

Population-Based Novelty Searches Can Converge

R. Paul Wiegand

wiegandrp@winthrop.edu

Department of Computer Science & Quantitative Methods

Winthrop University

Rock Hill, SC

Abstract

Novelty search is a powerful tool for finding sets of complex objects in complicated, open-ended spaces. Recent empirical analysis on a simplified version of novelty search makes it clear that novelty search happens at the level of the archive space, not the individual point space. The sparseness measure and archive update criterion create a process that is driven by a clear pair of objectives: spread out to *cover* the space, while trying to remain as efficiently *packed* as possible driving these simplified variants to converge to an ϵ -net in the sense defined by k Nearest Neighbor theory. Among the simplifications was omission of a population: the archive itself served as both the potential candidate set solution, as well as the source of parents in the evolutionary search, which has the potential to fundamentally hamstring the search. In this paper, we relax this crucial assumption, generalizing the previous simplified novelty search so that a traditional population and (μ, λ) dynamics are used to produce new search points, and the population and the archive are updated separately. We show empirically that it is still possible to converge, and that it can be *even more* likely to do so.

Introduction

Evolutionary algorithms (EAs) are increasingly being leveraged for creative endeavors from creating computer programs (Koza 1992) to producing art (Dreher 2014). Though there is a subfield of evolutionary computation dedicated to its ability to generate, create, and innovate (Goldberg 2002), there is not a lot of theory or foundational analysis for such applications.

One of the most successful and aggressively promoted concepts springing from innovative power of evolutionary methods is the idea of *Quality Diversity Algorithms* (Fontaine et al. 2019), such as *Novelty Search* (Stanley and Lehman 2015). The motivation of novelty search is that for some complex spaces, it is better to *ignore* the optimization objective and instead *explore* the space by discovering things that are increasingly “different” from those that the search process has encountered. Novelty search has been

surprisingly effective at finding good solutions to a number of challenging problems without even looking for those solutions directly, for instance discovering grasping behaviors for robotic arms (Huang et al. 2014), unsupervised feature learning for deep networks (Szerlip et al. 2015), discovering of gaits for quadruped locomotion (Morse et al. 2013), and goal-seeking in complex multiagent simulations (Lehman and Stanley 2011).

Advocates suggest at least three reasons novelty search has been so successful. The first is that novelty search avoids deceptive areas of the space by abandoning the objective in favor of a novelty metric (Stanley and Lehman 2015). Second, novelty search is typically used in conjunction with open-ended or generative representations. Finally, by using distance metrics within the behavior space rather than the raw genotype space, novelty search finds what is interesting in the space in which the designers are interested. Unfortunately, there is virtually no foundational analysis for novelty search to support such assertions.

This paper considers only the claim that novelty search has no objective, and that it is a “divergent” search process (Lehman, Wilder, and Stanley 2016; Lehman and Miikkulainen 2015; Stanley and Lehman 2015). This claim was addressed in a recent empirical analysis of a simplified form of novelty search (Wiegand 2020), which showed that under the most basic components of traditional novelty search (sparseness and an archive) the archive is driven toward increasingly better ϵ -covers of the search space while also trying to optimize the ϵ -packing of the archive. Simple variants of novelty search can converge, even in unbounded spaces.

But the algorithmic simplifications of this recent effort are severe and radically alter novelty search. In particular, though the simple novelty search evolutionary algorithm (SNSEA) that paper presents uses an archive to represent a candidate solution set, it does not use a population for evolutionary dynamics. Instead, parents are selected uniformly at random from the archive itself.

The current paper examines a population-based simple novelty search evolutionary algorithm (PSNSEA). We show empirically that adding a population and separate update dynamics for archives and populations, as in more traditional novelty search applications, does not necessarily alleviate

the problem that novelty search can converge, contradicting prior claims that they diverge.

Technical Approach

Sparseness and Archives

The *sparseness* of a candidate object y , ρ_y , is estimated by the average distance to the k nearest neighbors in a set A :

$$\rho_y := \frac{1}{k} \sum_{i=1}^k \delta(x_i^A, y), \quad (1)$$

where x_i^A is the i^{th} closest point in set A to y and δ is a distance function.

This sparseness score is used in two ways: 1) as a fitness value for evolution and 2) as an updated mechanism for an *archive* of points maintained during search. That is, parent and/or survival selection can be based on maximizing sparseness, and novelty search adds individuals to the archive only if the sparseness over the archive is above a certain threshold, ρ_{min} .

During search, the archive continues to expand as the search progresses, and the population is evolved toward new, novel search points. It is the assumption that search occurs at the level of these candidate points and the fact that the archive continues to expand that gives rise to previous assertions that novelty search “diverges”.

Archive-Based Searches

In a typical evolutionary algorithm, individuals in the population represent candidate solutions to some problem, and the search space is the space of potential candidate solutions to that problem. Given this, it’s easy to see why novelty search *appears* to be objectiveless: the algorithms ignore the optimization objective associated with that candidate solution, and one imagines search diverging in the sense that increasingly different candidates are progressively proposed in an ever-growing archive.

Still, there are other ways to understand search. Multi-objective optimization find approximations of Pareto non-dominating *sets* for some solution space (Seada and Deb 2018). *Cooptimization* (e.g., coevolutionary algorithms) are supersets of multiobjective optimization problems, in which simultaneously identifies the set of objectives and produces a set obeying some solution concept — typically also Pareto-based (Popovici et al. 2012).

Novelty search also searches a space of *novel archives*, not a space of candidate solutions. The true goal of this search is to *cover* a solution space as much as possible while remaining as efficiently *packed* as possible. That is, it develops a set that spreads out in the space and also keeps points in the set far apart.

The Population-Based Simple Novelty Search Evolutionary Algorithm

In a PBSNSEA, the parent population contains the current state of the exploration aspects of the search, while the archive contains the potential solution set at that moment of the search. Each generation, λ children are produced from

the μ parents, and the children are added to the archive if they meet the sparseness criterion as computed over the archive. Additionally, sparseness of the children with respect to the child population itself is computed, and that the μ individuals with the highest sparseness values are selected to be in the parent set for the next generation. This is repeated until some termination criteria is met. For this paper, we use $k = 3$ for the sparseness computation in all cases for simplicity. The termination criterion is simple a maximum number of generations (500 in all cases here). We will consider a real-valued representations, where $d = 5$ dimensions. Gene values are unbounded, and mutation works by independently adding a value drawn from $\mathcal{N}(0, \sigma)$ — so-called *Gaussian mutation*. This allows the possibility for the archive to grow in an unbounded way by adding individuals that are “novel” in the sense that they are distant from previous individuals. Distance is computed using the L_2 norm, Euclidean distance.

Note that our distance calculations are in the genotype space, whereas traditionally novelty search computes distance in “behavior space”. However, this is not an important distinction in this analysis since we are only evaluating the claim about whether an objective exists, not what it looks like for a given problem instance.

In all cases, experiments were replicated for 50 independent trials and run to 500 generations. All experiments in this paper use $\sigma = 0.1$ and $\rho_{min} = 0.45$.

Packing and Covering

To develop bounds on how efficiently distance-based, lazy methods like k -nearest neighbor algorithms are able to develop hypotheses for a given space, the theory community for that field have developed several formal definitions (Clarkson 1999).

The first idea is the ϵ -cover of a set: How well does the set cover some space?

Definition 1 An ϵ -cover of some space $Z = \langle U, \delta \rangle$, where $\delta : U \times U \mapsto \mathbb{R}$ is a distance measure over U , is a set $A \subset U$ such that $\forall x \in U, \exists a \in A$ with $\delta(x, a) < \epsilon$.

The second idea is ϵ -packing of a set: How efficient is the distribution of points in the set?

Definition 2 Given the space $Z = \langle U, \delta \rangle$, where $\delta : U \times U \mapsto \mathbb{R}$ is a distance measure over U , a set $A \subset U$ is an ϵ -packing iff $\delta(a, b) > 2\epsilon \forall a, b \in A$, where $a \neq b$.

Put more simply, a set that *covers* a space well is one that is spread out over that space such that no point in the space is too far from at least one point in the set, while a set that is *packed* well is one in which points inside the set aren’t too close together. Note that for cover, smaller is better (points in U are closer to points in A); however, for packing, larger is better (points in A are further apart).

Definition 3 An ϵ -net $A \subset U$ is a set that is an ϵ -cover of U and an $(\epsilon/2)$ -packing.

In k -NN theory, an ϵ -net is optimal: A set that is both efficiently packed and covers the space well.

We adopt the estimation measures for these from (Wiegand 2020). The packing estimator is simply half the maximum pairwise distance in the archive. Packing, thus, starts small and grows as points are added to the archive.

Directly computing cover is computationally infeasible. Instead, we approximate this as follows. First, we assume that the search will (with high probability) remain within the bound $\pm\sigma \cdot \text{maxGen}$ in all dimensions. We confirmed that though any 5D point was *possible* in principle, in all runs of all experimental groups, none were generated outside that region. Second, we sample points uniformly at random from inside that bounded hypercube. Finally, we find the closest point in the archive to each sample point. The maximum of such distances is reported as our estimate for ϵ -cover. Cover estimates will tend to start large, then drop as it becomes increasingly less probable that the algorithm will select new points that fill in the gaps of the space searched so far.

Results

The SNSEA variant from (Wiegand 2020) does not have a population in the sense that traditional novelty search has. It instead uses the archive as the population. Seen this way, the SNSEA is most comparable to a $(\mu + 1)$ -EA, where parent selection occurs randomly from the population (archive) and truncation survival selection is being used. Of course, it differs from this in that the fitness of an individual depends on the current population (archive), which grows.

So the algorithm *does* have selective pressure and progresses in a particular direction; it is not random search. Generally, the algorithm gradually grows the size of the archive, adding new points as they are discovered. The SNSEA will steadily increase packing and decrease cover.

Our (μ, λ) population-based variant instead maintains a separate population and archive. This permits the search dynamics to work independently of the archive state. In such a case, it is better to think of the archive as the “*saved best solution found so far*.” Still, since the archive update mechanism makes use of the sparseness criteria, the algorithm’s updates of the archive are not random or arbitrary. Again, the update process itself applies pressure to reduce packing (make the archive more efficient) and reduce the cover (finding new points). The key question is whether the conflation of population dynamics and archive update in (Wiegand 2020) is the reason convergence was observed.

The PSNSEA Converges

Ideally, a “diverging” novelty search will continue to find increasingly more efficient packings of the existing archive while developing an archive whose coverage of the space also improves. When cover improves and packing does not, this suggests that novel points are being added but the archive is missing large swaths of the space — novelty is being discovered but inefficiently. When packing improves and cover does not, this is a very different matter. The algorithm is inserting more points inside the convex hull of the archive but not really discovering anything new (as defined by the sparseness metric). As long as packing remains below cover, true novel discovery is occurring.

Unbounded 5D Euclidean Space

Gaussian Mutation $\sigma = 0.1$

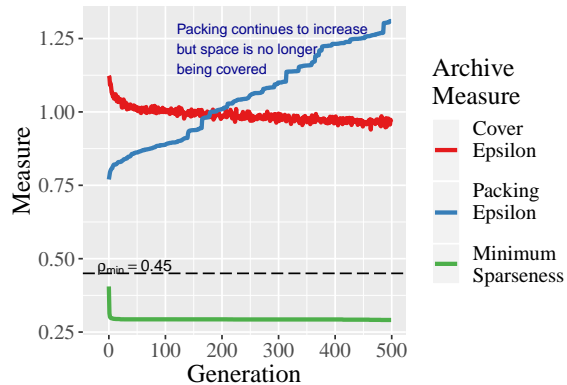


Figure 1: $(2, 8)$ -PSNSEA applied to an unbounded real-valued space with $n = 5$, $\rho_{min} = 0.45$, and $\sigma = 0.1$ averaged over 50 independent trials.

Figure 1 illustrates an example in a population-based approach where packing increases and cover converges. It is clear that it is driven by the packing and cover objectives described above. Indeed, the $(2, 8)$ -PSNSEA converges much faster than non-population approaches show in (?).

Populations Can Speed Up Convergence

To test whether or not populations can solve the problem of our simple novelty search EA converging, which considered several different population configurations for μ and λ , specifically, we ran 50 independent trials of eight experimental groups: $(2, 8)$; $(2, 16)$; $(2, 32)$; $(4, 16)$; $(4, 32)$; $(8, 32)$; $(16, 32)$ no-population. The goal is to see how long it takes before convergence occurs in the sense that coverage has dropped below packing. Our experiments were run for 500 generations, and in trials of all population-based experimental groups this occurred within that time. The only experimental group where packing does not overtake cover were the control group trials where children are selected directly from the archive. Here, we run a LOESS regression curve for both packing and coverage and record when the two curves *would have* intersected with 95% confidence.

Novelty search is searching *archive space* rather than the underlying Euclidean space, and that the population represents part of this potential archive space. In that sense, it makes the most sense to examine coverage and packing of the archive by generation. As such, Figure 2 shows the average number of generations it took for the different groups to converge. The control group (no-population) actually took longer than any population-based group.

Discussion

Novelty search can be a useful tool for discovery of complex objects in open-ended spaces by leveraging three key pieces: a sparseness metric and archive, generative representations,

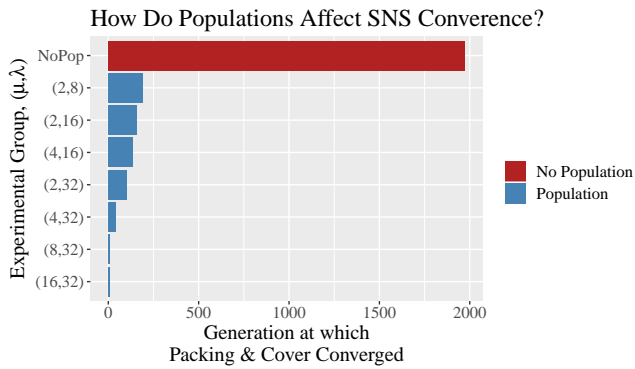


Figure 2: PSNSEAs applied unbounded 5D Euclidean space under different population conditions. Each experimental group was run for 50 independent trials. The bar length represents the *generation* in which packing overtakes cover and cover has converged.

and distance measures in the actual solution space (e.g., behaviors) rather than genotype space. This paper expands on recent efforts to examine novelty search at a more foundational level by concentrating on the first piece (sparseness and archive) to evaluate claims that novelty search is “divergent” and “objectiveless”. Previous work has concluded that these claims aren’t always valid — SNSEAs can converge by optimizing the archive toward an ϵ -net that covers and packs a space in the sense defined by k -NN theory. Unfortunately, this was based on unrealistic simplifications where SNSEAs draw parent individuals directly from the archive rather than using a population, which is not how most novelty search methods work.

In this paper, we consider whether adding a population changes things. Do SNSEAs with populations diverge? We show that even when the PSNSEA uses a child and parent population, it can converge in terms of cover measures over the space. In fact, we provide evidence that certain population configurations can actually *accelerate* convergence because populations are not necessarily good representatives of the leading edge of an archive.

It may seem counter-intuitive that drawing a single child from the archive can improve novelty search — surely drawing a single parent, producing one child, then updating the archive represents a significant diversity loss? But this counter-intuition again mistakes the fundamental unit of search in novelty search, which *not* the genotype space but the *archive* space. Each child produced represents a potential *part* of how the solution can be grown, and as such it needs to be as representative of the leading edge of the archive as possible assuming the goal is to grow the cover of the archive. Drawing parents directly from the archive ensures the archive is at least minimally represented, whereas maintaining a separate population runs the risk that the cover of the archive and the cover of the population are very different.

Still, selecting random individuals from inside an archive is inefficient. As the archive grows, the ratio of the leading edge to its overall volume will reduce, making it harder

and harder to sample from areas that are likely to produce “novel” solutions. This problem gets worse as dimensionality is increased, of course. Taken together, these observations suggest potential new ideas for maintaining the novelty gradient: Bias parent selection to favor the leading edge of the archive. This may involve convex hull operations, or selection methods for population-based approaches that forestall packing rather than increase it.

References

- Clarkson, K. L. 1999. Nearest neighbor queries in metric spaces. *Discrete & Computational Geometry* 22(1):63–93.
- Dreher, T. 2014. *History of Computer Art*. chapter IV.3.
- Fontaine, M. C.; Togelius, J.; Nikolaidis, S.; and Hoover, A. K. 2019. Covariance matrix adaptation for the rapid illumination of behavior space. In *Proceedings of the 2019 IEEE Conference on Games*.
- Goldberg, D. 2002. *The Design of Innovation: Lessons from and for Competent Genetic Algorithms*. Springer.
- Huang, P.; Lehman, J.; Mok, A. K.; Miikkulainen, R.; and Sentis, L. 2014. Grasping novel objects with a dexterous robotic hand through neuroevolution. In *2014 IEEE Symposium on Computational Intelligence in Control and Automation*, 125–132.
- Koza, J. 1992. *Genetic Programming*. MIT Press.
- Lehman, J., and Miikkulainen, R. 2015. Enhancing divergent search through extinction events. In *Proc. of the Genetic and Evolutionary Computation Conference*, 951–958.
- Lehman, J., and Stanley, K. O. 2011. Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary Computation* 19(2):189–223.
- Lehman, J.; Wilder, B.; and Stanley, K. O. 2016. On the critical role of divergent selection in evolvability. *Frontiers in Robotics and AI*.
- Morse, G.; Risi, S.; Snyder, C. R.; and Stanley, K. O. 2013. Single-unit pattern generators for quadruped locomotion. In *Proceedings of the 15th Genetic and Evolutionary Computation Conference, GECCO ’13*, 719–726.
- Popovici, E.; Bucci, A.; Wiegand, R. P.; and de Jong, E. D. 2012. Coevolutionary principles. In *Handbook of Natural Computing*. 987–1033.
- Seada, H., and Deb, K. 2018. Non-dominated sorting based multi/many-objective optimization: Two decades of research and application. In *Multi-Objective Optimization - Evolutionary to Hybrid Framework*. 1–24.
- Stanley, K. O., and Lehman, J. 2015. *Why Greatness Cannot Be Planned - The Myth of the Objective*. Springer.
- Szerlip, P. A.; Morse, G.; Pugh, J. K.; and Stanley, K. O. 2015. Unsupervised feature learning through divergent discriminative feature accumulation. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2979–2985.
- Wiegand, R. 2020. The objective of simple novelty search. In *Proceedings from the 2020 Florida Artificial Intelligence Research Symposium Conference*.