Simulation Design

July 2024

1 Overview

This model aims to simulate how the possession of scientific knowledge impacts technological inventions and innovations in a general sense. It conceptualizes invention as a process of recombination. The NK fitness landscape model provides a useful framework for search processes of performance of different combinations, with scientific knowledge functioning as a map for agents to explore the hills and valleys of landscapes. In this simulation, scientific knowledge is characterized by three attributes: range, precision, and accuracy. The range is analogous to the map's size (represented by radius), precision to the map's granularity, and accuracy to the probability of correctness. The results compare performances after several periods among models with different variables.

2 Base Model

2.1 The Environment

The environment is modeled by a classical NK fitness landscape, representing the actual performance of different combinations. Each cell on the landscape symbolizes a system comprising N components, with each component interdependent on K others. Each component can be in one of two states, 0 or 1. The contributions (coefficient values) of each component and the interaction impacts are randomly drawn from the standard normal distribution.

2.2 The Representation

The representation simulates the effect of scientific knowledge, which is analogous to a map predicting the performance of different combinations. It is characterized by three attributes: radius, granularity, and inaccuracy.

- Radius (R): Centered on the initial position, the agent can predict the performance of the area within a radius of R.
- Granularity (G): Due to the abstract nature of scientific knowledge, the agent cannot know everything within the area limited by the radius.

Hence, the predicted results in the representation are average values of multiple cells, with the number of cells determined by the parameter G.

- Inaccuracy (U): The representation values are not the exact results from the actual environment but are multiplied by $1 + \epsilon$, where $\epsilon \sim \mathcal{N}(0, U)$ (each point has a different ϵ). The parameter U indicates the inaccuracy of an agent's scientific theory. The larger the variable U, the more biased the predicted results.

2.3 The Agent

In the base model, there is one agent in each landscape, representing an individual or party seeking innovation (better performance through recombination).

Decision Rules

- The agent can "see" the actual value of the nearest cells, representing local search.
- The agent has a cost function for actions, where the further the point, the larger the cost.
- The agent calculates all predicted performance values based on their representation.
- The objective function is defined as:

$$L := V_p - C \tag{1}$$

  where V_p := predicted performance (p represents predicted) and C := the cost to reach the designated point.

- The agent moves to $\arg\max L$. Despite predicting a certain point to have a great performance value, the agent might choose a point with a relatively lower performance value due to the high cost of reaching the predicted point.

2.4 Details

- Each simulation has a set of fixed parameters: (1)N; (2) K; (3) Radius R; (4) Granularity G; (5) Inaccuracy U; (6) Parameters involved in the objective function;
- The environment is generated based on N and K, and the agent is randomly positioned on the landscape.
- The agent's representation is determined by R, G, and U, based on which the objective function is computed. The agent moves to a random point within the area with the highest objective function value.

- The agent accesses the actual performance value of that point. The result of this simulation period is the actual value minus the cost of the agent's movement.
- These steps are executed iteratively, with a default of 10 iterations.
- The final performance of this simulation is recorded as the average result of these iterations.

Meanwhile, a model simulating the process without scientific knowledge accompanies this, involving a fixed probability of long jumps.

3 Potential Extensions

- **Competition**: Introducing competitive scenarios among multiple agents.
- **More detailed cost function**: Including factors such as the maintenance of previous knowledge and time penalties.