Evolving Bipedal Locomotion through Objective-based Search, Novelty Search, and Resource Search

Nafis Zaman and Dean Hougen

Abstract

Evolutionary algorithms, such as objective-based search and novelty search, have been compared on problems such as the maze problem and the bipedal locomotion problem. Resource search has outperformed both algorithms in solving mazes, but its viability on more complicated problems is unknown. We evolve mobile bipeds using resource search and compare its performance to objective-based search and novelty search. Over 30 runs, the best bipeds evolved by resource search walked the farthest. We conclude that resource search is applicable to complex domains outside the maze problem.

1 Introduction

Objective-based search is a common evolutionary algorithm that rewards agents who are "closer" to the desired objective. In a physical representation of a problem space, "closer" could be measured by physical distance with respect to the goal. By explicitly driving agents towards the goal, objective-based search fails to account for solutions that initially require steps away from the goal. This problem, known as deception, can manifest in various problem spaces. For example, in mazes, correct solutions may involve paths that initially move away from the goal in order to circumvent walls. In bipedal locomotion, robots that evolve oscillatory patterns necessary for walking may fall down without moving very far.

Novelty search aims to mitigate deception by judging fitness based not on the objective, but on the uniqueness of the agents themselves. Agents that explore new areas of the search space are rewarded regardless of how well they did in terms of the objective (Lehman and Stanley, 2011). In the biped problem, robots that exhibit novel movements or directions are considered to be fit even if they barely leave the starting location. However, by ignoring the goal, novelty search fails to take advantage of the objective at all.

Resource search attempts to combine the advantages of both objective-based search and novelty search. That is, it attempts to avoid deception while still encouraging agents towards the objective. A limited amount of food is scattered throughout the environment, with more food placed in areas closer to the goal. If an agent fails to collect the required food, it automatically dies. The limitation on resources prevents many agents from exploring the same area, while the existence of food away from the objective allows some agents to explore new solutions and still be selected for reproduction. In addition, the non-uniform distribution of food reduces competition for agents that approach the goal (McDorman and Hougen, 2014).

Compared to objective-based search and novelty search, resource search has proven effective for the maze problem. We would like to strengthen the validity of resource search by applying it to a more complex domain: bipedal locomotion. Walking effectively requires evolving an oscillatory pattern with a considerable amount of control and balance in three-dimensional space.

2 Hypotheses

Lehman and Stanley (2011) have already compared novelty search to objective-based search in the biped domain. We aim to replicate these results in our own environment. Moreover, we will evaluate resource search in the same environment against the other two algorithms.

The experiments will test the following three hypotheses:

H_1 Novelty search will evolve robots that walk farther than those evolved by objective-based search.

H_2 Resource search will evolve robots that walk farther than those evolved by objective-based search.

H_3 Resource search will evolve robots that walk farther than those evolved by novelty search.

3 Methods

In order to apply resource search as generally as possible (i.e., without bias towards the particular algorithm), we tried to keep the environment and experimental parameters consistent with the work of Lehman and Stanley (2011).

The simulated biped resembles the lower body of a two-legged organism, as pictured in Figure 1. The waist, upper legs, and lower legs are cylinders and the feet are spheres. Each joint between an upper leg and lower leg can bend forwards and backwards (i.e., pitch). Each joint between an upper leg and the waist can bend forwards and backwards (i.e., pitch) and twist (i.e., roll). This results in six degrees of freedom within its four joints. Table 1 shows the physical biped parameters we used in our experiment.

We used the Open Dynamics Engine (2014) to simulate the bipeds' physical motions. They perform on a flat plane that extends as far as necessary along both axes. It is important to note that the objective is to walk as far as possible in *any* direction. There is no concept of forwards or backwards.

Figure 1: 3D model of the simulated biped

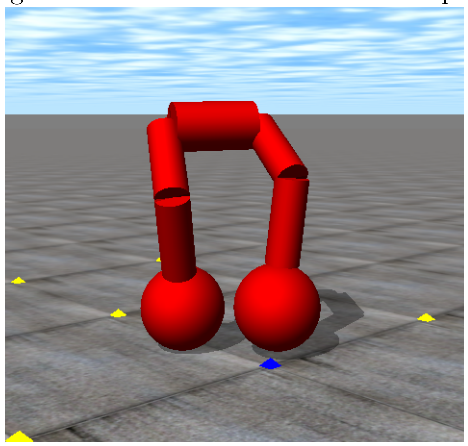


Table 1: The parameters used for the physical biped

Parameter	Value
Foot radius	0.17 m
Foot density	1 kg/m^3
Torso radius	0.1 m
Torso length	0.33 m
Torso density	1 kg/m^3
Leg segment radius	0.2 m
Leg segment length	0.33 m
Leg segment density	1.0 1 kg/m^3
Max torque	5.0 N·m
Proportional constant	9.0

The bipeds are controlled by continuous time-recurrent neural networks, which can properly represent the nonlinear dynamics inherent in walking. The neural network has two sensor inputs: a contact sensor on each foot. Each sensor can report “on” or “off” depending on whether the foot is touching the ground plane or not. The lack of other information, such as the position of the legs or orientation of the body, makes the problem even more difficult. The network outputs a target angle for each degree of freedom. which the controller uses to apply torque based on the joints’ current angles.

To evolve, the neural networks employ the NeuroEvolution of Augmenting Topologies (NEAT) method, which gradually adds to and modifies the actual topology of the network (Stanley and Miikkulainen, 2002). However, other neuroevolution techniques could be applied as long as it is used for all methods. We chose NEAT to maintain consistency with the previously conducted biped experiment. Table 2 shows the NEAT parameters we used in our experiment.

Table 2: The parameters used for the NEAT method

Parameter	Description	Value
Pop. size	Initial size of the population	500
Excess coefficient	Weight given to excess genes in compatibility distance	1.0
Disjoint coefficient	Weight given to disjoint genes in compatibility distance	1.0
Mutation diff. coefficient	Weight given to differences in connection weights	3.0
Prob. add link	Probability of adding a link during mutation	0.06
Prob. add node	Probability of adding a node during mutation	0.005
Prob. mutate time const.	Probability of changing the time constant during mutation	0.3
Prob. mutate bias	Probability of mutating the bias during mutation	0.3

During evaluation, a biped was allowed to perform for a maximum of 15 seconds. If the biped fell over before the time elapsed, the simulation was terminated. A biped was considered to have fallen over if the height of its torso was less than half of its starting height. The distance traveled was measured as the Euclidean distance between the biped's initial center of mass and final center of mass.

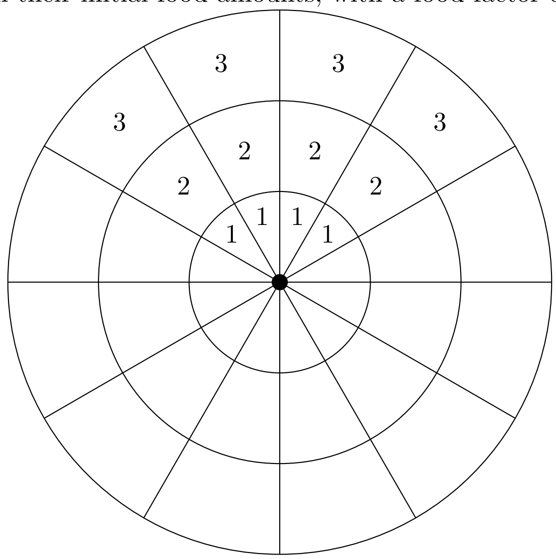
For a particular evolutionary algorithm, a single run consisted of a cycle of selection, reproduction, and evaluation of 500,000 agents. The selection process depended on the algorithm used. Reproduction occurred by crossing over and mutating the neural networks, as described by Stanley and Miikkulainen (2002) and used by Lehman and Stanley (2011), and was the same for all algorithms. Evaluation was also the same for all algorithms, as described above. We conducted 30 runs for each of the 3 algorithms of interest.

In objective-based search, bipeds were assigned fitness corresponding to the measured Euclidean distance between their initial and final locations.

In novelty search, bipeds were assigned fitness based on how novel their behavior was. Each biped's behavior was represented by a vector of 15 positions, with one position recorded per second during evaluation. If the biped falls, the remaining positions are filled with the biped's final position after the fall. The novelty distance between bipeds was equal to the sum of squares between their behavior vectors. Then, the novelty metric is based on a biped's k-nearest neighbors with $k = 15$.

In resource search, the environment is split into sectors based on a polar coordinate system, as shown in Figure 2. Concentric circles around the starting point divide the area into rings. The radii of the rings increases linearly. The area is also split into slices by lines extending outwards from the starting point.

Figure 2: Environment sectors for resource search with 12 slices. Some sectors are labeled with their initial food amounts, with a food factor of 1



Each slice covers the same arc angle. The intersections between these circles and lines create the sectors in which food is placed. There is an equal amount of food in sectors of the same ring. The amount of food increases linearly away from the starting point.

A biped's final sector is the sector containing the biped's final point. After running, a biped attempts to eat 2 units of food from its sector and 1 unit of food from each of its 8 surrounding sectors. This is the same consumption pattern used in the maze problem. If there isn't enough food to fulfill this requirement, then the biped dies and will not reproduce. Food is removed from the sectors regardless of whether the biped dies or not. Table 3 describes the resource search parameters and shows the values that we used in our experiment.

Table 3: The parameters used for resource search

Parameter	Description	Value
NUM_SLICES	The number of slices in each ring	20
SECTOR_LENGTH	The difference in radius between consecutive rings	0.05
FOOD_FACTOR	The factor by which food increases in the next sector	1

4 Results

Figure 3 shows the plot of the best agents' fitnesses evolved by each algorithm over time. The data was averaged over all 30 runs.

Figure 3: Average of best fitnesses

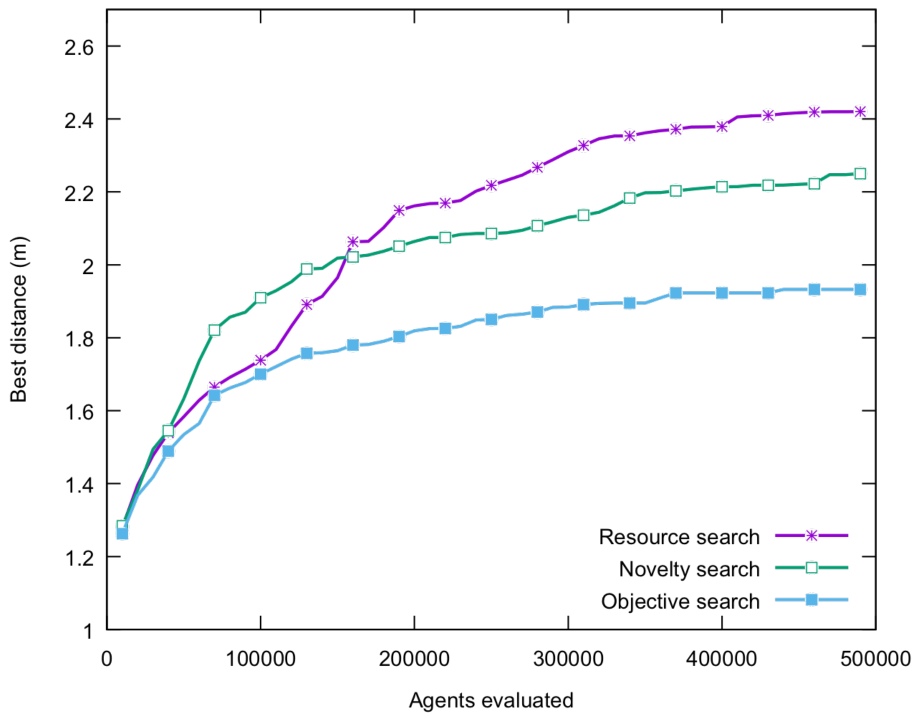


Figure 4 shows the plot of the agents' average fitnesses over time. The data was averaged over all 30 runs.

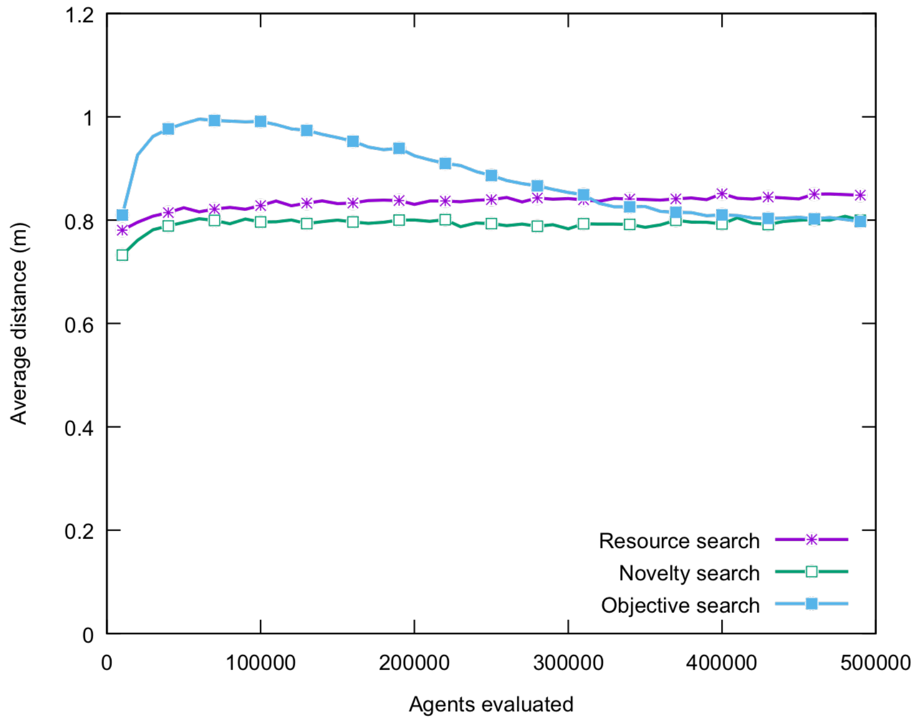
Figure 5 shows the population size of each algorithm over time. The data was averaged over all 30 runs.

To analyze the differences between the algorithms, we used a randomized version of analysis of variance (ANOVA). This version is more appropriate than standard ANOVA for performance curves due to carryover effects over time (Piater et al., 1998). Table 4 shows the resulting p-values.

Table 4: p-values from randomized ANOVA

	Algorithm effect	Interaction effect
Novelty vs. objective-based	0	0.0630
Resource vs. objective-based	0	0
Resource vs. novelty	0	0

Figure 4: Average of average fitnesses



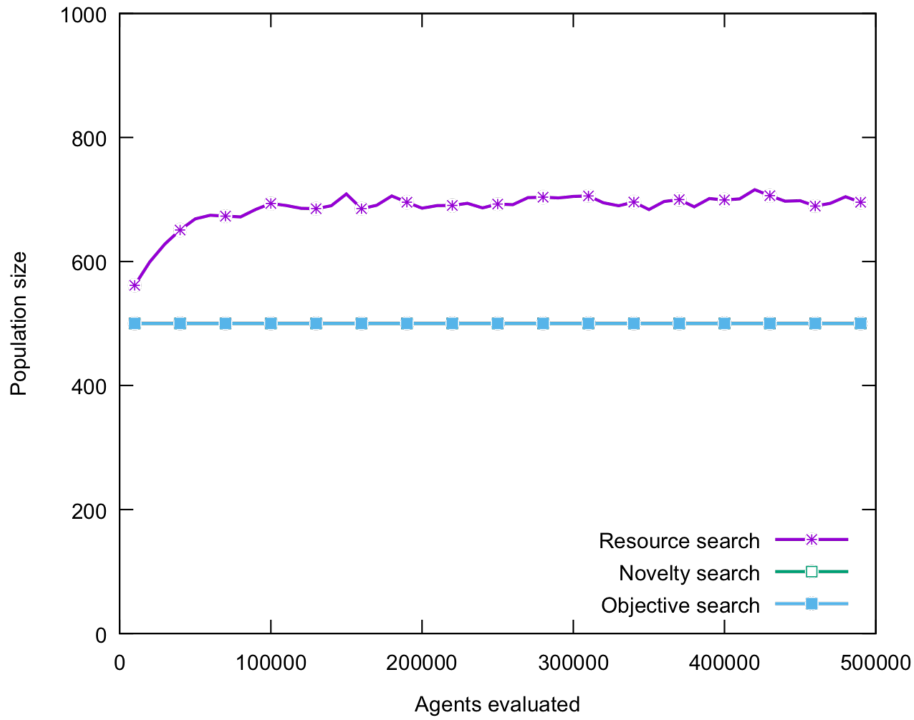
5 Discussion

Objective-based search and novelty search maintain a fixed population size throughout all generations. Resource search allows the population size to vary based on the number of agents that find sufficient resources. At the beginning, when most bipeds simply fall down, there is a lot of competition for the resources close to the starting point. These areas are quickly depleted, so the population size is small. As agents evolve and and take advantage of new niches, the population size grows.

Although the population sizes are different, our comparisons are valid because we use the number of agents evaluated, not the number of generations, as our independent variable.

For objective-based search and novelty search, our analysis of variance produced $p < 0.05$ for the algorithm effect. We reject the null hypothesis and conclude that the difference in performance is statistically significant. This is supported by the plot of best fitnesses, where the novelty search curve lies above the objective-based search curve. The best bipeds found by objective-based search walked 1.93 m on average, while those found by novelty search walked 2.26 m. Therefore, novelty search evolved bipeds that walked farther

Figure 5: Average population sizes



than those evolved by objective-based search, verifying hypothesis $\mathbf{H_1}$.

After the initial divergence, the curves appear to follow similar slopes. The identical shapes indicate that the relationship between performance and number of agents evaluated is comparable for objective-based search and novelty search. Our analysis of variance produced $p > 0.05$ for the interaction effect, so there is no evidence that the relationships are different.

For objective-based search and resource search, our analysis of variance produced $p < 0.05$ for the algorithm effect. We reject the null hypothesis and conclude that the difference in performance is statistically significant. This is supported by the plot of best fitnesses, where the resource search curve lies above the objective-based search curve. The best bipeds found by resource search walked 2.42 m on average. Therefore, resource search evolved bipeds that walked farther than those evolved by objective-based search, verifying hypothesis $\mathbf{H_2}$.

At the beginning, the resource search curve closely follows the objective-based search curve. However, shortly after evaluating 100,000 agents, the resource search curve quickly diverges and approaches a higher asymptote. The dissimilar shapes indicate that the relationship between performance and number of agents evaluated depends on the algorithm. Our analysis of variance supports this conclusion, producing $p < 0.05$ for the interaction effect.

For novelty search and resource search, our analysis of variance produced $p < 0.05$ for the algorithm effect. We reject the null hypothesis and conclude that the difference in performance is statistically significant. This is supported by the plot of best fitnesses, where the resource search curve levels out higher than the novelty search curve. Therefore, resource search evolved bipeds that walked farther than those evolved by novelty search, verifying hypothesis $\mathbf{H_3}$.

The novelty search curve follows steeper curves at the beginning. However, before the 200,000 agents evaluated mark, the resource search curve increases slope and crosses the novelty search curve. The dissimilar shapes indicate that the relationship between performance and number of agents evaluated depends on the algorithm. Our analysis of variance support this conclusion, producing $p < 0.05$ for the interaction effect. Although novelty search outperforms resource search initially, resource search catches up and passes novelty search after a reasonable number of agent evaluations.

6 Future Work

More work could be done to expand on the results in this paper:

- Further analysis on the bipeds evolved by the 3 algorithms could reveal differences between resource search and objective-based or novelty search. For example, do a significant number of bipeds get caught in deceptive behaviors? How many bipeds evolve oscillatory patterns? Are the bipeds stable (i.e., do small mutations cause significant drops in performance)?

- Modifying the parameters of resource search could improve results, as it is unlikely that we chose the parameters perfectly on the first try. In addition, this would test the robustness of resource search: From one set of parameters, can small perturbations result in significantly different performance?

- Implement sector clumping in the resource map. In other words, group successive sectors and put the same amount of food in each sector of a particular group. Food can increase linearly by group rather than by individual sectors.

- Apply evolutionary algorithms to more complex bipeds. For example, adding orientation sensors on the legs could help bipeds evolve more successful walking behaviors. Adding an upper body could help bipeds balance and avoid falling down after a bad step.

References

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

McDorman, B. and Hougen, D. (2014). A comparison of objective-based search, novelty search, combination search, and resource search on the maze problem. Unpublished manuscript, School of Computer Science, University of Oklahoma, Norman, OK.

Open Dynamics Engine (2014). Retrieved from http://ode-wiki.org/wiki/.

Piater, J. H., Cohen, P. R., Zhang, X., and Atighetchi, M. (1998). A randomized anova procedure for comparing performance curves. In *Proceedings of the Fifteenth International Conference on Machine Learning*, pages 430–438. Morgan Kaufmann Publishers.

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