CITS4403
Computational Modelling

Week 10 Lecture: Evolution

Dr. Siwen Luo

Semester 2, 2024
School of Computer Science,
University of Western Australia

W10 Lecture: Evolution

1. Theory of Evolution
2. Features to Simulate Evolution
3. Fitness Landscape
4. Simulation
   - Evidence of Evolution
   - Differential Survival
   - Mutation

THE UNIVERSITY OF WESTERN AUSTRALIA

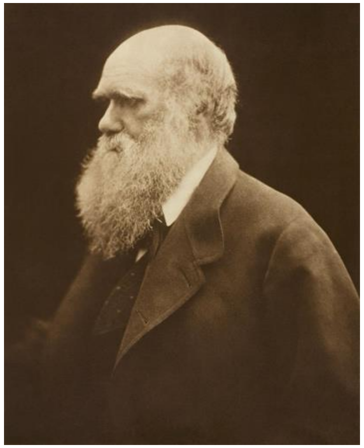
Natural Selection

Natural selection is a process in which inherited variations between individuals cause differences in survival and reproduction.

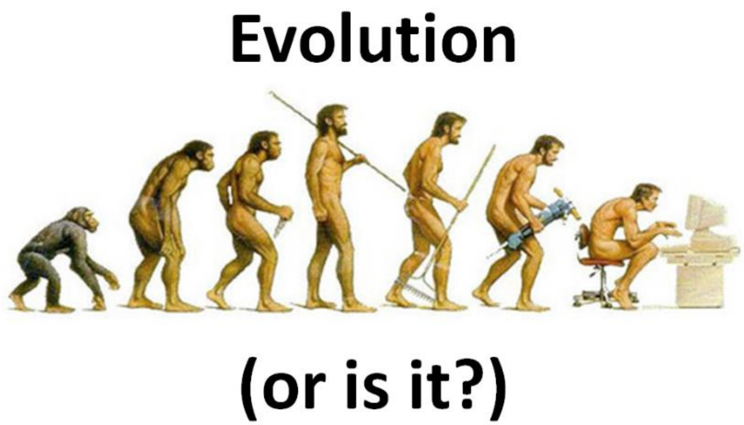
Theory of Evolution (Darwinian theory)

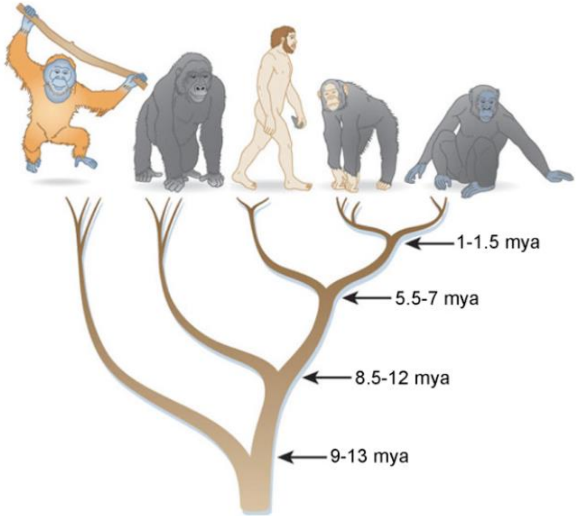
All species of organisms arise and develop through the natural selection of small, inherited variations that increase the individual's ability to ***compete, survive, and reproduce.***

Evolution holds that all species are related and gradually change over generations. In a population, the genetic variations affect the physical characteristics of an organism.



Beware the linear depiction of ape to man…





Evolution is an aggregate effect over populations. It's a process of continuous branching and divergence of populations of organisms.

According to Darwin, all current organisms are equally evolved organisms and share a common ancestor that lived about 580 million years ago and each organism alive now is the most evolved of its kind.

The theory of evolution explains the phenomena in natural world:

- **Adaptation**

  Natural selection favours adaptations that best enable creatures to survive under the circumstances in which they live.

- **Increasing Diversity**

  Over time the number of species on earth has generally increased (despite periods of mass extinction).

- **Increasing Complexity**

  The history of life on earth starts with relatively simple life forms, with more complex organisms appearing later in the geological record.

Simulating Evolution

Features to produce evolution:

- **Replicator**: a population of agents that can reproduce in some way.
- **Variation**: variability in the population, that is, differences between individuals.
- **Differential Survival or reproduction**: The differences between individuals must affect their ability to survive or reproduce.

Simulating Evolution

The basic machinery we may need:

- Agents have genetic information, called their "genotype".
- A fitness value is assigned to each genotype to specify their adaptation to the current environment.
- Some agents would die over the generations and some agents are born.

Simulating Evolution

There are many modelling decisions to make:

- How to represent the genotype?
- How to setup the fitness value? How to relate it to different genotypes?
- Which agents will die over generations? How to choose them?
- How many agents will be reproduced? Which agents will be reproduced?

Fitness Landscapes

Genotype: the genetic information of each agent that to be copied when the agent replicates.

A genotype can be represented by a sequence of N binary digits (0 or 1), where N is a parameter that we can choose.

Fitness: a quantity related to the ability of an agent to survive or reproduce.

Fitness landscape: a function that maps genotype to fitness.

This fitness is the "height" of the landscape. Genotypes which are similar are said to be "close" to each other, while those that are very different are "far" from each other.

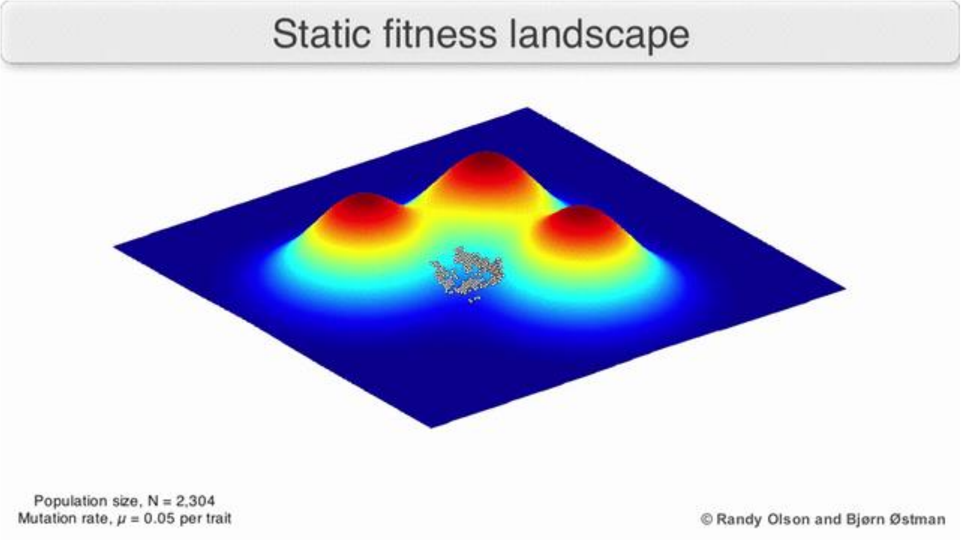
Each genotype corresponds to a location on the fitness landscape.

Fitness Landscapes

In all fitness landscapes, height is a visual ***metaphor*** for fitness.

If all genotypes have the same replication rate, on the other hand, a fitness landscape is said to be flat.

An evolving population typically climbs uphill in the fitness landscape, by a series of small genetic changes, until – in the infinite time limit – a local optimum is reached.



Fitness Landscape

The **genotype** of an agent, which corresponds to its **location** in the fitness landscape, is represented by a NumPy array of zeros and ones called loc.

```
class FitnessLandscape:

    def __init__(self, N):
        """Create a fitness landscape.

        N: number of dimensions
        """
        self.N = N
        self.set_values()

    def set_values(self):
        # random fitness value
        self.one_values = np.random.random(self.N)
        self.zero_values = np.random.random(self.N)

    def fitness(self, loc):
        """Evaluates the fitness of a location.

        loc: array of N 0s and 1s (each genotype corresponds to a location in an N-dimensional space)

        returns: float fitness
        """
        # where True, yield x (self.one_values), otherwise yield y (self.zero_values).
        fs = np.where(loc, self.one_values, self.zero_values)
        return fs.mean()
```

4 Simulation

In each step of simulation: 1) get the array containing the fitness of each agent; 2) decides which agents to die; 3) decides which agents reproduce

```python
class Simulation:

    def __init__(self, fit_land, agents):
        """Create the simulation:

        fit_land: fit_land
        num_agents: int number of agents
        agent_maker: function that makes agents
        """
        self.fit_land = fit_land
        self.agents = np.asarray(agents)
        self.instruments = []

    def step(self):
        """Simulate a time step and update the instruments.
        """
        n = len(self.agents)
        fits = self.get_fitnesses()

        # see who dies
        index_dead = self.choose_dead(fits)
        num_dead = len(index_dead)

        # replace the dead with copies of the living
        replacements = self.choose_replacements(num_dead, fits)
        self.agents[index_dead] = replacements

        # update any instruments
        self.update_instruments()
```

Differential Survival: The differences between individuals must affect their ability to survive.

```
def choose_dead(self, ps):
    """Choose which agents die in the next timestep.

    ps: probability of survival for each agent

    returns: indices of the chosen ones
    """
    n = len(self.agents)
    is_dead = np.random.random(n) < 0.1
    # np.nonzero: Return the indices of the elements that are non-zero.
    # which are True elements in is_dead array
    index_dead = np.nonzero(is_dead)[0]
    return index_dead
```

Does the differential survival implement here?

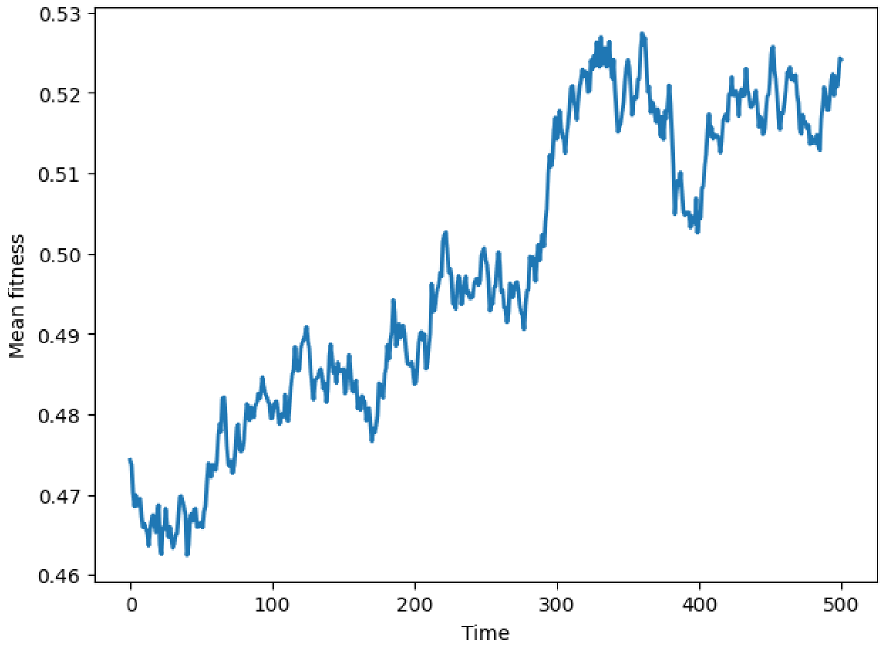
Every agent has the same probability of dying: 0.1

Evidence of Evolution

The most inclusive definition of evolution is a change in the distribution of **genotypes** in a population. Evolution is an aggregate effect: in other words, individuals don't evolve; populations do

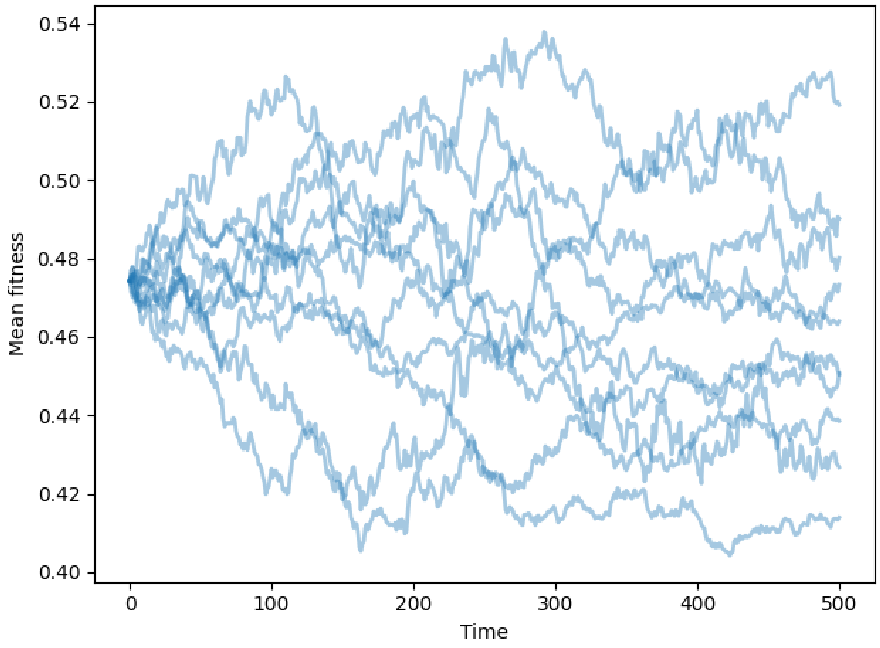
If the genotypes change, we expect their fitness to change as well. So, we will use ***changes in the distribution of fitness*** as evidence of evolution.

Distribution of Fitness



N=8, start with one agent at every location.

Distribution of Fitness

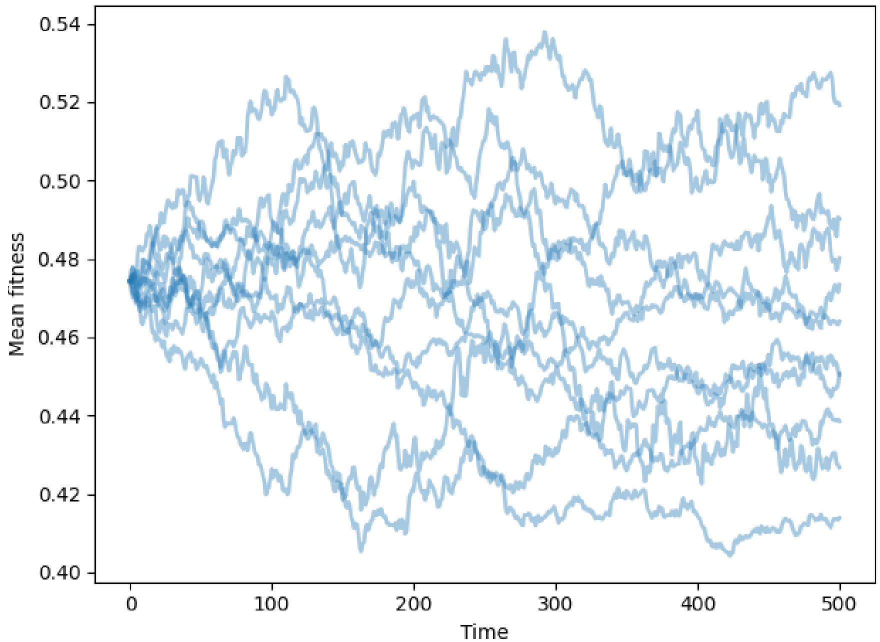


Result of running this simulation 10 times:

- N=8, start with one agent at every location.

THE UNIVERSITY OF WESTERN AUSTRALIA

Evidence of Evolution



Does this kind of evolution explain how biological species change over time, or how new species appear?

1) Adaptation? 2) Increasing diversity? 3) Increasing complexity?

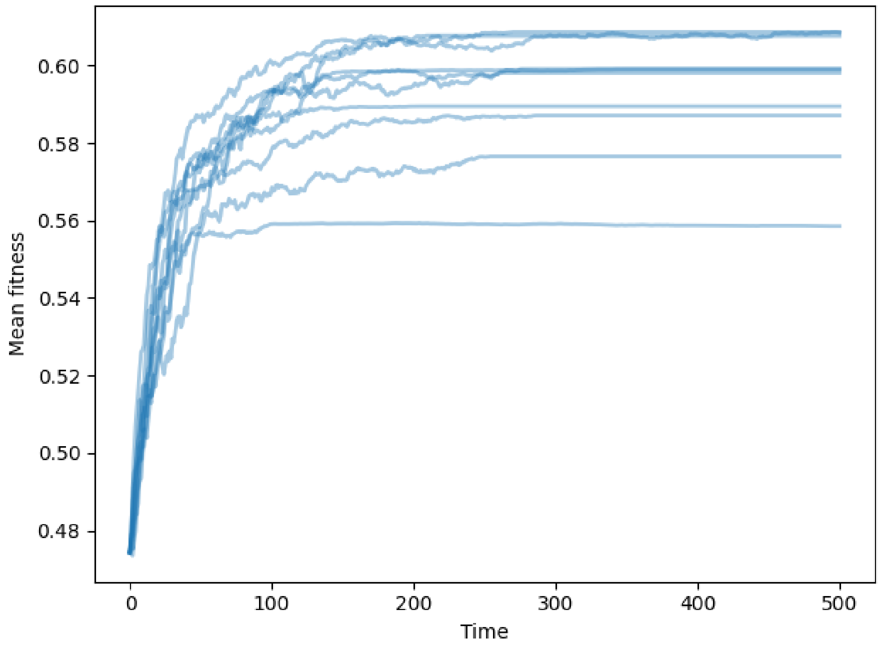
Add Differential Survival

To enable differential survival, make the probability that an agent survives each time step depend on its **fitness**.

```
def choose_dead(self, fits):
    n = len(self.agents)
    is_dead = np.random.random(n) > fits
    index_dead = np.nonzero(is_dead)[0]
    return index_dead
```

Agents with low fitness are more likely to die, agents with high fitness are more likely to survive long enough to reproduce.

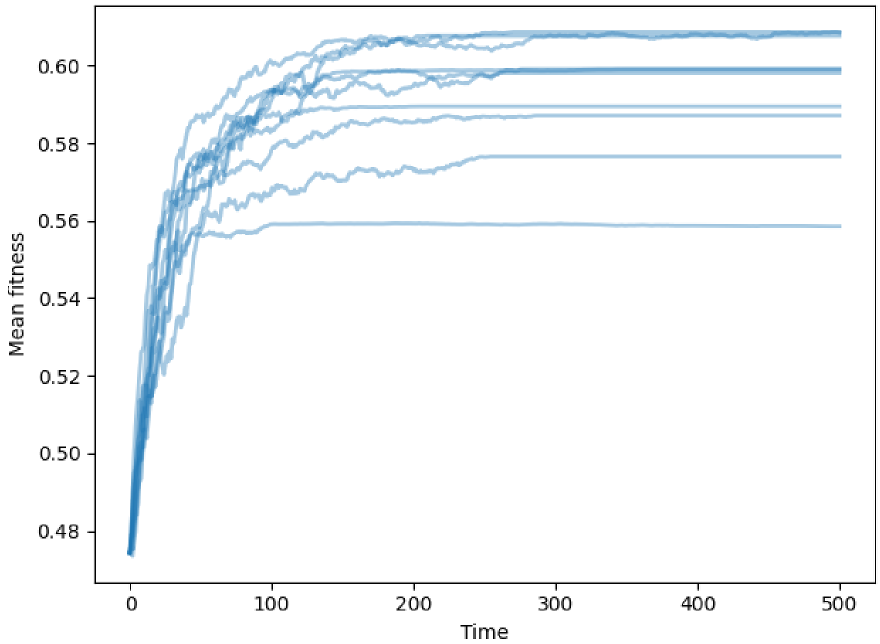
Distribution of Fitness



Result of running this simulation 10 times with differential survival:

- N=8, start with one agent at every location.

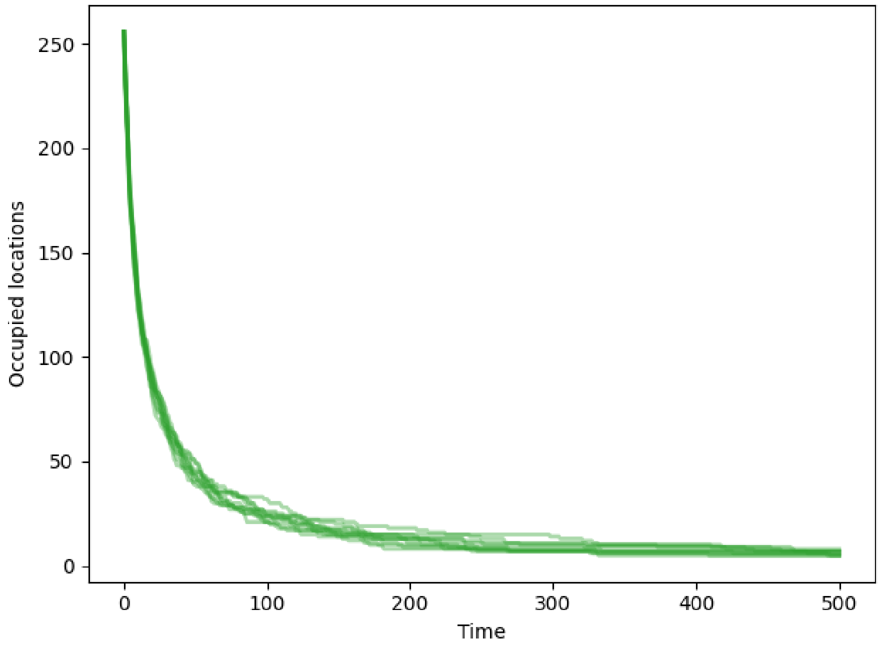
Distribution of Fitness



This simulation starts to explain **adaptation**:

- ~~*increasing fitness means that the species is getting better at surviving in its environment.*~~

Distribution of Occupied Locations



If we count the number of occupied locations on fitness landscape, we can see the number of occupied locations decreases over time, so **this model does not explain increasing diversity** at all.

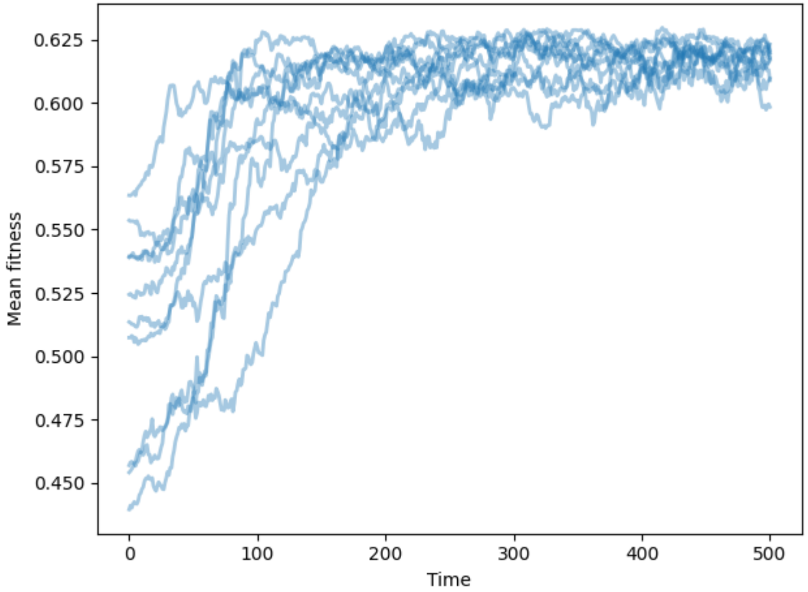
Mutation

With perfect copying in our model, we never see increasing diversity. But if we add mutation, along with differential survival or reproduction, we get a step closer to understanding evolution in nature.

```python
class Mutant(Agent):

    def copy(self, prob_mutate=0.05):
        if np.random.random() > prob_mutate:
            loc = self.loc.copy()
        else:
            direction = np.random.randint(self.fit_land.N)
            loc = self.mutate(direction)
        return Mutant(loc, self.fit_land)

    def mutate(self, direction):
        """Computes the location in the given direction.

        Result differs from the current location along the given axis.

        direction: int index from 0 to N-1

        returns: new array of N 0s and 1s
        """
        new_loc = self.loc.copy()
        new_loc[direction] ^= 1
        return new_loc
```

Distribution of Fitness

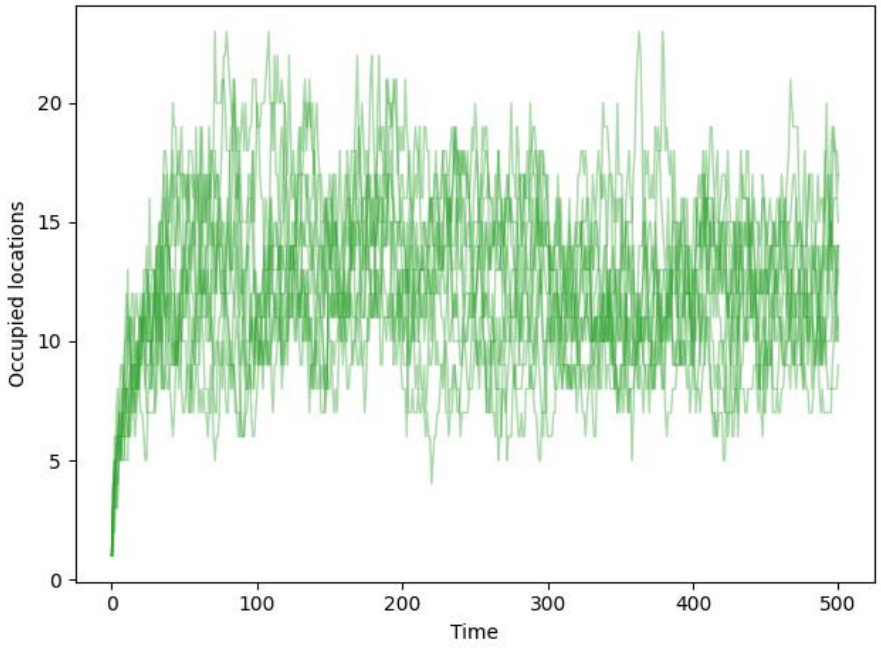
With **mutation**, we can start with the minimum variability -- all agents start at the same location.



Result of running this simulation 10 times with differential survival and mutation:

- N=8, start with 100 agents at the same location.

Distribution of Occupied Locations



Number of occupied locations over time for 10 simulations with mutation and differential survival and reproduction.

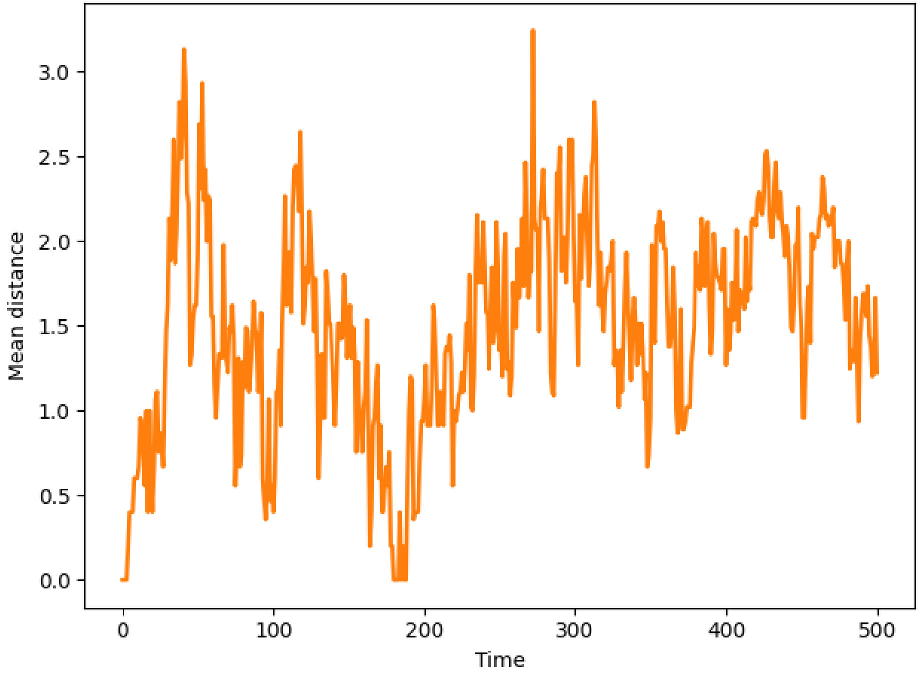
Speciation

Generally, a population is considered a species if their genotypes form a cluster, that is, if the genetic differences within the population are small compared to the differences between populations.

Since locations are represented with arrays of bits, we'll define distance as the number of bits that differ between locations. **To quantify the dispersion of a population, we can compute the mean of the distances between pairs of agents.**

THE UNIVERSITY OF WESTERN AUSTRALIA

Quantify Speciation



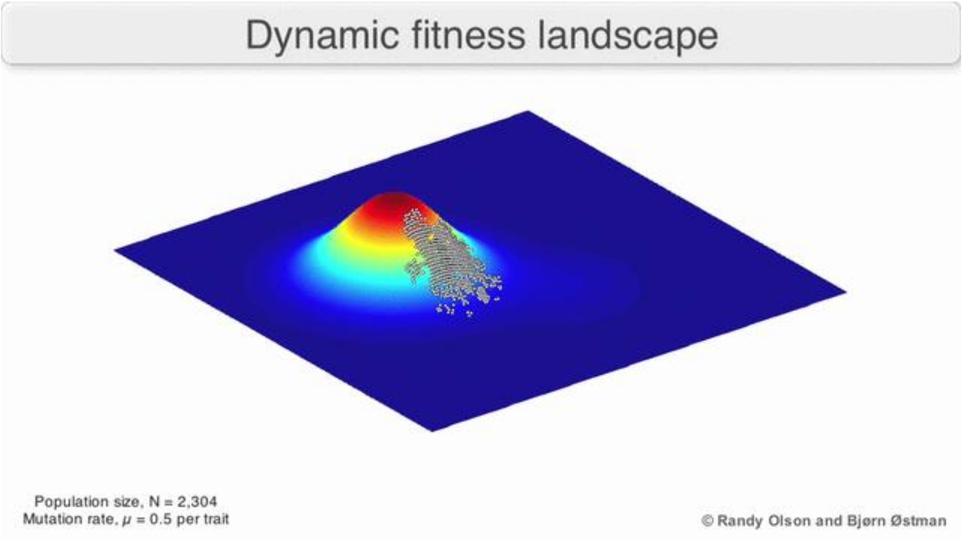
N=8, start with 100 agents at the same location, *with differential survival and mutation.*

As mutations occur, mean distance increases, reaching a maximum while the population migrates across the landscape.

Change Environment

Suppose we either change the environment, some features that increased fitness in the old environment might decrease it in the new environment, and vice versa.

We can run the simulation by change the fitness landscape after a certain steps of simulation.



Some final thoughts…

This model is not meant to be realistic.

Rather, it is a demonstration that the features of the model are ***sufficient*** to produce the behaviour we are trying to explain.

This doesn't prove that evolution in nature is caused by these mechanisms **alone**. But it is reasonable to think that they at least contribute to natural evolution.

Similarly, it doesn't prove that these mechanisms always cause evolution. But the results are robust, i.e. in almost any model that includes these features — ***imperfect replicators, variability, and differential reproduction*** — evolution happens.