Searching for Novelty in Pole Balancing

Chien-Lun (Allen) Huang
Department of Computer Science
University of Cape Town
Cape Town, South Africa
allen@allenhuang.net

Geoff Nitschke
Department of Computer Science
University of Cape Town
Cape Town, South Africa
gnitschke@cs.uct.ac.za

David Shorten
Department of Computer Science
University of Cape Town
Cape Town, South Africa
dpshorten@gmail.com

***Abstract*—*Novelty Search* (NS) has been proposed as an alternative search approach for black-box optimization methods where the fitness function is replaced and only novel solutions are searched for. NS has been demonstrated as advantageous when the fitness landscape is highly deceptive and misdirects the search process towards local optima. In this research we test the efficacy of NS in comparison to a purely objective based approach and a hybrid approach that combines NS and a fitness function in combination with two *behavior characterization* schemes. The task is non-Markovian double-pole balancing. Results indicate that the success of NS strongly depends upon the behavior characterization scheme used, given that NS performed the best under one scheme and relatively poorly under the other scheme.**

Keywords*—*Novelty Search; Neuro-Evolution; Pole Balancing; Deception; Evolutionary Algorithms

I. Introduction

Neuro-Evolution (NE) [6] has been successfully demonstrated in a diverse range of tasks, producing *Artificial Neural Networks* (ANN) controllers that yield pole balancing [13], finless rocket control [12] behaviors, as well as behaviors that solve various evolutionary robotics tasks [33] such as maze navigation and biped locomotion [22]. Currently, most NE controller design methods are objective based, using a fitness function to evaluate solution quality [27, 35, 16]. *Novelty Search* (NS) [22] is a non-objective based approach, where the search for novel solutions replaces a fitness function.

This research investigates the efficacy of NS compared to objective based NE for deriving ANN controllers able to solve the *non-Markovian double-pole balancing* task [11]. A hybrid approach, combining novelty and objective based search to guide the NE process, is also tested. The non-Markovian double-pole balancing task was selected as it is an established machine learning benchmark task that is a surrogate for complex noisy non-linear control tasks such as robot behavior automation [14]. NE has been demonstrated as an effective approach for solving such tasks [26].

NS has also been demonstrated as out-performing, in terms of the number of evaluations required for solution evolution, objective based NE methods in various tasks [19]. Such tasks include maze navigation [21], [30], biped locomotion [22], behavior-morphology co-evolution [20], an aggregation swarm robotics task [9], optimizing specially designed functions [31], and evolving plastic ANNs [34] and operant reward learning controllers in maze environments [38]. The increased task performance of NS in these cases has been attributed to NS functioning akin diversity maintenance techniques (such as *niching* and *crowding* [3, 7, 18, 25, 43]) common in evolutionary search methods [32], as well as its hypothesized benefits in *deceptive* search spaces [22].

A deceptive search space is one which causes a solution to become trapped in an area where it is unlikely to ultimately reach the objective [42]. Deception has been associated with task complexity as well as specific types of search spaces where a fitness function draws candidate solutions closer to an objective but ultimately the objective cannot be reached. Figure 1 presents an example of deception in a maze navigation task. Objective based search, where a fitness function measures how far a solution (behavior) is from the objective, would be deceived into rewarding behaviors that find dead-ends close to the objective [22].

In the context of this research, it is unclear if the search space of the non-Markovian double pole-balancing task is inherently deceptive. However, the closer the lengths of the two poles, the more difficult the task becomes [11]. Task difficulty and hence complexity of the search space has been equated by some as deception [22], [9], [4], in that as task complexity increases it becomes increasingly difficult to design an appropriate fitness function, and objective based search becomes more vulnerable to deception [42]. In the pole-balancing task we hypothesize that deception plays a role in that the likelihood of the fitness function misguiding the search of an NE process, causing convergence to local optima, increases with the task difficulty.

An essential component of NS is its *novelty metric* that determines how *novel* evolved solutions are, compared to previously evolved solutions stored in a *novelty archive* [22]. Dissimilar to objective based approaches, where selection pressure acts to adapt solutions according to a fitness function, NS consistently produces novel solutions via maintaining a constant pressure to produce novelty. However, it has been noted that NS performs poorly in tasks with large search (behavior) spaces [1], [24], [23], [30]. That is, whilst NS is not strictly a behavioral diversity maintenance method [32], it is tantamount to a process that optimizes for diversity (novelty). However, as the novelty archive increases in size it becomes increasingly costly for the novelty metric to find nearest neighbors [32]. Potential solutions include adapting the behavior metric or combining novelty and objective based search to form a *hybrid* approach [4]. Combining novelty and objective based search into a hybrid that uses a weighted sum has been demonstrated as yielding superior task performance in a deceptive version of *Tartarus* [1] and swarm robotic tasks [8] compared to NS or objective based approaches.

978-1-4799-7492-4/15/$31.00 ©2015 IEEE

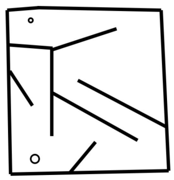

Fig. 1. **Maze Navigation Map.** The large circle represents the starting position of the robot and the small circle represents the goal. Dead-ends near the goal creates the potential for deception [22].

Given this, the research objective of this study was to comparatively test NS, objective based search and a novelty-fitness hybrid in a double pole balancing task. This is a part of continuing research that aims to characterize and define the types of tasks, behavior and genotype representations, fitness and novelty metrics (for example, deception [22], various types of behavior characterization [19], and various hybrid fitness-novelty metrics [30]) that benefit from NS versus objective-based and hybrid fitness-novelty search. A key motivation for this objective was that despite research that demonstrates NS as being beneficial in various types of tasks, with varying behavior and genotype representations, and novelty metrics [19], there is conflicting evidence for when NS [1] versus hybrid search methods [17] are advantageous.

To address this objective, this study comparatively tests NS, objective based search and a hybrid search method with respect to two types of behavior characterization schemes in the non-Markovian double pole balancing task [14].

II. Methods

The NE method used was *Neuro-Evolution of Augmenting Topologies* (NEAT) [39]. NEAT is an established NE method that was selected since it has been previously employed in similar studies that have applied NS versus objective based search [22], [41], [36], [37]. NEAT evolves both the topology and the weights of ANNs via a process called *complexification*. That is, at the start of artificial evolution, ANNs in the population are functionally simple, with a minimal number of nodes and connections. During the course of evolution, further nodes and connections are added to ANNs, where increasing the number of nodes and connections in an ANN increases the search space dimensionality. It is hypothesized that NEAT evolved ANNs are only as complex as is required to solve the given task [39]. This decreases the number of evaluations by finding a solution in a low dimensional search space. Other distinguishing features of NEAT are speciation, which protects innovation, and historical markings, which aid in the crossover of structurally different ANNs. Also, NEAT complements NS in that it encourages genotypic diversity whilst NS encourages phenotypic (behavioral) diversity [9].

A. ANN Controller

Given that the task is non-Markovian double pole balancing, controllers are only supplied with the cart position and the relative pole-angles (positions) and not the pole velocities. The ANN controller used in this study is the same as used in previous research [14], where three sensory input nodes accept the cart position and the two pole positions (angles). Sensory inputs are connected to an intermediate hidden layer which is in turn connected to a motor output node that yields the cart force. The number of input and output nodes are fixed, whereas NEAT adapts the number of hidden layer nodes and connectivity and connection weight values between layers.

TABLE I. **Experiment Parameters**: Parameters in **bold** are specific to NS and hybrid search. Fixed and Proportional Sampling apply only to NS and hybrid search. Other parameters apply to NS, objective and hybrid search.

Parameter	Fixed Sampling	Proportional Sampling
Novelty Threshold	0.05	0.05
Novelty Archive Size	Unbound	Unbound
Sampled Time Steps	10, 50, 100, 200	1/8, 2/8, 3/8, 4/8
-	500, 1,000, 5,000, 10,000	5/8, 6/8, 7/8, 8/8
Experiment runs	100	500
Simulation task trial iterations	100,000	100,000
Task trials per NEAT iteration	1	1
k-nearest neighbors	15	15
Pole lengths	[0.5, 0.05]	[0.5, 0.05]
Pole velocities	Hidden	Hidden

At each iteration of a pole balancing simulation task trial, the controller receives state variable values normalized to the range [-1.0, 1.0]. Controller output is cart force exerted such that simulation transitions to the next state which is then a new sensory input supplied to the controller. This sensory-motor cycle is repeated until a pole falls or the cart goes off the end of the track. Task simulation parameters are the same as those used in previous research [14], and the experimental setup is described in section III.

Adaptation of the ANN controller for pole balancing behavior was guided by either a *fitness function* (section III), *novelty search* (section II-B) or a *hybrid objective-novelty* based method (section II-C).

B. Novelty Search

NS encourages novelty and diversity in its solutions via keeping an archive of novel solutions. The novelty metric is applied to newly evolved solutions in order that the novelty of new solutions is ascertained. If a new solution's novelty exceeds the novelty metric's threshold, then that solution is added to the archive [22]. In the context of a search space of behaviors, if a solution is in a dense cluster of previously visited behaviors, it would be considered less novel and thus rewarded less. Thus, significantly novel behaviors are added to the archive over the course of an iterative search process.

In this study the fitness function of NEAT was replaced with a novelty metric, where novelty scores were computed via comparing pole-balancing behaviors (a vector of pole angles) with those behaviors stored in the novelty archive. However, we tested two behavior characterization schemes, each of which was defined by a procedure that sampled pole positions at given iterations during a pole balancing simulation. These behavior characterization schemes were known as *Fixed* and *Proportional Sampling* (section II-D).

1) Novelty Metric: In this study, the pole angles sampled at given iterations during a simulation task trial represented the pole-balancing behavior of a given ANN controller. Thus the novelty metric compared vectors of pole angles in order to ascertain if an evolved behavior was sufficiently different (novel) from those stored in the novelty archive. Hence, novel behavior was that which exceeded the novelty metric's threshold (table I) in the comparison of any two behaviors (pole position vectors).

Each genotype was a direct encoding (that is, no developmental encoding) of a vector of pole positions, for the two poles, for a given sample of simulation iterations (table I). The novelty metric was thus based on the *sparseness* (equation 1) of behaviors. Sparseness was measured as the average distance between the *k*-nearest neighbors of a given behavior [22], where *k* is fixed (table I). Neighbors are composed of other behaviors in the same generation as well as neighbors in the novelty archive.

$$\text{Sparseness}(x) = \frac{1}{k}\sum_{i=0}^{k} dist(x, \mu_i) \tag{1}$$

where μ is the *i*th-nearest neighbor of *x* with respect to the novelty metric, and where the distance component in equation 1 uses the Euclidean distance derived by the Pythagorean theorem [15] (equation 2).

$$dist(x, \mu) = \sqrt{(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2 + ... + (x_n - \mu_n)^2} \tag{2}$$

Given that the novelty metric records *N* samples (table I), each behavior is thus composed of *N* dimensions for each pole. When behavioral distances were computed, each dimension of a behavior (pole position vector component) was compared with the corresponding dimension in the other behavior.

C. Hybrid Novelty-Objective Based Search

Hybrid search methods, combining novelty metrics and fitness functions, have been demonstrated as yielding superior task performance, compared to NS or objective based search in a range of tasks [2], [10], [17]. This is hypothesized to be a result of hybrid methods introducing a bias in search towards optimal solutions which are potentially ignored by NS or not readily attainable by fitness functions given rapid loss of genotypic diversity [22].

The hybrid method proposed by Inden *et al.* [17] used selection as its defining feature. For example, half a genotype population was selected based on fitness and the other half based on novelty. Another approach is to take the most promising solutions found from NS thus far and further optimize them using a fitness function. This approach exploits the strengths of both methods, using NS to find good approximate solutions and then using objective based search to *tune* these approximate solutions in order to increase the likelihood that optimal solutions are attained [22, 29]. Specific hybrid methods include combining novelty and fitness functions linearly [2], and using a a multi-objective formulation that treats novelty and fitness as complementary objectives to be optimized [28].

The hybrid search method used in this study linearly combines varying biases for novelty and fitness [2]. Thus via adjusting novelty and fitness biases, the experimenter is able to control the complementary degrees of novelty and objective based search that the hybrid is optimizing for. The score that each solution receives for the hybrid method is defined as:

TABLE II. **SUMMARY OF EXPERIMENTS**: ρ DEFINES THE NOVELTY-FITNESS PROPORTION OF THE HYBRID SEARCH METHOD.

Method	ρ	NEAT Iterations
Objective based	1.0	100
Novelty based	0.0	100
Hybrid	0.25, 0.5, 0.75	100

$$\text{score}(i) = \rho.\overline{fit}(i) + (1-\rho).\overline{nov}(i) \tag{3}$$

Where, $\rho \in [0, 1]$ controls the relative weighting of fitness and novelty, which are normalized according to:

$$\overline{fit}(i) = \frac{fit(i) - fit_{min}}{fit_{max} - fit_{min}}, \overline{nov}(i) = \frac{nov(i) - nov_{min}}{nov_{max} - nov_{min}} \tag{4}$$

Where, nov_{min}, fit_{min} are the lowest novelty and fitness values in the population, respectively and nov_{max}, fit_{max} are the highest. A high value of ρ indicates a bias towards objective based search whilst lower values of ρ indicate a bias towards NS. Consequently, $\rho = 1$ means that only objective based search was used and $\rho = 0$ means that only NS was used [2].

D. Novelty and Hybrid Sampling Approaches

For the NS and hybrid methods, two methods for sampling pole positions during a task trial simulation were tested (table I). These were Fixed and Proportional Sampling, where Fixed Sampling sampled pole positions at [10, 50, 100, 200, 500, 1,000, 5,000, 10,000] iterations in a pole-balancing simulation. This was the same sampling range used in previous pole-balancing experiments [17], where 10,000 was set as the maximum simulation duration given that those controllers that did not achieve a pole balancing behavior by 10,000 iterations were highly unlikely to do so thereafter. However, Proportional Sampling sampled pole positions at one eighth portions of the simulation length (table I). We hypothesized that this would be a more appropriate behavior characterization scheme for this study, discussed in section IV.

III. EXPERIMENTS AND TASK

The task was non-Markovian double pole balancing with incomplete state information (pole velocities were hidden) and fourth-order Runge-Kutta integration [14]. Pole balancing simulation task trials ran for a maximum of 100,000 simulation iterations (table I), where a controller's sensory-motor cycle was processed every iteration, corresponding to approximately 0.01 second of real time. The same simulation parameters as previous research were used [14].

For the objective based and hybrid methods (section II-C), *fitness* was determined by the number of iterations a controller could keep both poles within a specified failure angle from vertical and the cart between the ends of the track. In line with previous research [14], the failure angle was 36 degrees, the range of cart track was [-2.4, 2.4], and pole lengths were set such that the second pole was to be 1/10th the length of the first pole (table I).

TABLE III. **NEAT PARAMETERS**: PARAMETERS USED FOR THE OBJECTIVE BASED, NOVELTY AND HYBRID METHODS.

Parameter	Objective	Novelty	Hybrid
Population Size	1000	1000	1000
Add Connection Probability	0.3	0.3	0.3
Mutate Connection Probability	0.1	0.1	0.1
Add Neuron Probability	0.01	0.01	0.01
Mutate Neuron Probability	0.01	0.01	0.01
Survival Threshold	0.4	0.4	0.4
Recurring Probability	0.2	0.2	0.2

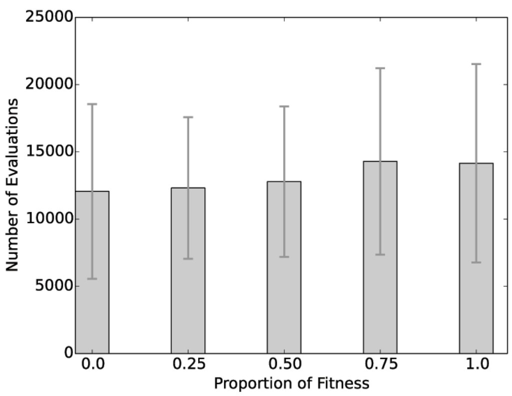

Fig. 2. *Fixed Sampling* behavior characterization. Average number of evaluations until a pole balancing behavior was achieved for increasing values of ρ, starting with pure novelty ($\rho = 0$) up to pure fitness ($\rho = 1$). Results from $\rho = 0$ are different with statistical significance (Mann-Whitney U test, $p \leq 0.05$) compared to all other ρ values.

Experiments comparatively tested three methods for directing the search process of NEAT (table II). That is, NS [22], a fitness function [40] and a hybrid that used a linear combination of novelty and fitness. Table II presents an overview of the experiments. Table I outlines parameters used for the comparative methods, where Fixed and Proportional Sampling are different behavior characterization (pole position sampling) schemes (section II-D). Also, the *Novelty Threshold*, *Novelty Archive Size*, and *Sampled Time Steps* parameters are specific to the NS and hybrid methods. Table III presents the NE (NEAT) parameters common to the NS, objective based and hybrid search methods.

The comparative task performance (average number of evaluations required to evolve a pole-balancing behavior) of NS, the objective based and hybrid search methods, with respect to two types of behavioral characterization used (Fixed and Proportional Sampling) were measured over 100 runs (section IV). This task performance measure was used as it is consistent with previous research [14].

IV. RESULTS AND DISCUSSION

Figure 2 presents results from the Fixed Sampling behavior characterization scheme, where NS, objective based and hybrid methods are tested via adjusting the ρ value (table II). Figure 2 presents the average number of evaluations (calculated over 100 runs per ρ value) required to reach the highest performing pole-balancing behavior, for NS ($\rho = 0$), objective based search ($\rho = 1$), and a range of ρ values in between indicative of varying complements of fitness and novelty for the hybrid method. A *Mann-Whitney U* statistical test [5] yielded a statistically significant difference ($p \leq 0.05$) between NS ($\rho = 0$) and all other ρ values, indicating the benefit of NS when Fixed Sampling behavior characterization was used. No statistically significant difference was found between the objective based and hybrid methods for Fixed Sampling.

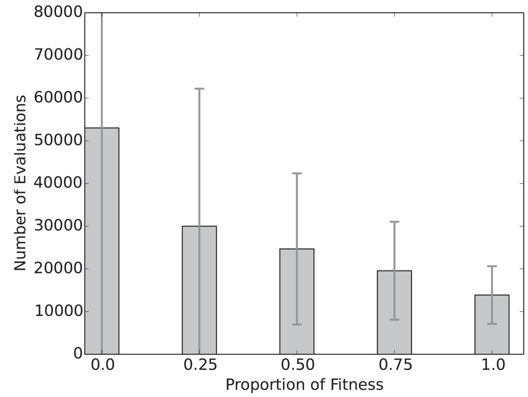

Fig. 3. *Proportional Sampling* behavior characterization. Average number of evaluations until a pole balancing behavior was achieved for increasing values of ρ, starting with pure novelty ($\rho = 0$) up to pure fitness ($\rho = 1$). All ρ values are different with statistical significance (Mann-Whitney U test, $p \leq 0.05$) compared to all other ρ values.

Thus NS out-performed the objective based and hybrid search methods, which is inconsistent with the results of Inden *et al.* [17] that indicated NS performed comparably to objective based search and poorly compared to a hybrid method in the non-Markovian double pole balancing task. A similar behavior characterization scheme (Fixed Sampling) was also used by Inden *et al.* [17], however the cart position was sampled instead of pole positions, as in this study. Also, Inden *et al.* [17] used a hybrid method based on selection [17]. These experimental differences are theorized to explain the difference in results presented here, though this is the subject of ongoing research.

A potential issue with Fixed Sampling (section II-D) is that pole-balancing behaviors would fail early in a simulation prior to a behavior sample (pole positions) being taken. This would result in default values (0) being placed in the pole position vector elements of the given behavior. Figure 4 illustrates an example of this, where behavior *A* fails before behavior *B* and thus includes more default 0 values. In the context of NS, the novelty metric would select for behaviors that are sufficiently different (that is, novel). That is, behaviors that have fewer default 0 values, and thus more pole position samples, would be selected by NS. This would correspond to improved pole balancing behaviors, that is, those behaviors that have few or no failures (0 values).

Similarly, in the case of the objective based and hybrid methods, the fitness function (section III) selects for behaviors that fail late in the simulation, or not at all. That is, genotypes with few or no default 0 values in. The mitigating factor of NS and the hybrid search method is that the novelty metric encourages selection from a diverse range of sufficiently different (novel) yet not necessarily optimal behaviors. The impact of progressively introducing increasing degrees of objective based

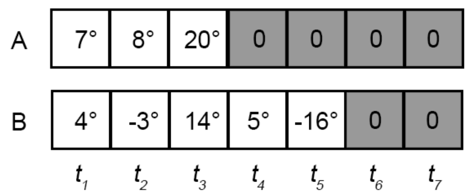

Fig. 4. Example pole-balancing behaviors, *A* and *B*, where behaviors consist of several pole angle samples taken during a pole-balancing simulation. The default value of 0 is given if the behavior has failed when the sample is taken.

search in the hybrid is evident in figure 2, where the average number of evaluations required by the hybrid to achieve a pole-balancing behavior increases such that NS ($\rho = 0$) yields a statistically significant higher task performance.

To remove any bias that potentially occurs with Fixed Sampling behavior characterization when solutions fail early in a simulation, and are thus explicitly selected against by the objective based or hybrid methods, the Proportional Sampling (section II-D) behavior characterization was devised. Proportional Sampling characterized pole-balancing behavior via sampling pole positions proportional to how many simulation iterations transpired before the behavior failed. In this study the sampling rate was *one eighth* meaning that a behavior (pole positions') sample was taken every 12,500 iterations and there were eight samples in total, where a pole balancing simulation could run for a maximum of 100,000 iterations (table I).

However, if the pole-balancing behavior failed before iteration 100,000, then samples would be taken at one eighth portions up until the iteration where the pole-balancing behavior failed, given that pole-positions were recorded at every simulation iteration. For example, if a pole balancing behavior failed at iteration 1000, then samples would be taken at every 125 iterations, in order that *this* behavior consists of eight pole positions. A key idea of Proportional Sampling was that behaviors that contained initially good pole angles, but that later failed during the simulation would not necessarily be selected against, given that a pole balancing behavior may be briefly achieved but then subsequently fail.

Figure 3 presents results from Proportional Sampling, where NS, the objective based and hybrid methods are tested via adjusting the ρ value (table II). Figure 3 presents the average number of evaluations (calculated over 500 runs per ρ value), required to attain the best performing pole-balancing behavior for NS ($\rho = 0$) and objective based search ($\rho = 1$), as well as a range of ρ values in between indicative of varying complements of fitness and novelty for the hybrid method. A Mann-Whitney U statistical test yielded a statistically significant difference ($p \leq 0.05$) between all ρ values, indicating that the pure objective based search ($\rho = 1$) out-performed NS ($\rho = 0$) and the hybrid method when Proportional Sampling behavior characterization was used.

However, this result is indicative of the fitness function selecting for behaviors that maximize the number of simulation iterations for which poles were kept between the failure angles (section III). That is, given that Proportional Sampling ensures that only samples prior to pole-balancing behavior failure are taken, the objective based method optimized within this search space of behaviors converging upon optima within this restricted behavior set. Thus, even though pole-balancing behaviors that maximize the number of iterations are selected for, behaviors that fail at some point during the simulation are invariably selected.

The sub-optimal nature of the Proportional Sampling behavior characterization scheme is evident in a comparison of Fixed and Proportional sampling for all methods (figure 5). Figure 5 presents the comparative number of average evaluations (calculated over 100 runs) required for all methods, using Fixed versus Proportional Sampling to attain the highest performing pole-balancing behavior. This indicates that search methods using Fixed Sampling outperform Proportional Sampling, with statistical significance ($p \leq 0.05$ using the Mann-Whitney U statistical test), where on average approximately 40,000 fewer evaluations were required by methods using Fixed Sampling behavior characterization (figure 5). However, within Fixed Sampling, NS yielded the highest task performance, with statistical significance, over the objective based and hybrid methods (figure 2).

These results support the notion that behavioral characterization has a significant impact upon the success of the search method and the evolution of effective problem solving behaviors [9]. In this study there was a direct genotype to behavior mapping, so the composition of genotypes (vectors of pole positions) had a significant impact upon the fitness and novelty values assigned to genotypes (behaviors) and thus the success of a given search process.

This was especially the case for the objective based and hybrid methods using Fixed Sampling, where genotypes (behaviors) with few or no default 0 values were selected, meaning that Fixed compared to Proportional Sampling, indirectly resulted in the selection of more robust behaviors. Fixed Sampling also aided NS in the selection of effective behaviors given that the novelty metric selected behaviors that were sufficiently different from those already in the novelty archive. Given that behaviors containing many default 0 values were initially placed in the archive, as they were initially novel, NS subsequently selected novel behaviors which were those with relatively few or no default 0 values. That is, the comparatively higher performance of NS is attributable to the novelty metric which measures behavioral distance [19], and selects accordingly different (novel) high performing behaviors from across a broad spectrum of the behavior space.

Furthermore, NS is theorized to be comparatively successful in the pole-balancing task given that many points in the search space correspond to the same type of behavior [21]. That is, in the pole balancing task, there are many permutations of pole positions that can comprise a genotype, where all of these permutations correspond to an effective pole-balancing behavior. Another factor that we theorize as resulting in a deficiency in the hybrid method, leading to its comparatively inferior performance (figure 2), was the use of a linear combination of fitness and novelty metrics [2]. These results indicate that a linear combination of metrics is not always an appropriate approach given that hybrid fitness-novelty metrics are highly sensitive to their relative fitness and novelty weightings. Also, these results support the notion that

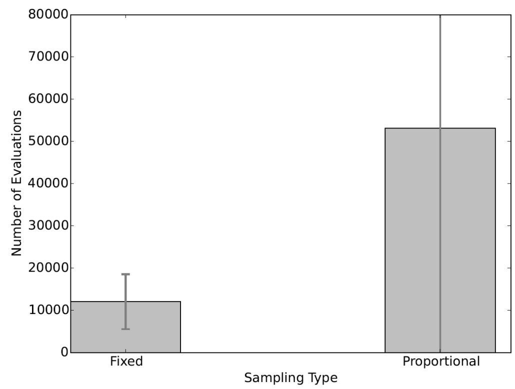

Fig. 5. Average number of evaluations comparing *Fixed* and *Proportional Sampling* behavior characterization methods for the search methods tested. Results are statistically significant (Mann-Whitney U test, $p \leq 0.05$).

linear combinations of novelty and fitness metrics are sensitive to behavior characterization and genotype representation.

The combination of factors leading to effective hybrid fitness-novelty guided search [2], [17], [9] remains the subject of current research, however one potential improvement in the pole-balancing would be the use of multi-objective evolutionary algorithms that combine novelty and fitness as complementary objectives [30].

V. Conclusion

This research presented a comparative study that tested *novelty*, *objective* based and *novelty-objective* hybrid search methods applied to behavior evolution in the non-Markovian double pole balancing task. Also, two different behavior characterization schemes were tested in combination with each method in order to ascertain the impact of behavior characterization on the respective search methods. In this task behaviors were characterized by a set of pole positions sampled at given iterations during a pole balancing simulation, and there was a direct mapping between genotypes and behaviors. The behavior characterization schemes were *Fixed* and *Proportional Sampling*, where behaviors were characterized by pole positions sampled at fixed intervals and in proportion to the simulation length, respectively.

Results indicated that methods using Fixed Sampling behavior characterization significantly out-performed those using Proportional Sampling in terms of the number of the average number of evaluations required to evolve the best performing pole balancing behavior. Within the context of methods using Fixed Sampling, NS was found to yield a higher performance compared to the objective based and hybrid search methods. However, under the Proportional Sampling behavior characterization scheme, NS performed relatively poorly compared to the other search methods. Thus results indicate the importance of behavior characterization when using NS, given that the novelty metric relies upon behavioral differences in its selection process.

Future research will test other hybrid novelty-fitness metrics, such as multi-objective optimization approaches [30], as well as test novelty search with various behavior characterizations [19] in tasks with more complex and deceptive search spaces [9].

References

[1] G. Cuccu and F. Gomez. When novelty is not enough. In C. Di Chio, S. Cagnoni, C. Cotta, M. Ebner, A. Ekart, A. Esparcia-Alcazar, J. Merelo, F. Neri, M. Preuss, H. Richter, J. Togelius, and G. Yannakakis, editors, *Applications of Evolutionary Computation*, pages 234–243. Springer, Berlin, Germany, 2011.

[2] G. Cuccu and F. Gomez. When novelty is not enough. In *Applications of Evolutionary Computation*, pages 234–243. Springer, 2011.

[3] P. Darwen and X. Yao. A dilemma for fitness sharing with a scaling function. In *Proceedings of Congress on Evolutionary Computation*, volume 1, pages 166–171. IEEE Press, 1995.

[4] S. Doncieux and J-B. Mouret. Beyond black-box optimization: a review of selective pressures for evolutionary robotics. *Evolutionary Intelligence*, 71(1):1–23, 2014.

[5] B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical Recipes*. Cambridge University Press, Cambridge, 1986.

[6] D. Floreano, P. Dürr, and C. Mattiussi. Neuroevolution: from architectures to learning. *Evolutionary Intelligence*, 1(1):47–62, 2008.

[7] S. Galan and O. Mengshoel. Generalized crowding for genetic algorithms. In *Proceedings of Genetic and Evolutionary Computation Conference*, pages 775–782. ACM, 2010.

[8] J. Gomes and A. Christensen. Generic behaviour similarity measures for evolutionary swarm robotics. In *Proceedings of the International Conference on Genetic and Evolutionary Computation*, pages 199–206, Amsterdam, the Netherlands, 2013. ACM Press.

[9] J. Gomes, P. Urbano, and A. Christensen. Evolution of swarm robotics systems with novelty search. *Swarm Intelligence*, 7(2):115–144, 2013.

[10] J. Gomes, P. Urbano, and A. Christensen. Evolution of swarm robotics systems with novelty search. *Swarm Intelligence*, 7(2-3):115–144, 2013.

[11] F. Gomez and R. Miikkulainen. Solving non-markovian control tasks with neuroevolution. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1356–1361, Stockholm, Sweden, 1999. Morgan Kaufmann.

[12] F. Gomez and R. Miikkulainen. Active guidance for a finless rocket using neuroevolution. In *Proceedings of Genetic and Evolutionary Computation Conference*, pages 2084–2095. Springer, 2003.

[13] F. Gomez, J. Schmidhuber, and R. Miikkulainen. Efficient non-linear control through neuroevolution. In *Machine Learning: European Conference on Machine Learning*, pages 654–662. Springer, 2006.

[14] F. Gomez, J. Schmidhuber, and R. Miikkulainen. Accelerated neural evolution through cooperatively coevolved synapses. *Journal of Machine Learning Research*, 9(1):937–965, 2008.

[15] J. Gower. Euclidean distance geometry. *Mathematical Scientist*, 7(1):1–14, 1982.

[16] F. Gruau, D. Whitley, and L. Pyeatt. A comparison

between cellular encoding and direct encoding for genetic neural networks. In *Proceedings of the First Annual Conference on Genetic Programming*, pages 81–89. MIT Press, 1996.

[17] B. Inden, Y. Jin, R. Haschke, H. Ritter, and B. Sendhoff. An examination of different fitness and novelty based selection methods for the evolution of neural networks. *Soft Computing*, 17(5):753–767, 2013.

[18] H. Jin and M. Wong. Adaptive diversity maintenance and convergence guarantee in multiobjective evolutionary algorithms. In *Proceedings of Congress on Evolutionary Computation*, volume 4, pages 2498–2505. IEEE Press, 2003.

[19] S. Kistemaker and S. Whiteson. Critical factors in the performance of novelty search. In *Proceedings of Genetic and Evolutionary Computation Conference*, pages 965–972. ACM, 2011.

[20] P. Krcah. Solving deceptive tasks in robot body-brain co-evolution by searching for behavioral novelty. In *Proceedings of the International Conference on Intelligent Systems Design and Applications (ISDA2010)*, pages 284–289, Cairo, Egypt, 2010. ACM Press.

[21] J. Lehman and K. Stanley. Exploiting open-endedness to solve problems through the search for novelty. In *Proceedings of the Artificial Life Conference (ALife XI)*, pages 329–336, Winchester, UK, 2008. ACM Press.

[22] J. Lehman and K. Stanley. Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary computation*, 19(2):189–223, 2011.

[23] J. Lehman, K. Stanley, and R. Miikkulainen. Effective diversity maintenance in deceptive domains. In *Proceedings of the International Conference on Genetic and Evolutionary Computation*, pages 215–222, Amsterdam, the Netherlands, 2013. ACM Press.

[24] J. Lehman and K. O. Stanley. Revising the evolutionary computation abstraction: Minimal criteria novelty search. In *Proceedings of the International Conference on Genetic and Evolutionary Computation*, pages 103–110, Portland, USA, 2010. ACM Press.

[25] B. Luo, J. Zheng, J. Xie, and J. Wu. Dynamic crowding distance? a new diversity maintenance strategy for moeas. In *Proceedings of International Conference on Computing, Networking and Communications*, volume 1, pages 580–585. IEEE Press, 2008.

[26] R. Miikkulainen. Neuroevolution. In C. Sammut and G. Webb, editors, *Encyclopedia of Machine Learning*, pages 716–720. Springer, New York, USA, 2010.

[27] D. Moriarty. *Symbiotic evolution of neural networks in sequential decision tasks*. PhD thesis, 1997.

[28] J. Mouret. Novelty-based multiobjectivization. In *New Horizons in Evolutionary Robotics*, pages 139–154. Springer, 2011.

[29] J. Mouret and S. Doncieux. Overcoming the bootstrap problem in evolutionary robotics using behavioral diversity. In *Proceedings of Congress on Evolutionary Computation*, pages 1161–1168. IEEE Press, 2009.

[30] J-B. Mouret. Novelty-based multi-objectivization. In *New Horizons in Evolutionary Robotics: Extended contributions of the 2009 EvoDeRobWorkshop*, pages 139–154. Springer, Berlin, Germany, 2011.

[31] J-B. Mouret and S. Doncieux. Using behavioral exploration objectives to solve deceptive problems in neuro-evolution. In *Proceedings of the 11th annual conference on Genetic and evolutionary computation*, pages 627–634, Montreal, Canada, 2009. ACM Press.

[32] J-B. Mouret and S. Doncieux. Encouraging behavioral diversity in evolutionary robotics: An empirical study. *Evolutionary Computation*, 20(1):91–133, 2012.

[33] S. Nolfi and D. Floreano. *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines.* MIT Press, Cambridge, USA, 2000.

[34] S. Risi, S. Vanderbleek, C. Hughes, and K. Stanley. How novelty search escapes the deceptive trap of learning to learn. In *Proceedings of the 11th Annual Conference on Genetic and evolutionary computation*, pages 153–160, Montreal, Canada, 2009. ACM Press.

[35] N. Saravanan and D. Fogel. Evolving neural control systems. *IEEE Intelligent Systems*, 10(3):23–27, 1995.

[36] D. Shorten and G. Nitschke. How evolvable is novelty search? In *SSCI 2014*, pages 125–132, Orlando, USA, 2014. IEEE Press.

[37] D. Shorten and G. Nitschke. Evolving generalised maze solvers. In *Proceedings of EvoStar 2015*, Copenhagen, Denmark, 2015. Springer.

[38] A. Soltoggio and B. Jones. Novelty of behaviour as a basis for the neuro-evolution of operant reward learning. In *Proceedings of the International Conference on Genetic and Evolutionary Computation*, pages 169–176, Montreal, Canada, 2009. ACM Press.

[39] K. Stanley and R. Miikkulainen. Evolving neural networks through augmenting topologies. *Evolutionary Computation.*, 10(2):99–127, 2002.

[40] K. Stanley and R. Miikkulainen. Evolving neural networks through augmenting topologies. *Evolutionary computation*, 10(2):99–127, 2002.

[41] R. Velez, , and J. Clune. Novelty search creates robots with general skills for exploration. In *Proceedings of the conference on Genetic and evolutionary computation*, pages 737–744, Vancouver, Canada, 2014. ACM.

[42] D. Whitley. Fundamental principles of deception in genetic search. In *Foundations of Genetic Algorithms*, pages 221–241. Morgan Kaufmann, San Mateo, USA, 1991.

[43] B. Yuan and M. Gallagher. On the importance of diversity maintenance in estimation of distribution algorithms. In *Proceedings of Genetic and Evolutionary Computation Conference*, pages 719–726. ACM, 2005.