Objective vs. Divergent Search: A Reference for a Small Demo

Gabriel E. Lima

October 15, 2025

Abstract

This document is a self-contained reference for a small, visual Genetic Algorithm (GA) simulation contrasting three search strategies: *objective* (fitness-driven), *novelty*, and *surprise*. Simulation code available at: https://github.com/Lima001/PPGInf-Coursework/tree/main/Bio-inspired%20Computing/Novelty-Surprise%20Search

1 Introduction

This simulation focuses on a simple pedagogical scenario: agents try to navigate from a fixed *Start* to a fixed *Goal* inside a bounded 2D world. A deceptive *Trap* sits between the start and goal, providing a local reward that acts as an attractive lure. The experiment explores three alternative search strategies:

- **Objective search**: selects individuals based on how much they maximize a task reward (closeness to the goal).
- **Novelty search**: selects individuals based on how different their behavior is from previously seen behaviors.
- **Surprise search**: selects individuals that deviate from *recent* behavioral trends.

Objective search can efficiently advance towards the goal but may become trapped by deceptive rewards. Novelty and surprise are intrinsic motivations that encourage exploration, often enabling an escape from such deceptive traps.

2 Simulation Environment

The environment is a square region where agents, represented as points, execute short plans (genomes). The key elements are:

- **Start**: A common initial position for all agents.
- **Goal**: The desired destination.
- **Trap**: A region that gives a local reward but is not the goal.
- **Trajectory**: The sequence of positions produced by executing a genome.

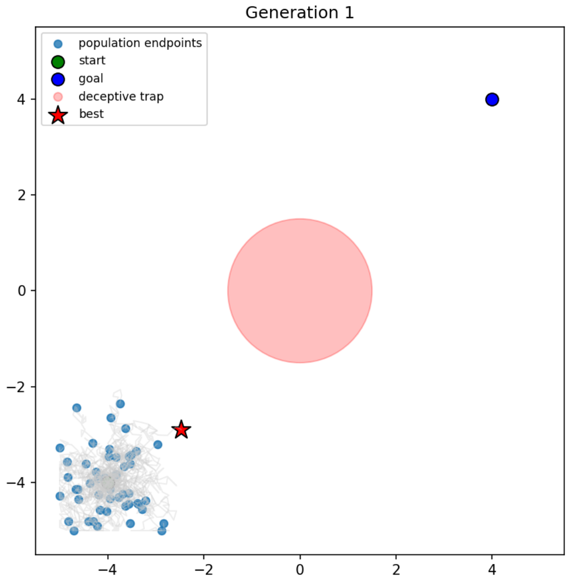

Figure 1: Visual representation of the simulation environment, showing the Start and Goal points, the central deceptive trap, and example trajectories.

3 The Evolutionary Algorithm

3.1 Genome Representation and Trajectory

A genome encodes a short sequence of 2D movement commands. When executed from the start position, it yields a trajectory – a time-ordered list of positions. Let $T \in \mathbb{N}$ be the number of timesteps. A genome is defined as:

$$\mathbf{g} = (a_1, a_2, \ldots, a_T), \qquad a_t \in \mathbb{R}^2.$$

Starting at s_0, the trajectory is generated by the rule:

$$s_{t+1} = \text{clip}\left(s_t + \alpha\, a_t\right), \quad t = 0, \ldots, T-1,$$

where $\alpha > 0$ is a step-size multiplier and clip restricts the agent's coordinates to remain within the world boundaries.

3.2 Genetic Operators

The algorithm uses the following genetic operators:

- **Tournament Selection:** A subset S of k individuals is sampled uniformly without replacement; the one with the highest score is selected: $\arg\max_{i \in S} \text{score}_i$.
- **One-Point Crossover:** A crossover point $c \in \{1, \ldots, 2T-1\}$ is chosen. The first child receives the first c genes from parent A and the remainder from parent B.
- **Per-Gene Mutation:** For each gene, a random perturbation is applied with a given probability p:

$$g'_j = \begin{cases} g_j + \epsilon_j, & \epsilon_j \sim \mathcal{N}(0, \sigma^2), \text{ with probability } p, \\ g_j, & \text{otherwise.} \end{cases}$$

4 Objective Search

Objective search uses a task-specific reward (fitness) as its selection signal. In this demo, fitness rewards closeness to the goal, while the trap provides a local bonus that can create a deceptive local optimum.

For a trajectory $S = (s_0, \dots, s_T)$ and goal g, the base fitness is the negative distance to the goal:

$$f_{\text{base}}(S) = -\|s_T - g\|.$$

Let $\delta_{\text{trap}} = \min_t \|s_t - c_{\text{trap}}\|$. A bonus is awarded for approaching or entering the trap (of radius R):

$$b_{\text{trap}}(S) = \begin{cases} B\left(1 - \dfrac{\delta_{\text{trap}}}{R}\right), & \delta_{\text{trap}} \le R, \\ 0, & \text{otherwise.} \end{cases}$$

An additional bonus is given if the goal is reached (within radius r):

$$b_{\text{goal}}(S) = \begin{cases} G, & \|s_T - g\| < r, \\ 0, & \text{otherwise.} \end{cases}$$

The total fitness is the sum of these components:

$$f(S) = f_{\text{base}}(S) + b_{\text{trap}}(S) + b_{\text{goal}}(S).$$

Interpretation: Despite being far from the true goal (a negative base fitness), visiting the trap can produce a strongly positive total fitness, which is what makes the trap deceptive.

5 Novelty Search

Novelty search rewards behaviors that are *different* from those seen before. In this demo, a behavior is described by its *endpoint* (the final position of a trajectory). Novelty is the average distance to the k nearest neighbors in a combined set of the current population and a historical archive. This archive encourages long-term exploration.

Given endpoints $\{d_i\}_{i=1}^N$ and an archive $\mathcal{A} = \{a_j\}_{j=1}^M$, the combined pool is $S = \{d_1, \dots, d_N, a_1, \dots, a_M\}$. For an endpoint d, its novelty is calculated from the k smallest distances $r_{(1)} \le \cdots \le r_{(k)}$ to other points in S:

$$\text{novelty}(d) = \frac{1}{k} \sum_{i=1}^{k} r_{(i)}.$$

Each generation, the most novel endpoints are added to the FIFO archive.

6 Surprise Search

Surprise rewards behaviors that *deviate from recent trends*. While novelty measures difference against all past behaviors (via the archive), surprise focuses on how unexpected a behavior is compared to only the last few generations. This is inherently temporal. In this demo, flattened trajectories are used for simplicity to capture the full path shape and timing.

Let the current generation's flattened trajectories be $D_{\text{cur}} = \{d_i^{(0)}\}$. The history pool H contains trajectories from the last m generations:

$$H = \bigcup_{t=1}^{m} D^{(-t)}.$$

For a current descriptor d, we find the distances to all descriptors in H. Let the k smallest distances be $s_{(1)} \leq \cdots \leq s_{(k)}$. The surprise is:

$$\text{surprise}(d) = \frac{1}{k} \sum_{i=1}^{k} s_{(i)}.$$

If H is empty (in early generations), surprise is defined as 0.

How Surprise Differs from Novelty

Novelty, based on *endpoints*, measures "going somewhere new." Surprise, based on *flattened trajectories*, measures "moving in a new way recently." These two are not necessarily correlated; you can go to a new location with a familiar path (high novelty, low surprise) or take a novel curved path to a familiar location (low novelty, high surprise).

7 Combining Metrics: The Composite Score

Fitness, novelty, and surprise operate on different scales. To combine them fairly, each metric is normalized per generation (min-max scaling), and a weighted sum is computed. This composite score is used for selection.

Given vectors for fitness (f), novelty (n), and surprise (s) over the population, normalization is defined as:

$$\widehat{v}_i = \frac{v_i - \min_j v_j}{\max_j v_j - \min_j v_j + \varepsilon}, \quad \varepsilon \ll 1.$$

The composite score is then:

$$\text{composite}_i = w_F \widehat{f}_i + w_N \widehat{n}_i + w_S \widehat{s}_i,$$

with non-negative weights w_F, w_N, w_S chosen by the experimenter.

8 Population Diversity Metric

Diversity summarizes how spread out the population's behaviors are. A common metric is the mean pairwise Euclidean distance between all endpoints. It helps monitor whether the population is exploring or converging.

Given endpoints $p_1, \ldots, p_N$, the diversity is:

$$\text{diversity} = \frac{2}{N(N-1)} \sum_{i<j} \|p_i - p_j\|.$$

9 References

J. Lehman and K. O. Stanley. "Abandoning Objectives: Evolution through the Search for Novelty Alone". *Evolutionary Computation*, 2011.

D. Gravina, A. Liapis, G. N. Yannakakis. "Surprise Search: Beyond Objectives and Novelty". *GECCO*, 2016.

M. Mitchell. *An Introduction to Genetic Algorithms.* MIT Press, 1996.

J. Schmidhuber. "Formal Theory of Creativity, Fun, and Intrinsic Motivation". *IEEE Transactions on Autonomous Mental Development*, 2010.