CO3091 - Computational Intelligence and Software Engineering

Lecture 04


# Evolutionary Algorithms — Part I

Leandro L. Minku

# Overview

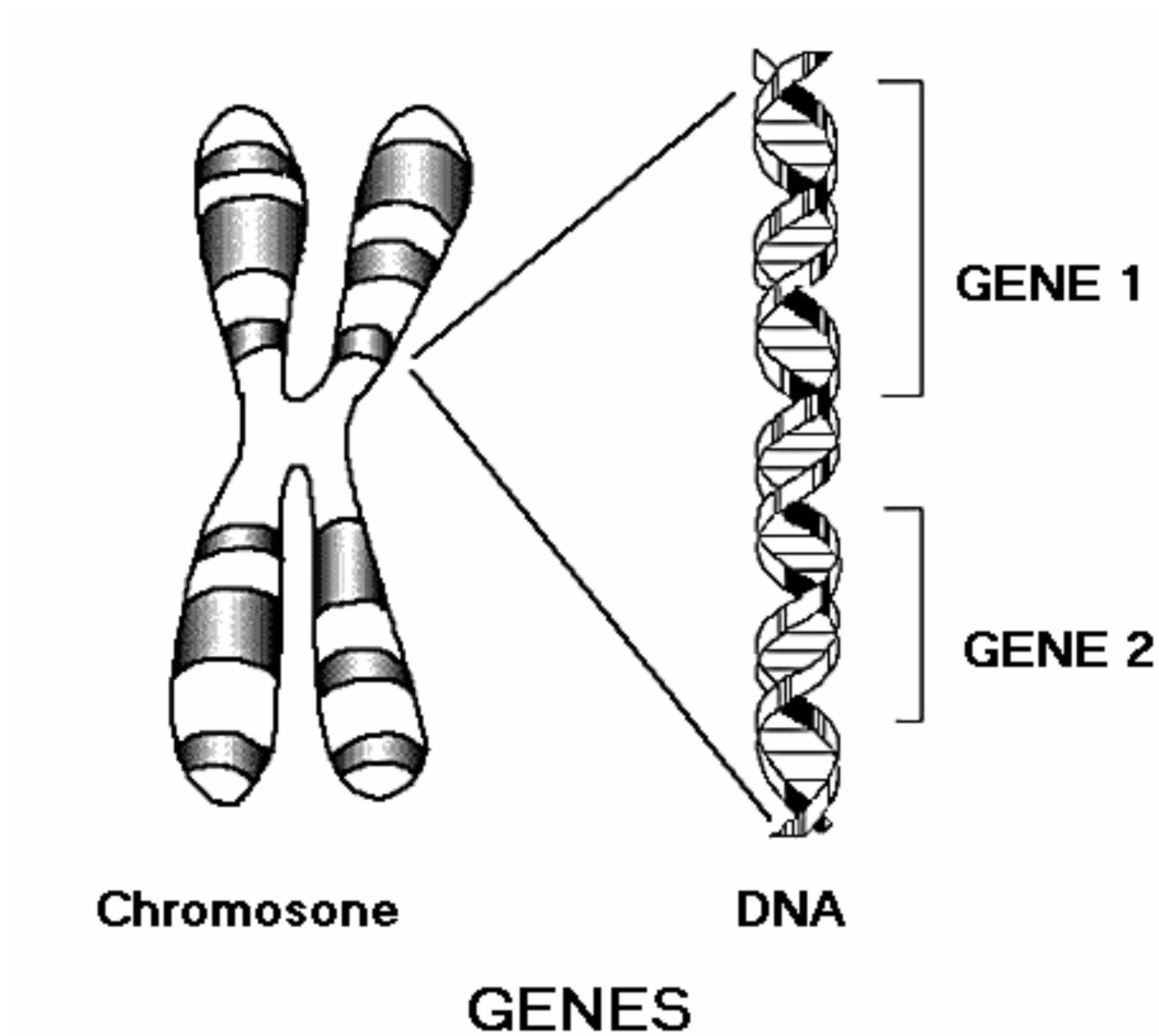
- Some concepts: genes, genotypes, phenotypes, inheritance, etc.
- What is evolution?
- How does evolution occur?
- What is the result of evolution?
- Evolutionary algorithms.
    - To be continued in the next lecture.

# Announcements

- Next class will be part II of the current lecture.
- Problem class to discuss remaining exercises from surgery I on Thursday.
- Friday lab sessions: only in weeks 12, 15 and 17.
- Wednesday lab session: in week 13.

# Some Genetics Concepts

- In the natural world:
  - DNA: chemical compound containing the instructions needed for developing organisms activities.
  - Gene: region of DNA that influences a particular characteristic in an individual.
  - Allele: alternative form of a gene.
  - Chromosomes: pack a molecule of DNA.




https://www2.warwick.ac.uk/fac/sci/maths/research/events/2008_2009/workshops/isscngc/genes.gif

# Genotypes and Phenotypes

Genotype: complete set of DNA.

# Genotypes and Phenotypes

Phenotype: observed characteristics.

# Evolution

Evolution is the change in the **inherited** traits of a **population** of organisms **through successive generations**.

- Only occurs when there is a **change in allele frequency** within a population over time.
  - E.g.: most beetles in a population are green, and very few are brown.
  - Some generations later, most beetles are brown, and few are green.
  - The frequency of the alleles for brown / green became higher / lower.

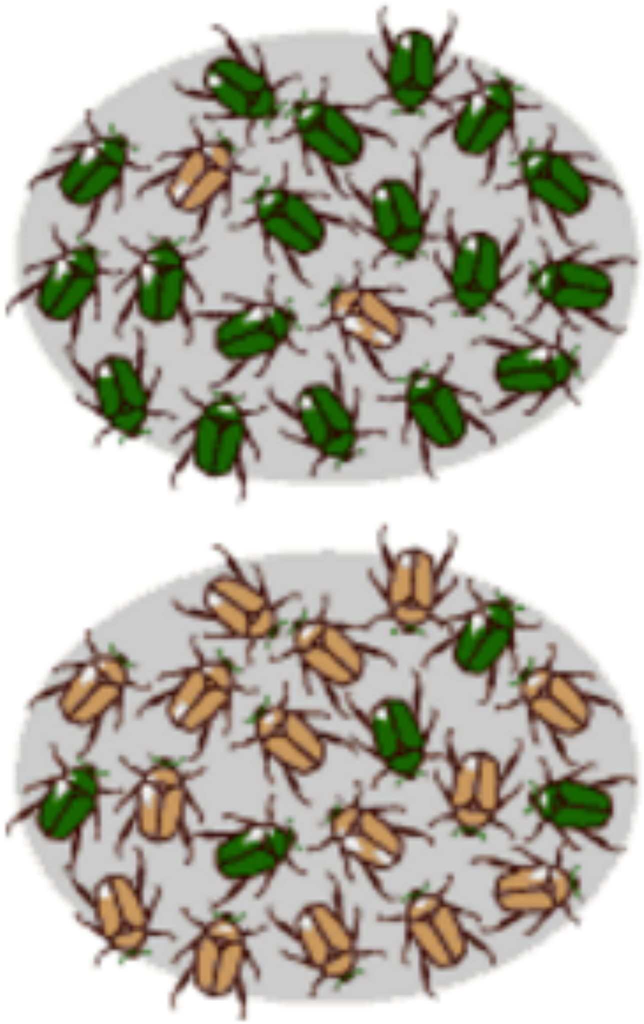


Image from: http://evolution.berkeley.edu/evolibrary/article/0_0_0/evo_15

# Fundamental Forces of Evolution

- Genetic variation
- Natural selection

# Genetic Variation

- Genetic variation: some individuals are genetically different from others.
    - Sexual reproduction: can introduce new combinations of genes into a population.

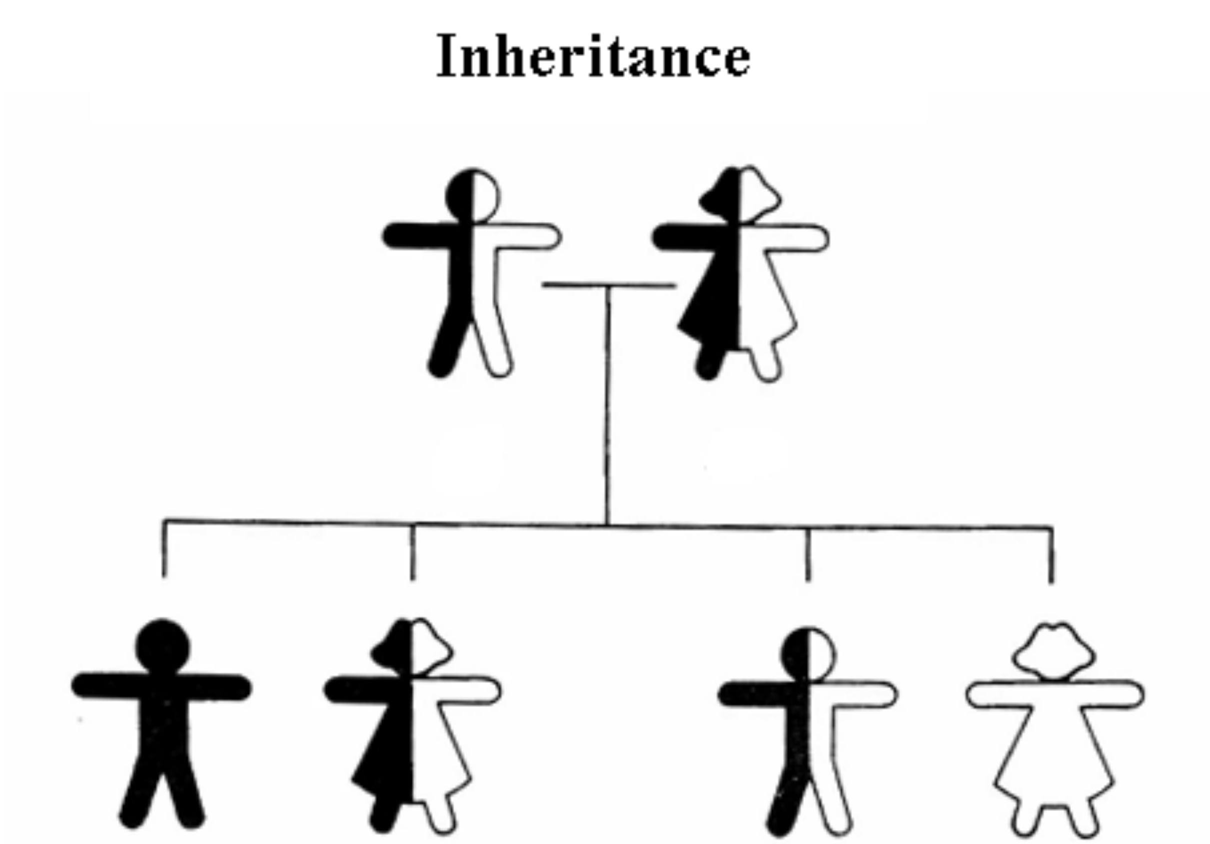
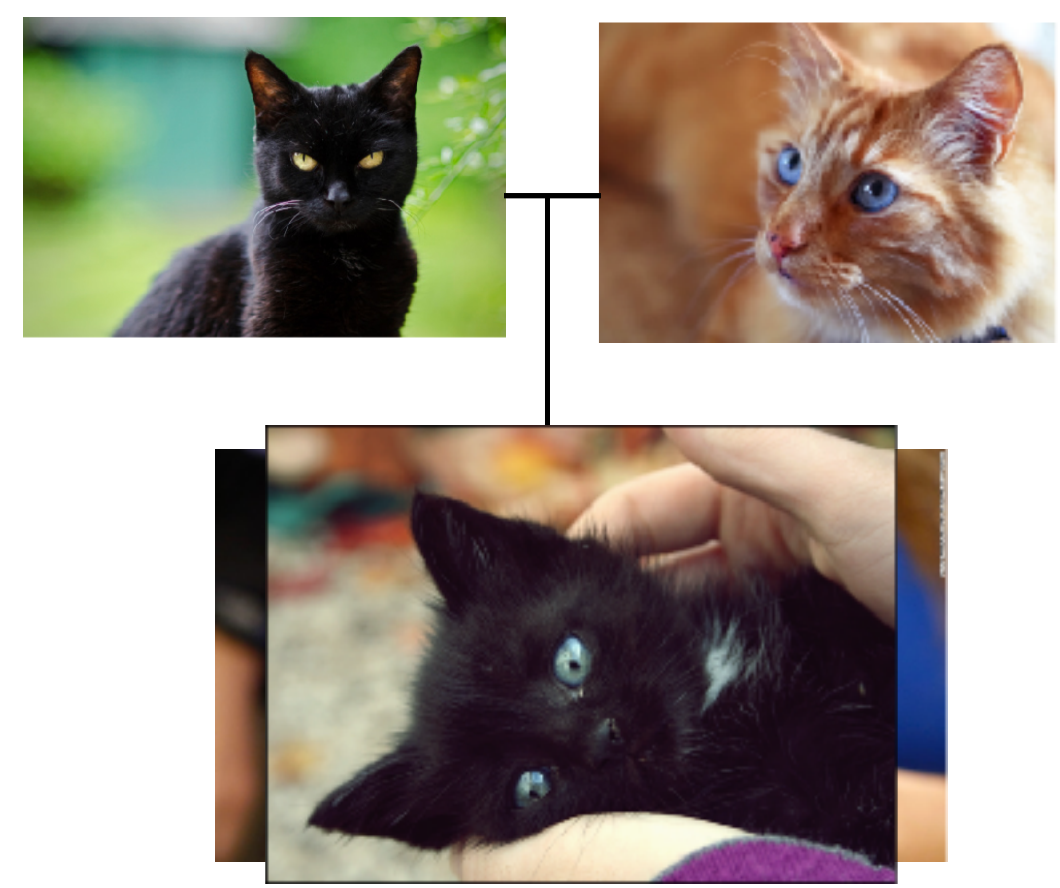



Image from: http://faculty.ccp.edu/faculty/wberman/bio106/images/inheritance.gif



# Genetic Variation

- Genetic variation: some individuals are genetically different from others.
    - Mutation: natural process that alters a DNA sequence. It allows variations not present in the parents.

# Genetic Variation

- Genetic variation: some individuals are genetically different from others.
    - Mutation: natural process that alters a DNA sequence. It allows variations not present in the parents.

Image from: https://upload.wikimedia.org/wikipedia/commons/b/b5/Periophthalmus_gracilis.jpg

# Fundamental Forces of Evolution

- Genetic variation
- **Natural selection**

# Natural Selection

Natural selection is the **differential survival and reproduction** of individuals due to **differences in phenotype**.

- E.g., green beetles may be easier for birds to spot and therefore eat.
- Then, brown beetles will be more likely to survive to produce offspring.

Image from: http://evolution.berkeley.edu/evolibrary/article/evo_25

# Natural Selection

Natural selection is the **differential survival and reproduction** of individuals due to **differences in phenotype**.

Image from: http://www.fazendavisconde.com.br/images/Pavao_Azul_Pavo_cristatus_Fazenda_Visconde_4.jpg

- E.g., sexual selection: male peecocks maintain elaborate tails to increase their chances with females.
- More attractive male peecocks are more likely to mate and produce offspring.

# Darwin's Theory of Evolution by Natural Selection

1. Variation in traits

3. Inheritance

2. Differential survival and reproduction

4. End result after a long enough time

Image from: http://evolution.berkeley.edu/evolibrary/article/evo_25

We refer to genotypes more likely to leave offspring for the next generation as **fitter** genotypes. Fitness depends on the environment.

# Result of Evolution by Natural Selection: Adaptation

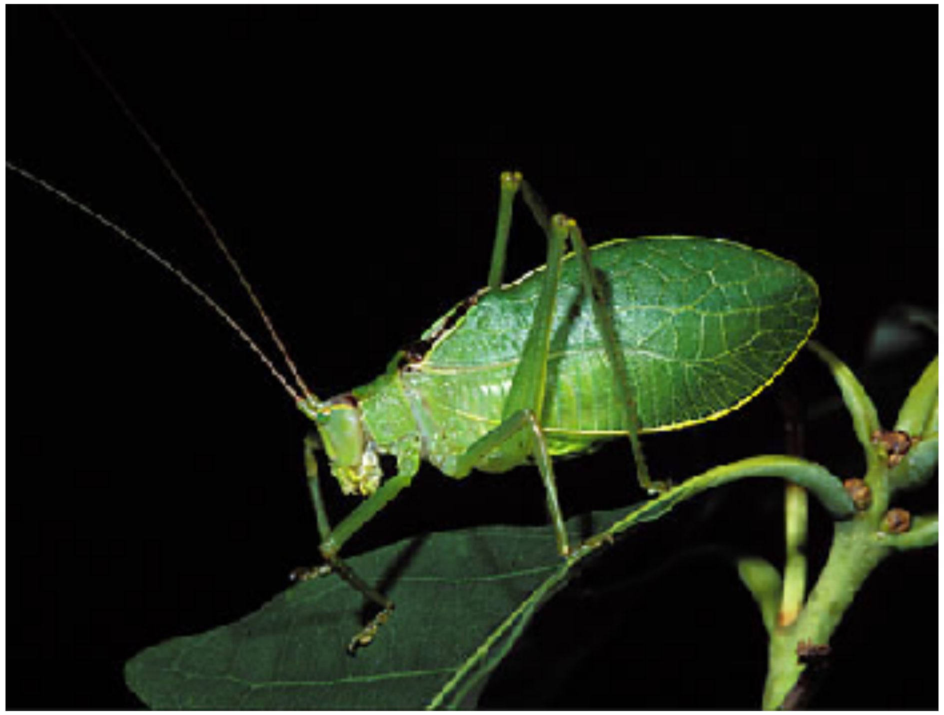


Image from: http://www.fcps.edu/islandcreekes/ecology/Insects/True%20Katydid/141pm2.jpg

Image from: http://blog.nus.edu.sg/lsm1303student2013/files/2013/03/milk_coral_snakes-1c2gopq.jpg

Image from: http://www.orpingtonfieldclub.org.uk/article002pics/wasponflyorchidbest.jpg

# Natural Evolution vs Evolutionary Algorithms (EAs)

| Natural Evolution | Evolutionary Algorithms |
|---|---|
| Fitter individuals are the ones more likely to survive and reproduce in a given environment. | Fitter solutions are the ones that are better in terms of our objective function, and thus more likely to generate new solutions. |
| After many generations, we get adaptations to the environment. | After many iterations, we get better and better solutions given our objective function. |
| Adaptation takes millions of years. | Each generation passes much quicker in a computer. |

# Evolutionary Algorithm (EA)'s Pseudocode

Evolutionary Algorithm

1. Initialise population
2. Evaluate each individual (determine their fitness)
3. Repeat (until a termination condition is satisfied)

    3.1 Select parents

    3.2 Recombine parents with probability Pc

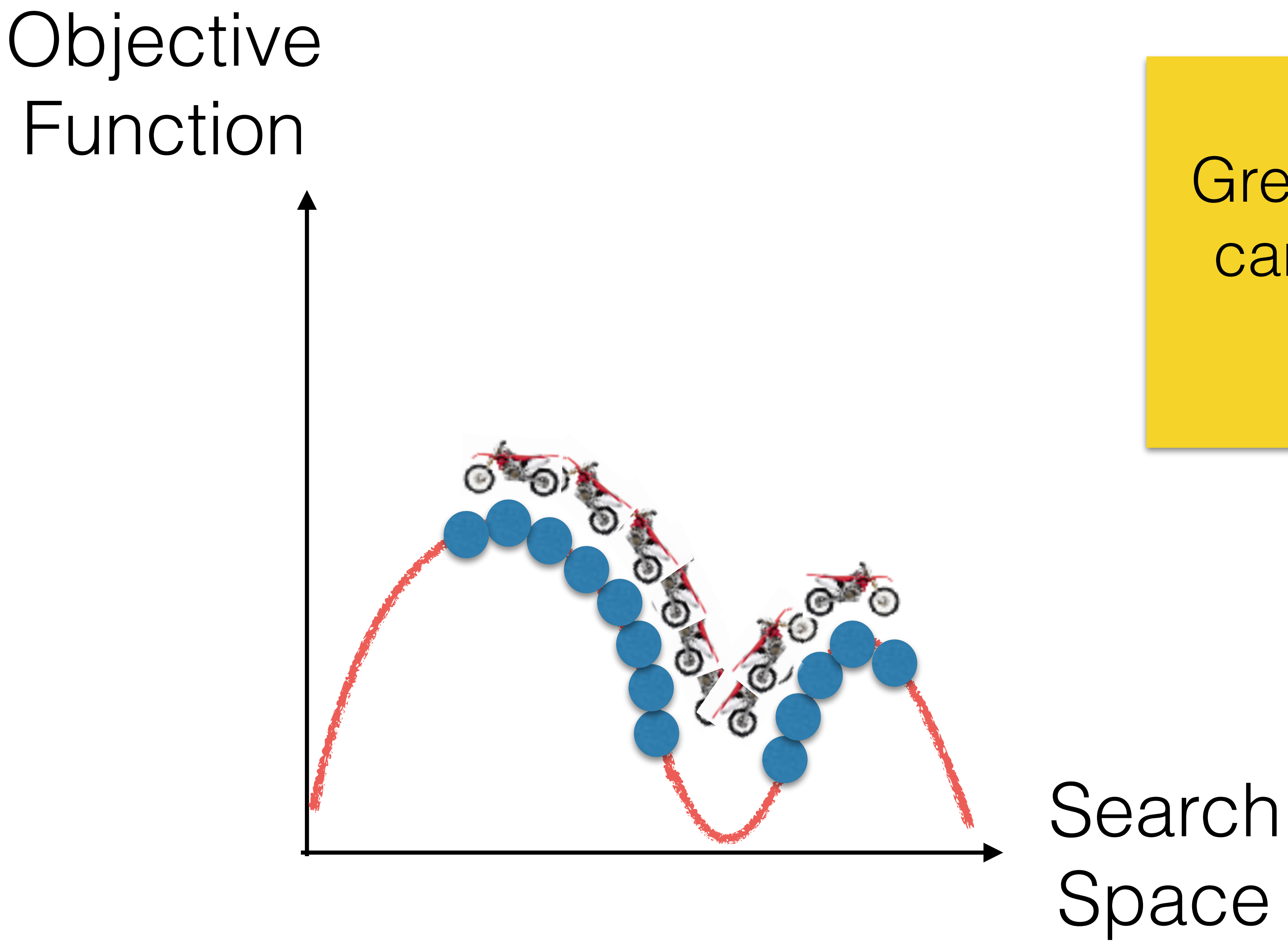
    3.3 Mutate resulting offspring with probability Pm

    3.4 Evaluate offspring

    3.5 Select survivors for the next generation

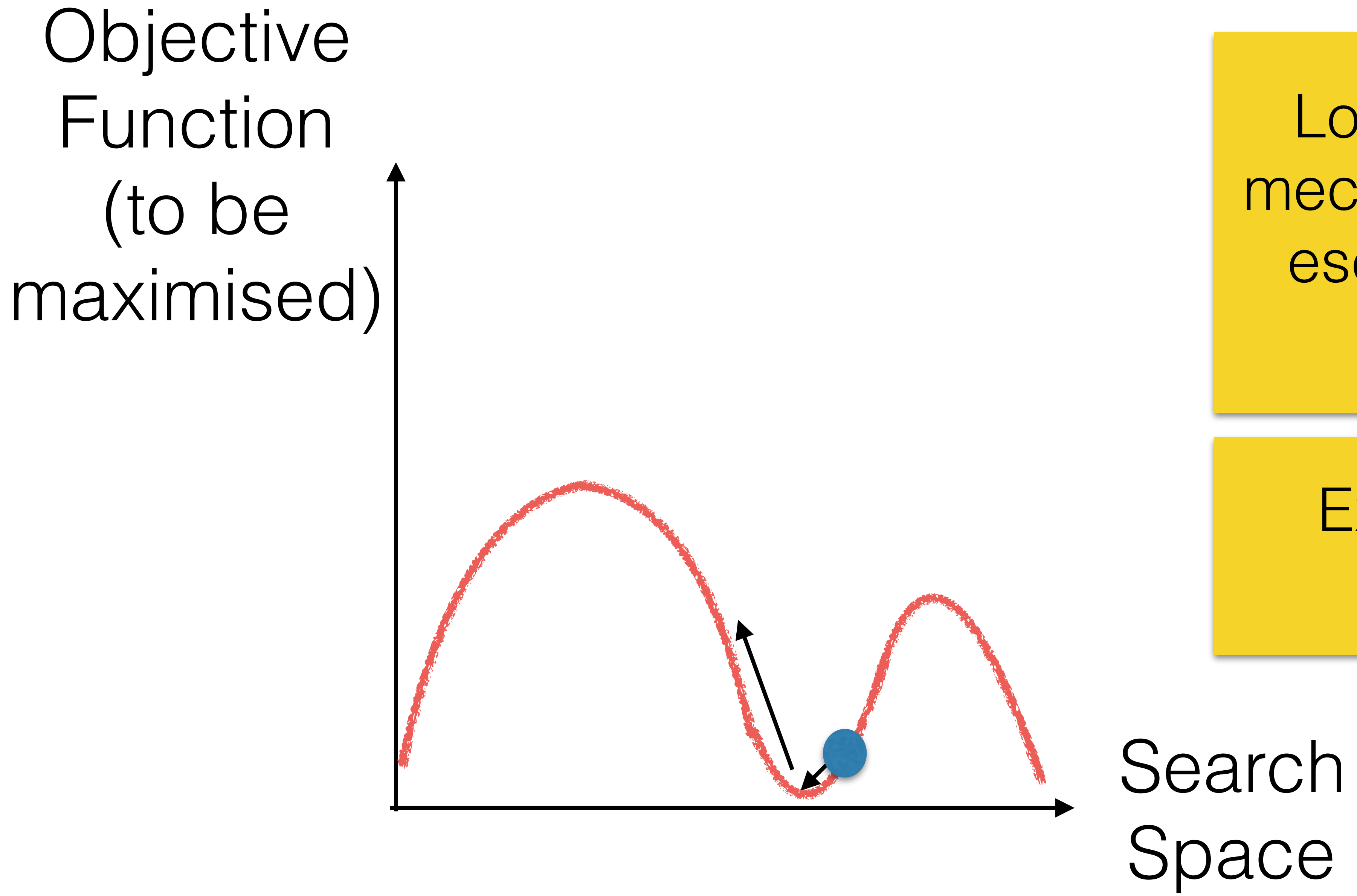
Parents and/or survivors selection is based on a selective **pressure** towards fitter (better) individuals.

# Hill-Climbing




Greedy local search can get trapped in local optima.

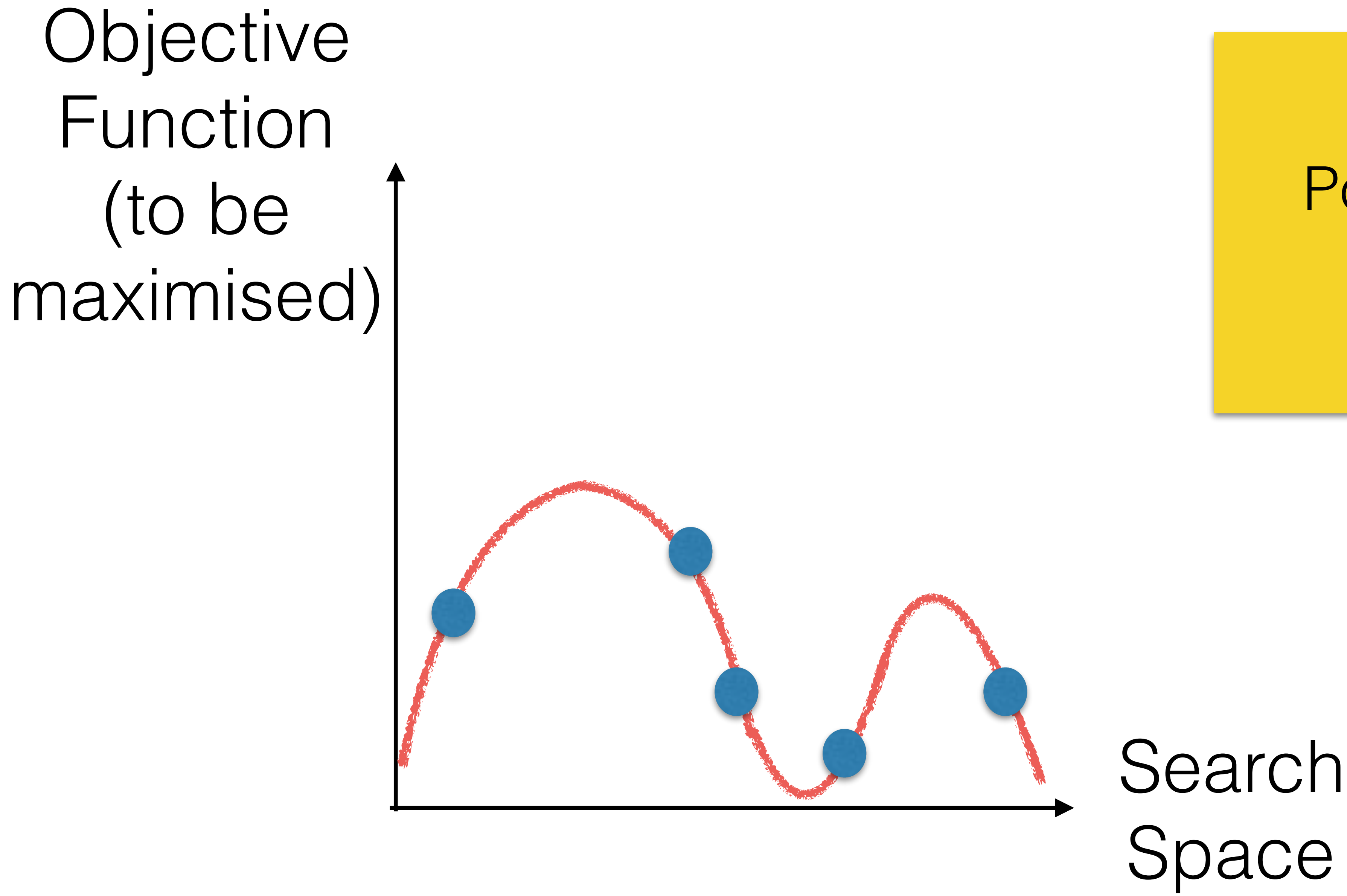
# Simulated Annealing




Local search with mechanisms to avoid escape from local optima.

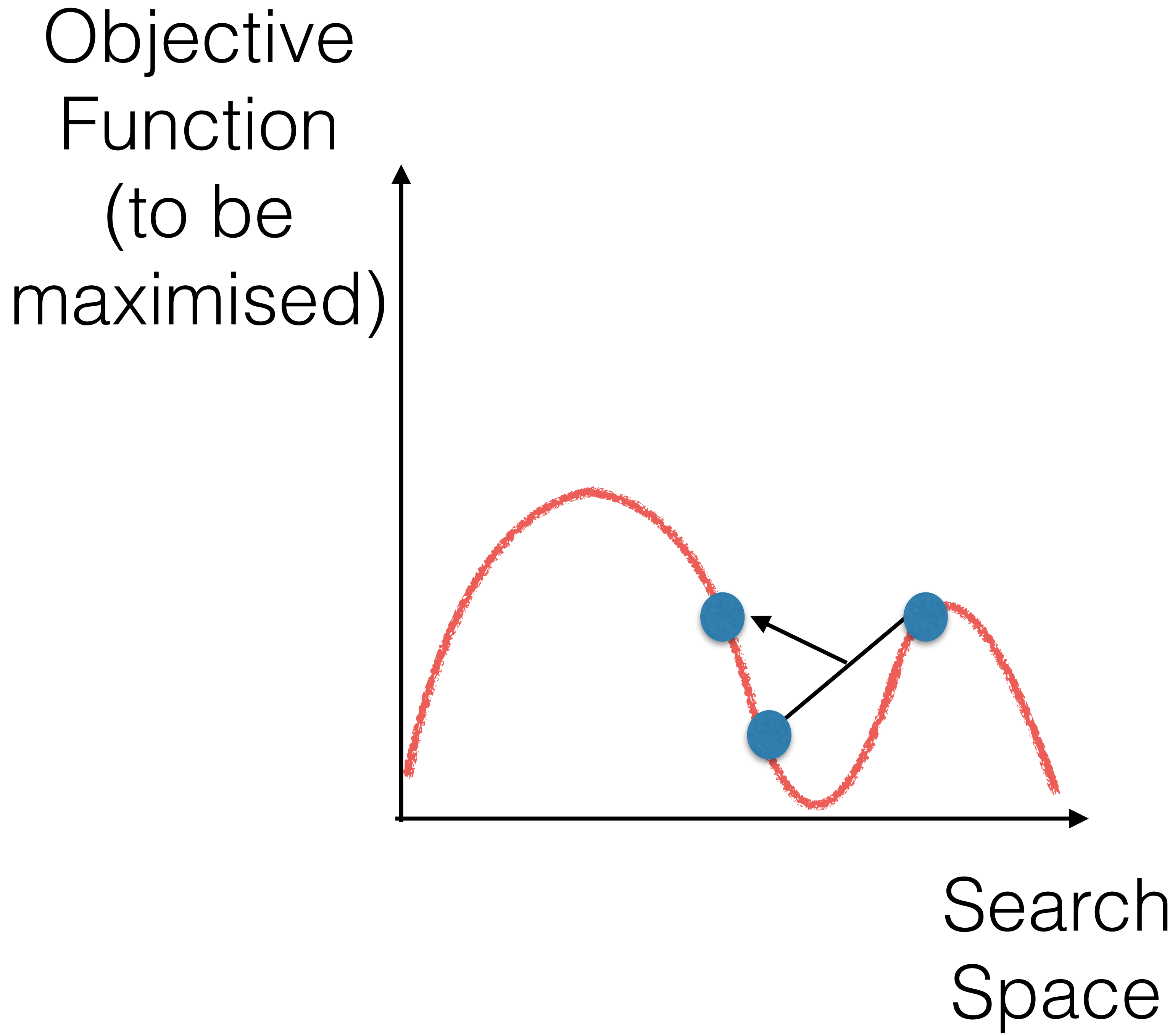
Exploration and exploitation.

# Evolutionary Algorithms (EAs)




Population helps exploration.

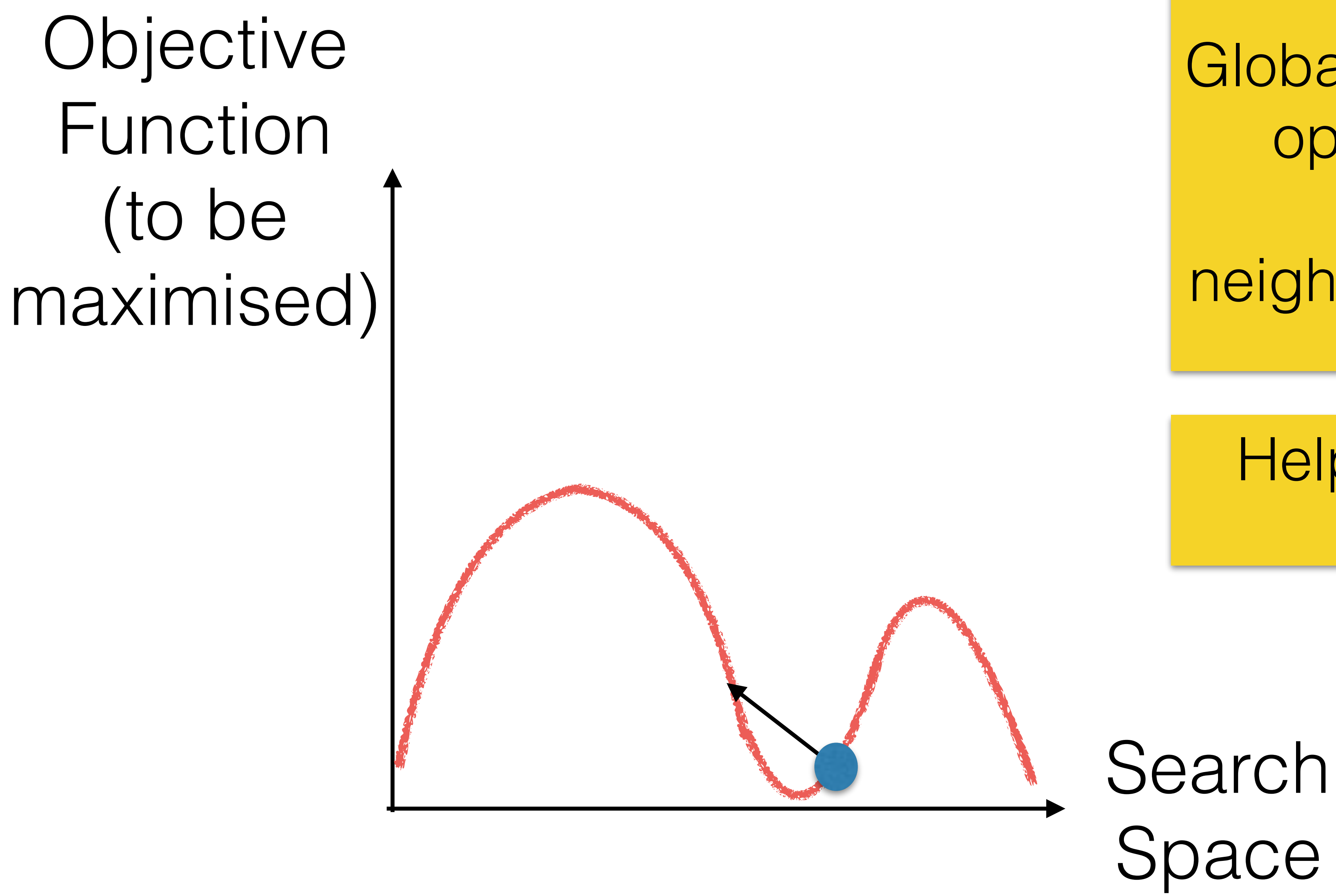
# Evolutionary Algorithms (EAs)




Bad individuals have some (small) chance to reproduce.

Helps to avoid local optima.

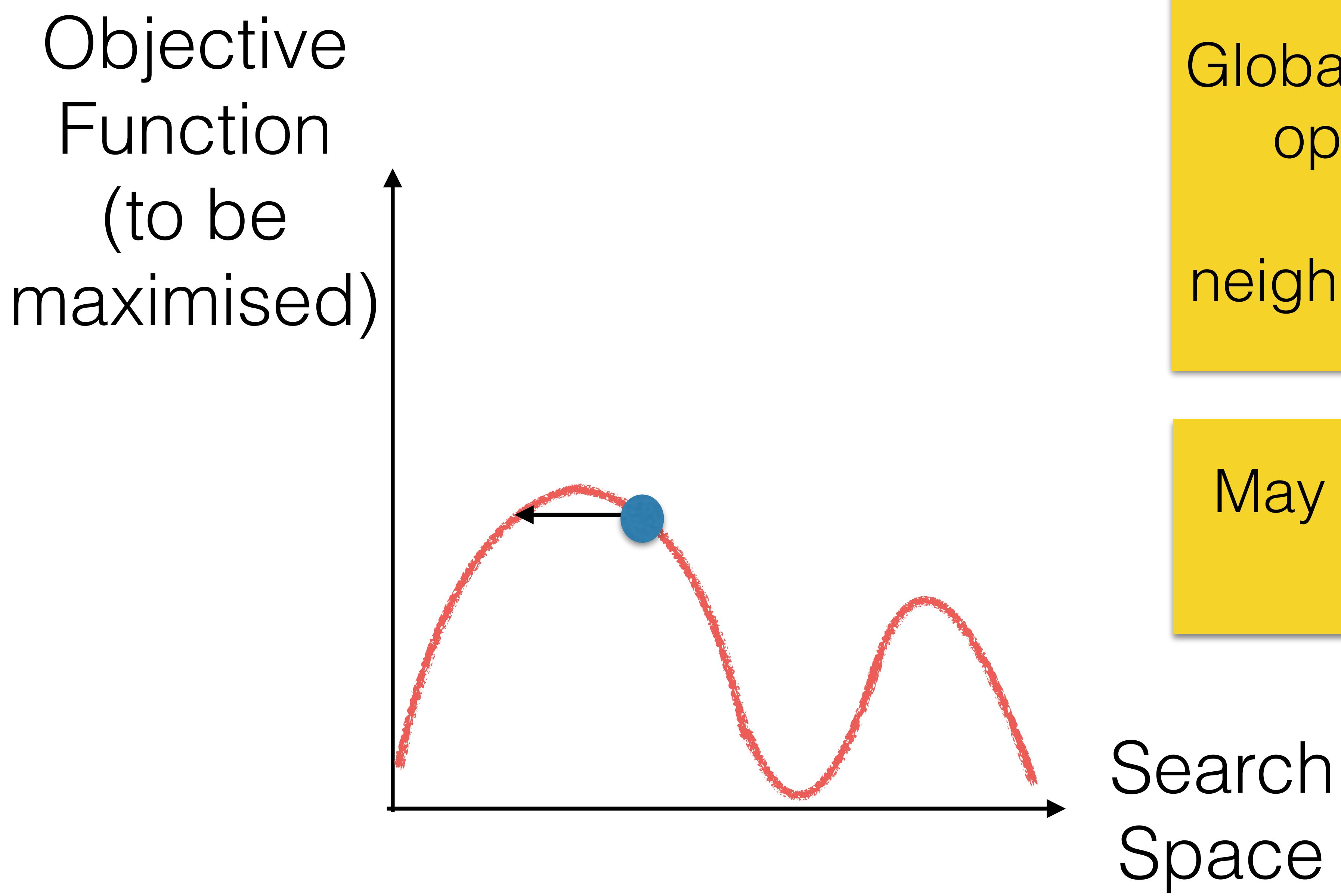
# Evolutionary Algorithms (EAs)




Global search: variation operators are not restricted to neighbouring solutions.

Helps to avoid local optima.

# Evolutionary Algorithms (EAs)




Global search: variation operators are not restricted to neighbouring solutions.

May jump passed the optimum.

# Typical Behaviour of an EA

Early phase:
quasi-random population distribution

Mid-phase:
population arranged around/on hills

Late phase:
population concentrated on high hills

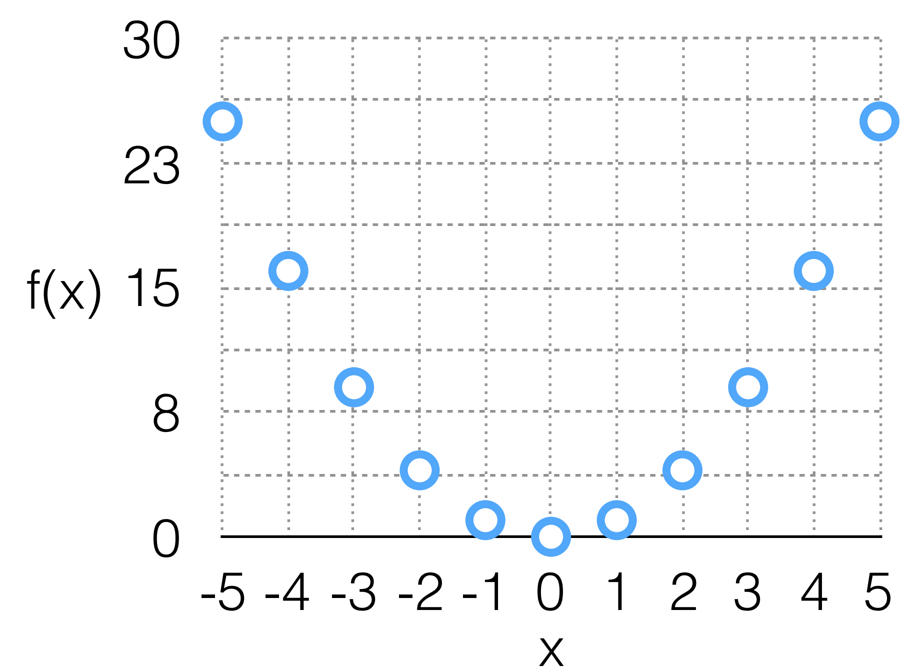
# Illustrative Optimisation Problem

- Problem: maximise $f(x) = x^2$, $x \in \{-15,-14,\ldots,0,1,2,\ldots,15\}$.
  - Design variable:
    $x \in \{-15,-14,\ldots,0,1,2,\ldots,15\}$.
    - Search space:

      $\{-15,-14,\ldots,0,1,2,\ldots,15\}$.
  - Objective function:

    $f(x) = x^2$ (to be maximised)
  - No constraints.

Illustrative plot for $-5 \leq x \leq 5$:



# Representation (=Encoding)

Genotype (Representation)

Image from: http://uioslonorway.files.wordpress.com/2014/05/dna.jpg

Phenotype

Image from: http://i2.cdn.turner.com/cnnnext/dam/assets/150324154010-04-internet-cats-restricted-super-169.jpg

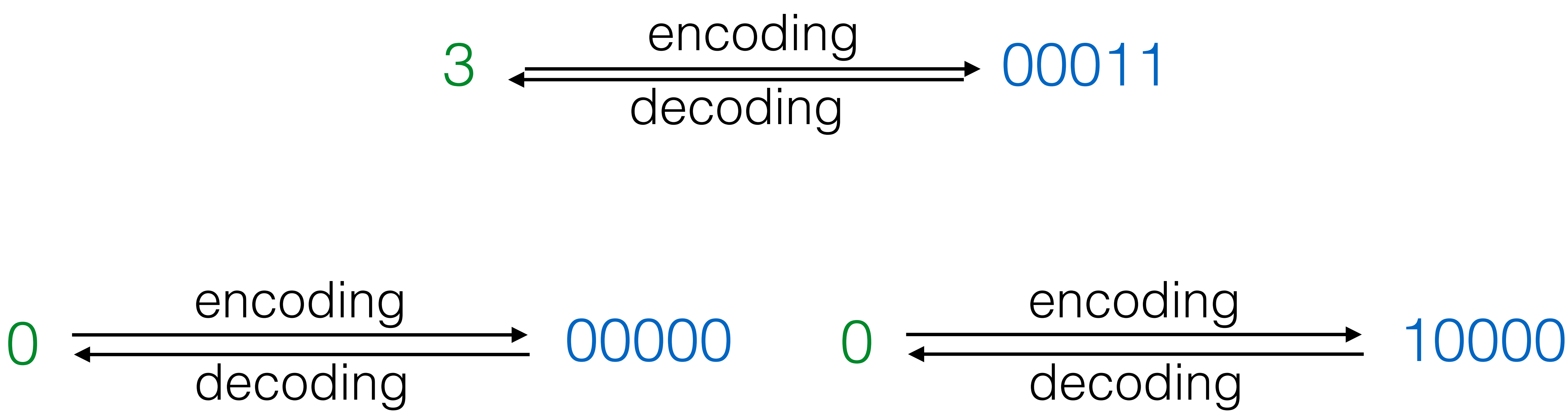
A phenotype that cannot be represented in the genotypic space cannot exist.

# Representation

- Several different representations could be used for a given problem.
- Ideally, representations should allow all feasible solutions to be represented.
- It is helpful if the representation is easy to manipulate by the algorithm.

# Binary Representation

- Genotype space = $\{0,1\}^L$
- Example:
  - Problem: maximise $f(x) = x^2$, $x \in \{-15,-14,\ldots,0,1,2,\ldots,15\}$
  - Representation: $\{0,1\}^5$, where the first bit represents x's sign (1 for negative and 0 for positive).



# Representation

- Binary vector.
    - E.g., for the lorry problem.
- Integer vector.
    - E.g., for the problem of maximising $f(x) = x^2$, $x \in \{-15,-14,\ldots,0,1,2,\ldots,15\}$.
    - E.g., if your design variable is categorical (e.g., in {Toyota, Volkswagen, Fiat, Vauxhall}).
- Floating-point vector.
    - E.g., for the problem of maximising $f(x_1, x_2) = x_1 + x_2$; $x_1, x_2 \in [0,1]$
- Permutations.
    - E.g., for the traveling salesman problem.
- Matrices.
    - E.g., for staff allocation problems.
- Etc.

# EA's Pseudocode

Evolutionary Algorithm

**1. Initialise population**

2. Evaluate each individual (determine their fitness)

3. Repeat (until a termination condition is satisfied)

    3.1 Select parents

    3.2 Recombine parents with probability Pc

    3.3 Mutate resulting offspring with probability Pm

    3.4 Evaluate offspring

    3.5 Select survivors for the next generation

# Initialisation

- Initialisation usually done at random.
- Need to ensure even spread and mixture of possible allele values.
- Can include existing solutions, or use problem-specific heuristics, to “seed” the population.

# Example: Random Initialisation

- Create each individual of the population randomly.
  - E.g., create one individual, considering the following representation $\{0,1\}^5$

0 0 1 1 1

# EA's Pseudocode

Evolutionary Algorithm

1. Initialise population

**2. Evaluate each individual (determine their fitness)**

3. Repeat (until a termination condition is satisfied)

    3.1 Select parents

    3.2 Recombine parents

    3.3 Mutate resulting offspring

    3.4 Evaluate offspring

    3.5 Select survivors for the next generation

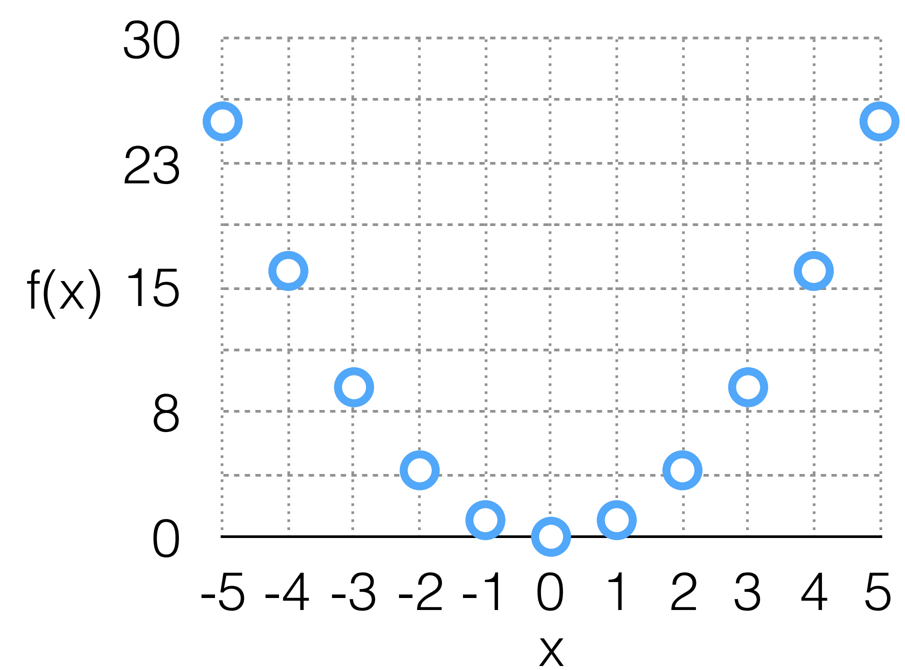
# Fitness Function

- Problem-dependent.
  - Comes from the objective or quality function.
- Represents the requirements that the population should adapt to.
- Assigns a single real-valued fitness to each phenotype.
- Typically we talk about fitness being maximised.
- Some problems may be best posed as minimisation problems, but conversion is possible.

# Illustrative Optimisation Problem

- Problem: maximise $f(x) = x^2$, $x \in \{-15,-14,\ldots,0,1,2,\ldots,15\}$.
  - Design variable:
    $x \in \{-15,-14,\ldots,0,1,2,\ldots,15\}$.
    - Search space:
      $\{-15,-14,\ldots,0,1,2,\ldots,15\}$.
  - **Objective function:**
    **$f(x) = x^2$ (to be maximised)**
  - No constraints.

Illustrative plot for $-5 \leq x \leq 5$:



# EA's Pseudocode

Evolutionary Algorithm

1. Initialise population
2. Evaluate each individual (determine their fitness)
3. Repeat (until a termination condition is satisfied)

    **3.1 Select parents**

    3.2 Recombine parents with probability Pc

    3.3 Mutate resulting offspring with probability Pm

    3.4 Evaluate offspring

    3.5 Select survivors for the next generation

# Parent Selection

- Usually probabilistic:
  - High quality solutions more likely to become parents than low quality.
  - Even the worst in current population usually has non-zero probability of becoming a parent.

- This stochastic nature can help to escape from local optima.


Image from: https://lh6.googleusercontent.com/-wCEtlOfs4lI/TXjes2fSfaI/AAAAAAAABEg/7yOX_b1D2Ho/s1600/pavoesMenor.jpg

# How Many Parents to Select?

- This is a design choice of the algorithm.
- Frequently, if your population size is *S*, you choose the number of parents so as to produce *S* children.
- E.g., if each pair of parents can produce 2 children by recombination, you could select *S* parents to produce *S* children.

# Parent Selection Mechanisms

- Roulette Wheel
- Tournament Selection
- Ranking Selection

# Parent Selection Mechanisms

- **Roulette Wheel**
- **Tournament Selection**
- Ranking Selection

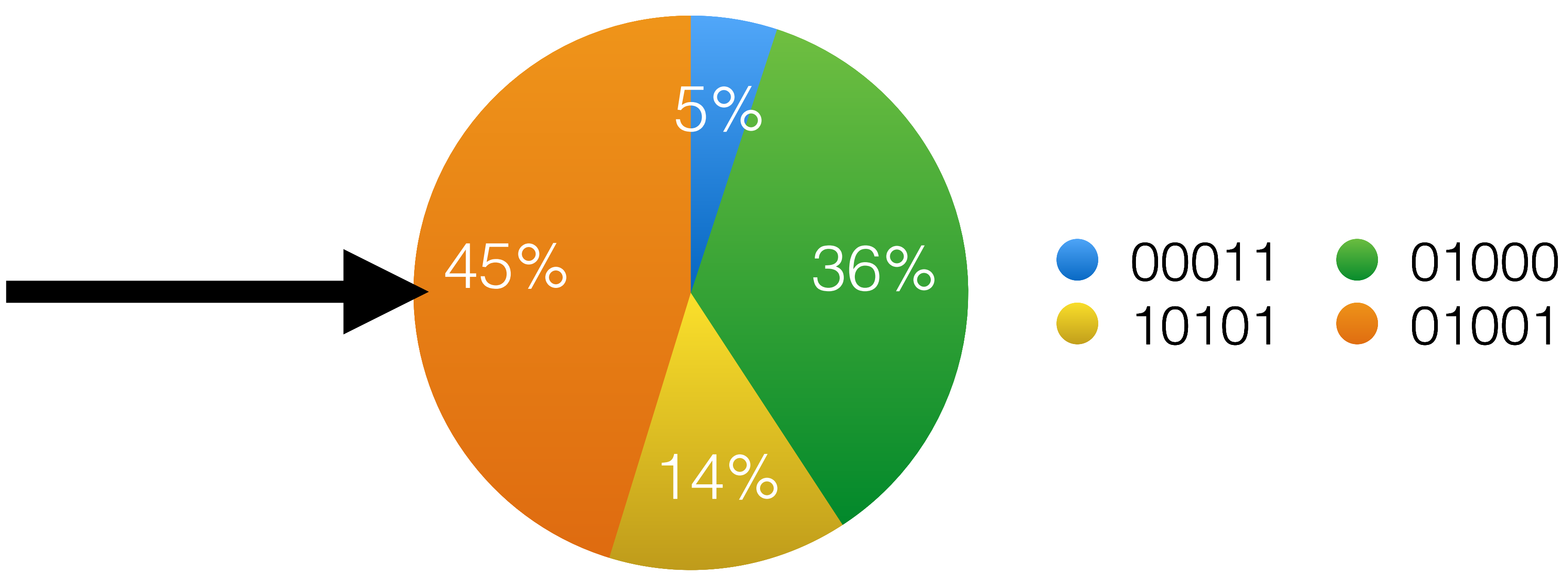
# Roulette Wheel Parents Selection

- Probability of an individual to be selected as parent is proportional to its fitness. Assuming maximisation of positive fitnesses: $f(x) / \Sigma f(x)$.
- Example:
  - Problem: maximise $f(x) = x^2$, $x \in \{-15,-14,\ldots,0,1,2,\ldots,15\}$
  - Representation: $\{0,1\}^5$.

| Genotypes | Phenotypes | Fitnesses | Probability |
|---|---|---|---|
| 00011 | 3 | 9 | 9/179 = 0.0503 |
| 01000 | 8 | 64 | 64/179 = 0.3575 |
| 10101 | -5 | 25 | 25/179 = 0.1397 |
| 01001 | 9 | 81 | 81/179 = 0.4525 |
| Sum (Σ): | | 179 | 1 |

# Roulette Wheel Parents Selection — Selecting 4 Parents

| Genotypes | Phenotypes | Fitnesses | Probability |
|---|---|---|---|
| 00011 | 3 | 9 | 9/179 = 0.0503 |
| 01000 | 8 | 64 | 64/179 = 0.3575 |
| 10101 | -5 | 25 | 25/179 = 0.1397 |
| 01001 | 9 | 81 | 81/179 = 0.4525 |
| Sum (Σ): | | 179 | 1 |




Randomly selected Parents:

01001
10101
01000
01000

# Problems of Roulette Wheel Parents Selection

- Outstanding individuals may take over the population very quickly, causing premature convergence.
- When fitness values are very close to each other, there is almost no selection pressure.
- The mechanism behaves differently on transposed versions of the same function.

# Problems of Roulette Wheel Parents Selection

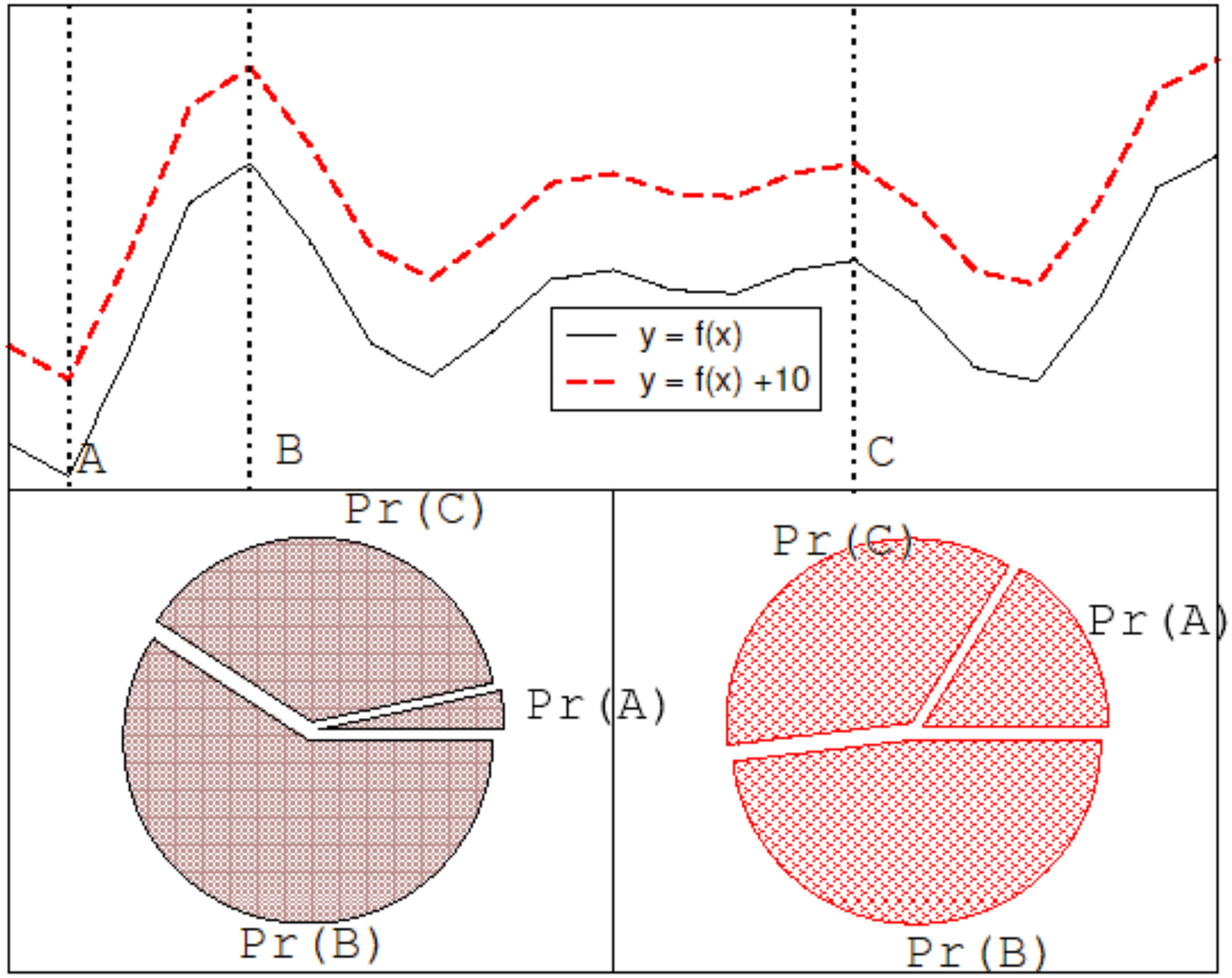



Image from Eiben and Smith's slides.

# Tournament Selection

- Informal Procedure:
  - Pick *k* members at random then select the best of these.
  - Repeat to select more individuals.

E.g.: k = 2, assuming maximisation

| | Genotypes | Phenotypes | Fitnesses |
|---|---|---|---|
| → | 00011 | 3 | 9 |
| | 01000 | 8 | 64 |
| | 10101 | -5 | 25 |
| → | 01001 | 9 | 81 |

Parent: 01001

# EA's Pseudocode

Evolutionary Algorithm

**1. Initialise population**

**2. Evaluate each individual (determine their fitness)**

**3. Repeat (until a termination condition is satisfied)**

**3.1 Select parents**

3.2 Recombine parents with probability Pc

3.3 Mutate resulting offspring with probability Pm

3.4 Evaluate offspring

3.5 Select survivors for the next generation

# Further Reading

- Mechanisms: the processes of evolution
  http://evolution.berkeley.edu/evolibrary/article/0_0_0/evo_14

- Eiben and Smith, Introduction to Evolutionary Computing, Chapter 2 (What is an Evolutionary Algorithm?) and Chapter 3 (Genetic Algorithms), Springer 2003.