 5b that the direct encoding chains often terminate with a good-fitness (dark purple) local optimum. Looking across Figures 5a-5d, the divergence in fitness distributions can be seen. The CPPN LON in Figure 5d contains a lot of nodes which are very pale purple, indicating their fitness is in the first quartile

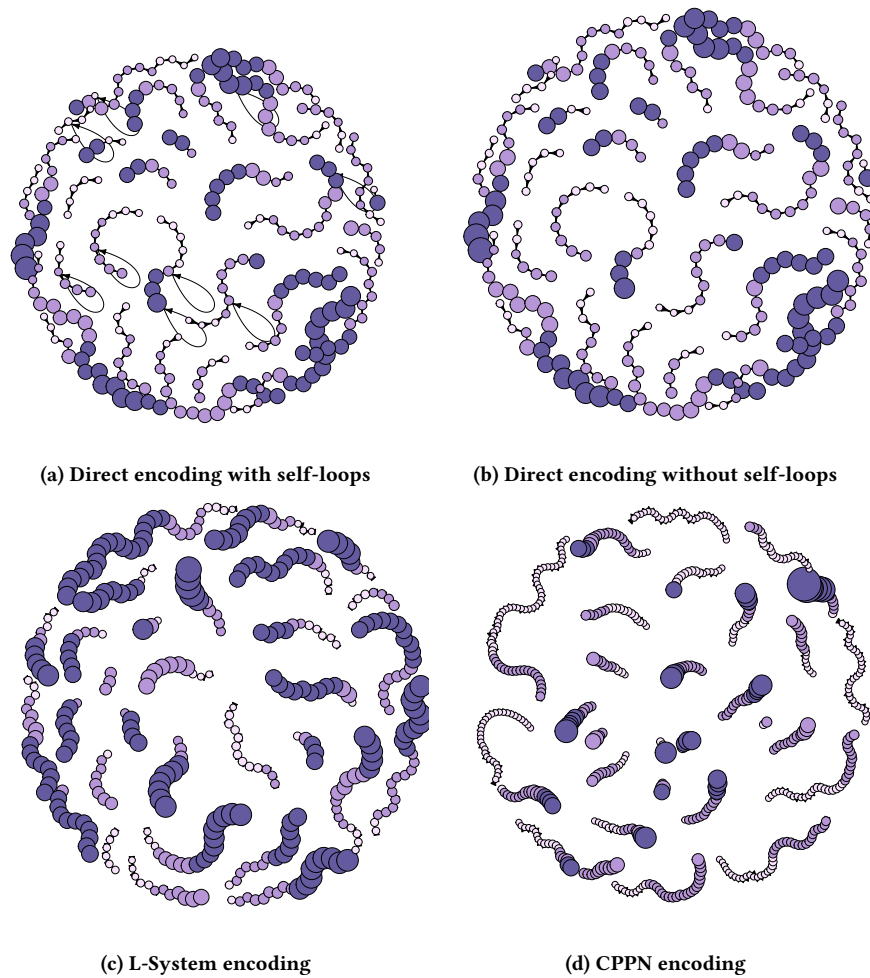


Figure 5: LONs for the three robot design encodings. larger size and darker colour of nodes indicate better fitness. Node size is proportional to fitness. For colour: nodes with fitness in the first quartile of the overall fitness distribution across the three LONs (this is a fitness of less than 9.59) are low-fitness and are coloured very pale purple. Nodes within the interquartile range are middle-fitness and are light purple in colour. Finally, nodes with fitness in the upper quartile – which is 24.48 or above – are high-fitness and are dark purple. Self-loops are represented with edges which curve out of and back into the right of a node.

of the three-LON fitness distribution. Comparing that observation with Figure 5c, where the LON for the L-System encoding is shown, we note that there are more chains which contain the dark purple; that is, high-fitness (in the upper quartile) nodes are reached in more searches. The chains for this LON appear to be longer than those evident in the direct encoding LON (Figure 5a), and they very often end up with fitness in the third quartile (above 24.48). These observations, when taken together, indicate the *navigability* associated with an L-System encoding.

Another observation from the CPPN LON in Figure 5d is that some of the chains are lengthy but do not display any notable improvement in fitness. This tells us that the fitness landscape associated with using this encoding is not easily navigable: there appear to be large poor-quality (where poor-quality is having a fitness in the first quartile: less than 9.59) local optima plateaus which are difficult to escape (under the studied search operations

and parameters). Looking at the direct and L-System encodings in Figures 5a and 5c, and remark that they appear to be associated with better fitness landscape navigability than the CPPN: the chains in their LONs more often display incremental improvements in fitness, which can be seen in the increase in size of node throughout a chain. This is particularly evident with the L-System LON in Figure 5c. It seems that the search becomes trapped at low-quality local optima with these encoding less often. This helps to explain the higher performance which has been observed using L-System design encoding in the literature [50].

5.2 Features

In this Section we will first consider statistics about the ILS runs conducted with the different encodings. Table 2 shows the rate of mutation acceptance; the rate of design acceptance (this is computed as the number of accepted mutations which were associated

with a changed design); the number of unique designs present; and the total number of attempted mutations. From Table 2, we notice that the mutation acceptance is lowest for direct encoding and the CPPN encoding has the lowest design acceptance. The highest rates are associated with the L-System encoding. The other two encodings tried out substantially lower numbers of unique designs than the L-System, despite having a larger number of attempted mutations. These facts, taken together, help to ratify the observations in the previous section that the direct encoding struggles to widely explore and find fitness improvement, while the L-System encoding facilitates wide exploration and moving towards better fitness.

Table 2: ILS run statistics for the three encodings

metric	direct	L-System	CPPN
<i>mutation acceptance</i>	10.52%	47.65%	33.77%
<i>design acceptance</i>	32.15%	38.10%	19.38%
<i>unique designs</i>	165 800	227 411	147 443
<i>attempted mutations</i>	352 438	297 964	328 970

Table 3 reports measurements for the LONs associated with each encoding type. We introduce now the metrics whose names or meanings may not be intuitive: *components* is the number of weakly connected components in the network. The metric *path length* is the average edge distance between any two nodes in the network (excluding nodes which are unreachable to each other); *degree* is the average degree of a node, including self-loops; and *infeasible* is the percentage of nodes which had fitness indicating that the robot activated the killswitch and can therefore be considered infeasible.

Table 3: LON statistics

metric	direct	L-System	CPPN
<i>nodes</i>	269	360	546
<i>edges</i>	248	330	516
<i>components</i>	30	30	30
<i>path length</i>	4.37	6.32	12.66
<i>degree</i>	1.84	1.83	1.89
<i>infeasible</i>	0%	0.003%	34.43%

The plots in Figure 6 show distributions for run and LON metrics across the three encoding types. For each, there is a boxplot with a horizontal line indicating the median. We also overlay as a swarm the actual data points. For each pair of distributions, the u-test result is shown. Figure 6a displays the distribution of all fitnesses in the sampled LON nodes (local optima) of each of the three robot design encoding choices: direct, L-System, and CPPN.

We can notice from visual inspection that the three distributions have very different medians, and the swarms show that the overall shape of the distributions are divergent to one another as well. In terms of IQR and median, the L-System has the best fitness range, with the direct encoding being next-best. This was also shown in a previous study [50]. The CPPN encoding is especially distinct from the other two; when including outliers, it has a much wider range, and includes more lower-quality fitnesses than those of the other two. Another interesting phenomenon is that in the case of L-System and CPPN, there appear to be groups of local optima with the same fitness: at approximately 25 for the L-System and direct

encodings, and approximately 10 for the CPPN. Despite having the lowest IQR (a finding which ratifies those in the literature [49]), the CPPN encoding resulted in the highest fitness sampled. Figure 6b shows distributions for the *maximum* fitness found in each run. We can see from visual inspection that the CPPN is associated with the widest range. Indeed, it has both the highest value and the lowest. The encoding which shows the highest (best) IQR and median is the L-System type. The plot in Figure 6c shows, for each of the three encodings, the distribution for the number of fitness evaluations used by the 30 runs. Notice from comparing the boxes visually that the distributions are similar for all three encodings.

In Figure 6d we present the data concerning fitness deltas. A delta is here defined as the fitness increase associated with a directed LON edge (recall that the edges are always directed towards equal or better fitness, so the delta is always zero or above). For the plot, we compute the *median* delta within each component in the LON; these correspond to the 30 runs. Those medians are the data shown in the Figure. Observing the three encodings, we note that the direct encoding is associated with the largest fitness deltas. This can be seen with its higher IQR and median line. Additionally, the largest value in across all encodings is achieved by the direct encoding. The L-System distribution has the lowest and narrowest IQR, implying that the fitness deltas are the smallest (but perhaps most consistent in magnitude) of the three. In terms of significant difference, the L-System and CPPN pair are statistically the same; however, the direct/L-System and direct-CPPN pairs differ, with $p < 0.005$ and $p < 0.05$ respectively. Finally, the distributions of chain lengths in the LON (there is one for each of the 30 runs) are captured in Figure 6e. A chain is the path of local optima followed by one run of the ILS. We observe from the plot that the direct encodings had the shortest chains; this was followed by the L-System, and then CPPN with the longest chains. The direct-L-system pair has an associated $p < 0.05$, the L-system-CPPN pair are $p < 0.05$ different and the direct-CPPN are statistically different with $p < 0.0005$.

5.3 The Elephants in the Room

One consideration of our approach is that we did not use a crossover operator. The reader may wonder whether the ILS used here actually explores promising regions of the search space, or whether instead all the sampled local optima are poor when compared to those that might be found by crossover-based approaches. To this end, we note that the fitness levels reached as shown in Figure 6b are similar to those shown in previous literature — where an evolutionary algorithm with crossover was used [50] with the same virtual environment and both direct and L-System encoding. It is therefore reasonable to assume that the ILS approach reaches fitness levels approximately equivalent to those achieved by a crossover-based algorithm, and that at least some of the local optima sampled are of acceptable quality. Another factor to consider when interpreting the results is that we used mutation parameters from the literature ([49]). Those were tuned for an algorithm which also included crossover, so it is probable that they are not the optimal settings when applied in an iterated local search. This being said, being consistent with the parameters from previous work is a reasonable place to start when beginning a new line of research. In this study, we looked at different possible genotype encoding to explore one

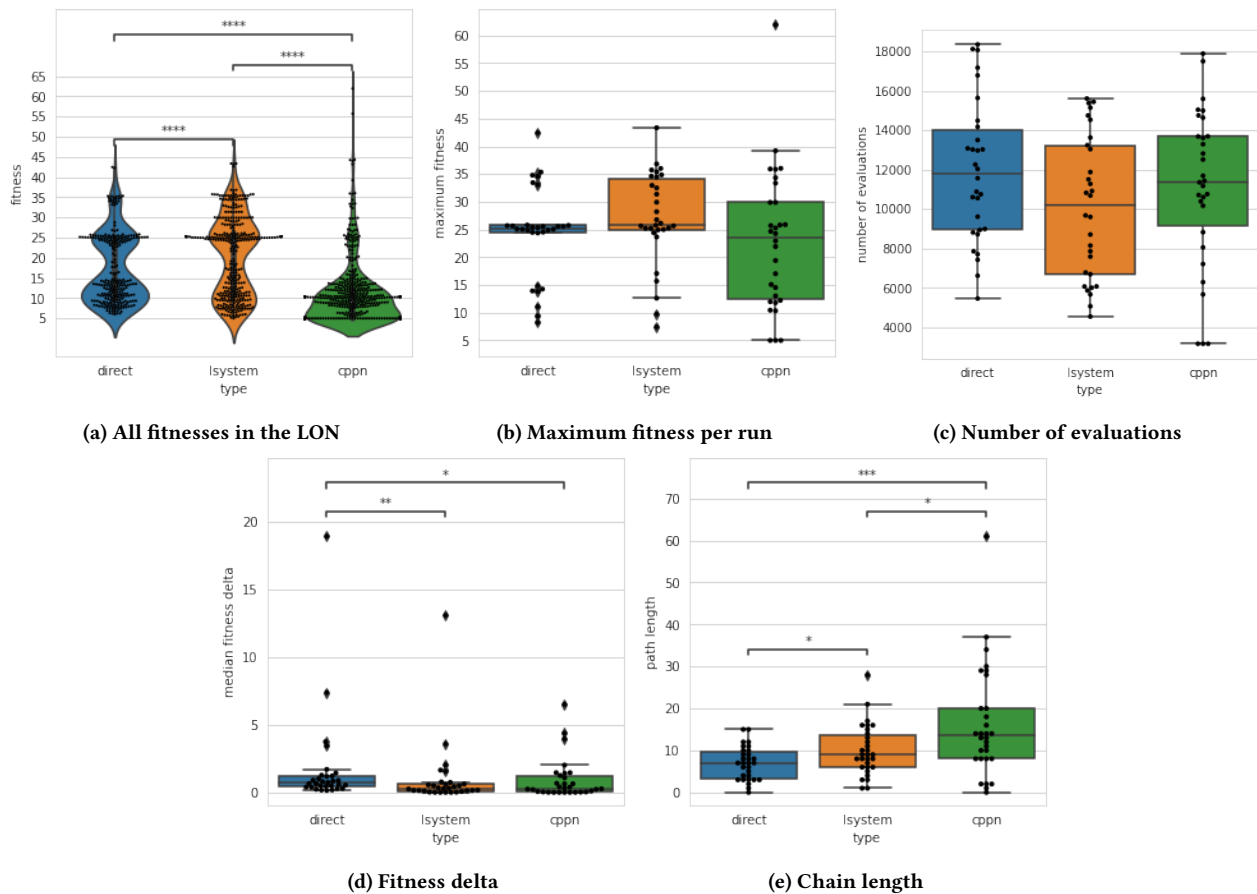


Figure 6: Distribution for various run and LON metrics across the three three encoding types. If there is a significant difference (according to a u-test) between two distributions, this is indicated with a horizontal bar and p -value level: $*p < 0.05$; $p < 0.005$; $***p < 0.0005$; $****p < 0.00005$**

phenotype space: a tree-based structure. Therefore, the present findings are limited to this phenotype space. In future work, it will be interesting to look at how different encodings explore different design representations such as voxel-based.

6 CONCLUSIONS

We have conducted a study with the aim of providing insight into performance differences associated with different robot design encodings which have been observed in the literature. To this end, we carried out fitness landscape analysis with local optima networks (LONs). Three design encodings for a virtual 2D robot were considered: a direct tree structure representation, an L-System, and a type of neural network (a CPPN). We sampled the LONs associated with each of these, visualising them and comparing their features. The results showed that an L-System encoding is associated with good fitness landscape navigability: search has the ability to escape local optima and discover a large number of unique designs. In contrast, although being able to produce a large number of unique designs, the CPPN encoding appears to be linked with low-quality local optima from which it is hard to escape towards better fitness.

This observation indicates that CPPN encoding is not suited for creating the structure used to define a robot morphology in this study: a tree-based structure. We also found that a direct encoding can lead to search struggling to leave local optima; however, good fitnesses can sometimes be found when it does. These findings provide insight into previously-observed performance differences between encodings in the literature. We therefore propose that LONs can be a valuable tool for explaining phenomena seen in observed in evolutionary robotics search can inform the result of better algorithms and operators in future. LONs bring additional benefit compared to simply running evolution under different encodings: they facilitate visualisation of high-dimensional spaces and an understanding of the local optima level in the genotype space. Code and data associated with this work are available in a public Zenodo repository⁵.

REFERENCES

- [1] Mike Angus, Edgar Buchanan, Léni K Le Goff, Emma Hart, Agoston E Eiben, Matteo De Carlo, Alan F Winfield, Matthew F Hale, Robert Woolley, Jon Timmis,

⁵<https://zenodo.org/records/1094345>

- et al. 2023. Practical hardware for evolvable robots. *Frontiers in Robotics and AI* 10 (2023).
- [2] Joshua Auerbach, Deniz Aydin, Andrea Maesani, Przemyslaw Kornatowski, Titus Cieslewski, Grégoire Heitz, Pradeep Fernando, Ilya Loshchilov, Ludovic Daler, and Dario Floreano. 2014. Robogen: Robot generation through artificial evolution. In *Artificial Life Conference Proceedings*. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ..., 136–137.
- [3] Wojciech Bożejko, Andrzej Gnatowski, Teodor Niżyński, Michael Affenzeller, and Andreas Beham. 2018. Local optima networks in solving algorithm selection problem for tsp. In *International Conference on Dependability and Complex Systems*. Springer, 83–93.
- [4] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. 2016. Openai gym. *arXiv preprint arXiv:1606.01540* (2016).
- [5] Luzius Brodbeck, Simon Hauser, and Fumiya Iida. 2015. Morphological evolution of physical robots through model-free phenotype development. *PLoS one* 10, 6 (2015), e0128444.
- [6] Nick Cheney, Robert MacCurdy, Jeff Clune, and Hod Lipson. 2014. Unshackling evolution: evolving soft robots with multiple materials and a powerful generative encoding. *ACM SIGEVOlution* 7, 1 (2014), 11–23.
- [7] Christopher W Cleghorn and Gabriela Ochoa. 2021. Understanding parameter spaces using local optima networks: a case study on particle swarm optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. 1657–1664.
- [8] Matteo De Carlo, Eliseo Ferrante, and AE Eiben. 2020. Comparing indirect encodings by evolutionary attractor analysis in the trait space of modular robots. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*. 73–74.
- [9] Peter Dittrich, Andre Skusa, Wolfgang Banzhaf, and Wolfgang Kantschik. 1999. Dynamical properties of the fitness landscape of a GP Controlled random morphology robot. In *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation-Volume 2*. Citeseer, 1002–1008.
- [10] Agrim Gupta, Silvio Savarese, Surya Ganguli, and Li Fei-Fei. 2021. Embodied intelligence via learning and evolution. *Nature communications* 12, 1 (2021), 5721.
- [11] Matthew F Hale, Edgar Buchanan, Alan F Winfield, Jon Timmis, Emma Hart, Agoston E Eiben, Mike Angus, Frank Veenstra, Wei Li, Robert Woolley, et al. 2019. The are robot fabricator: How to (re) produce robots that can evolve in the real world. In *Artificial Life Conference Proceedings*. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ..., 95–102.
- [12] Emma Hart and Léni K Le Goff. 2022. Artificial evolution of robot bodies and control: on the interaction between evolution, learning and culture. *Philosophical Transactions of the Royal Society B* 377, 1843 (2022), 20210117.
- [13] Ting Hu, Gabriela Ochoa, and Wolfgang Banzhaf. 2023. Phenotype Search Trajectory Networks for Linear Genetic Programming. In *European Conference on Genetic Programming (Part of EvoStar)*. Springer, 52–67.
- [14] Tomihisa Kamada, Satoru Kawai, et al. 1989. An algorithm for drawing general undirected graphs. *Information processing letters* 31, 1 (1989), 7–15.
- [15] Sam Kriegman, Douglas Blackiston, Michael Levin, and Josh Bongard. 2020. A scalable pipeline for designing reconfigurable organisms. *Proceedings of the National Academy of Sciences* 117, 4 (2020), 1853–1859.
- [16] Léni K Le Goff, Emma Hart, Alexandre Coninx, and Stéphane Doncieux. 2020. On pros and cons of evolving topologies with novelty search. In *Artificial Life Conference Proceedings* 32. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ..., 423–431.
- [17] Wei Li, Edgar Buchanan, Léni K Le Goff, Emma Hart, Matthew F Hale, Matteo De Carlo, Robert Woolley, Alan F Winfield, Jonathan Timmis, and Andy M Tyrrell. 2023. Evaluation of frameworks that combine evolution and learning to design robots in complex morphological spaces. *IEEE Transactions on Evolutionary Computation* (2023).
- [18] Aristid Lindenmayer and Hans Jürgensen. 1992. Grammars of development: discrete-state models for growth, differentiation, and gene expression in modular organisms. In *Lindenmayer systems: Impacts on theoretical computer science, computer graphics, and developmental biology*. Springer, 3–21.
- [19] Hod Lipson, Vytas Sunspirai, Josh Bongard, and Nicholas Cheney. 2016. On the difficulty of co-optimizing morphology and control in evolved virtual creatures. In *Artificial life conference proceedings*. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ..., 226–233.
- [20] Helena R Lourenço, Olivier C Martin, and Thomas Stützle. 2003. Iterated local search. In *Handbook of metaheuristics*. Springer, 320–353.
- [21] Henry B Mann and Donald R Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* (1947), 50–60.
- [22] Karine Miras. 2021. Constrained by design: Influence of genetic encodings on evolved traits of robots. *Frontiers in Robotics and AI* 8 (2021), 672379.
- [23] Karine Miras and Agoston E Eiben. 2019. Effects of environmental conditions on evolved robot morphologies and behavior. In *Proceedings of the Genetic and Evolutionary Computation Conference*. 125–132.
- [24] Karine Miras, Eliseo Ferrante, and Agoston E Eiben. 2020. Environmental influences on evolvable robots. *PLoS one* 15, 5 (2020), e0233848.
- [25] Karine Miras, Evert Haasdijk, Kyrre Glette, and AE Eiben. 2018. Effects of selection preferences on evolved robot morphologies and behaviors. In *Artificial Life Conference Proceedings*. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ..., 224–231.
- [26] Karine Miras, Evert Haasdijk, Kyrre Glette, and Agoston E Eiben. 2018. Search space analysis of evolvable robot morphologies. In *Applications of Evolutionary Computation: 21st International Conference, EvoApplications 2018, Parma, Italy, April 4-6, 2018, Proceedings 21*. Springer, 703–718.
- [27] Paul Mitchell, Gabriela Ochoa, and Romain Chassagne. 2023. Local optima networks of the black box optimisation benchmark functions. In *Proceedings of the Companion Conference on Genetic and Evolutionary Computation*. 2072–2080.
- [28] Sebastián Muñoz-Herrera and Karol Suchan. 2022. Local Optima Network Analysis of Multi-Attribute Vehicle Routing Problems. *Mathematics* 10, 24 (2022), 4644.
- [29] Martín Naya-Varela, Andrés Faiña, and Richard J Duro. 2023. Guiding the Exploration of the Solution Space in Walking Robots Through Growth-Based Morphological Development. In *Proceedings of the Genetic and Evolutionary Computation Conference*. 1230–1238.
- [30] Gabriela Ochoa and Francisco Chicano. 2019. Local optima network analysis for MAX-SAT. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. 1430–1437.
- [31] Gabriela Ochoa, Katherine M Malan, and Christian Blum. 2021. Search trajectory networks: A tool for analysing and visualising the behaviour of metaheuristics. *Applied Soft Computing* 109 (2021), 107492.
- [32] Gabriela Ochoa, Marco Tomassini, Sébastien Vérel, and Christian Darabos. 2008. A study of NK landscapes' basins and local optima networks. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*. 555–562.
- [33] Gabriela Ochoa and Nadarajen Veerapen. 2018. Mapping the global structure of TSP fitness landscapes. *Journal of Heuristics* 24, 3 (2018), 265–294.
- [34] Gabriela Ochoa and Nadarajen Veerapen. 2022. Neural Architecture Search: A Visual Analysis. In *International Conference on Parallel Problem Solving from Nature*. Springer, 603–615.
- [35] Gabriela Ochoa, Nadarajen Veerapen, Fabio Daolio, and Marco Tomassini. 2017. Understanding phase transitions with local optima networks: number partitioning as a case study. In *Evolutionary Computation in Combinatorial Optimization: 17th European Conference, EvoCOP 2017, Amsterdam, The Netherlands, April 19-21, 2017, Proceedings 17*. Springer, 233–248.
- [36] Gabriela Ochoa, Sébastien Verel, Fabio Daolio, and Marco Tomassini. 2014. Local optima networks: A new model of combinatorial fitness landscapes. *Recent advances in the theory and application of fitness landscapes* (2014), 233–262.
- [37] Lucas Marcondes Pavelski, Marie-Éléonore Kessaci, and Myriam Delgado. 2021. Local Optima Network Sampling for Permutation Flowshop. In *2021 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 1131–1138.
- [38] Rolf Pfeifer and Josh Bongard. 2006. *How the body shapes the way we think: a new view of intelligence*. MIT press.
- [39] Federico Pigozzi, Federico Julian Camerota Verdù, and Eric Medvet. 2023. How the Morphology Encoding Influences the Learning Ability in Body-Brain Co-Optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference*. 1045–1054.
- [40] Karl Sims. 1994. Evolving 3D morphology and behavior by competition. *Artificial life* 1, 4 (1994), 353–372.
- [41] Tom Smith, Andrew Philippides, Phil Husbands, and Michael O'Shea. 2002. Neutrality and ruggedness in robot landscapes. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, Vol. 2. IEEE, 1348–1353.
- [42] Kenneth O Stanley. 2007. Compositional pattern producing networks: A novel abstraction of development. *Genetic programming and evolvable machines* 8 (2007), 131–162.
- [43] Kenneth O Stanley and Risto Miikkilainen. 2002. Evolving neural networks through augmenting topologies. *Evolutionary computation* 10, 2 (2002), 99–127.
- [44] Sara Tari and Gabriela Ochoa. 2021. Local search pivoting rules and the landscape global structure. In *Proceedings of the Genetic and Evolutionary Computation Conference*. 278–286.
- [45] Robert Endre Tarjan. 1974. A new algorithm for finding weak components. *Inform. Process. Lett.* 3, 1 (1974), 13–15.
- [46] Sarah L Thomson, Nadarajen Veerapen, Gabriela Ochoa, and Daan van den Berg. 2023. Randomness in local optima network sampling. In *Proceedings of the Companion Conference on Genetic and Evolutionary Computation*. 2099–2107.
- [47] German Treimun-Costa, Elizabeth Montero, Gabriela Ochoa, and Nicolás Rojas-Morales. 2020. Modelling parameter configuration spaces with local optima networks. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*. 751–759.
- [48] Frank Veenstra, Andres Faina, Sebastian Risi, and Kasper Stoy. 2017. Evolution and morphogenesis of simulated modular robots: a comparison between a direct and generative encoding. In *Applications of Evolutionary Computation: 20th European Conference, EvoApplications 2017, Amsterdam, The Netherlands, April 19-21, 2017,*

- Proceedings, Part I 20*. Springer, 870–885.
- [49] Frank Veenstra and Kyrre Glette. 2020. How different encodings affect performance and diversification when evolving the morphology and control of 2D virtual creatures. In *Artificial Life Conference Proceedings 32*. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , 592–601.
- [50] Frank Veenstra, Martin Herring Olsen, and Kyrre Glette. 2022. Effects of encodings and quality-diversity on evolving 2D virtual creatures. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. 164–167.
- [51] Nadarajen Veerapen, Fabio Daolio, and Gabriela Ochoa. 2017. Modelling genetic improvement landscapes with local optima networks. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. 1543–1548.
- [52] Sébastien Verel, Fabio Daolio, Gabriela Ochoa, and Marco Tomassini. 2012. Local optima networks with escape edges. In *Artificial Evolution: 10th International Conference, Evolution Artificielle, EA 2011, Angers, France, October 24-26, 2011, Revised Selected Papers 10*. Springer, 49–60.