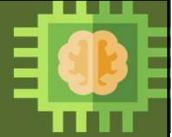


Elective Course

Course Code: CS4103

Autumn 2025-26



## Lecture #21

# Artificial Intelligence for Data Science

## Week-6: Introduction to Genetic Algorithm (GA) [Part-II]

Course Instructor:

Dr. Monidipa Das

Assistant Professor

Department of Computational and Data Sciences

Indian Institute of Science Education and Research Kolkata, India 741246

## Genetic Algorithm: An Overview



- **Step-1:** Start with a randomly generated population of individuals.  
The population size is  $n$ .
- **Step-2:** Calculate the fitness of each individual in the population.
- **Step-3:** Repeat the following steps until  $n$  children or offspring are created.
  - **Select** pair of parents from the population
  - With probability  $P_c$  perform crossover on these selected parents to generate two offspring
  - Mutate offspring with probability  $P_m$
- **Step-4:** Replace the old population with this new population
- **Step-5:** Go to step 2 until the termination condition is reached.

# Selection



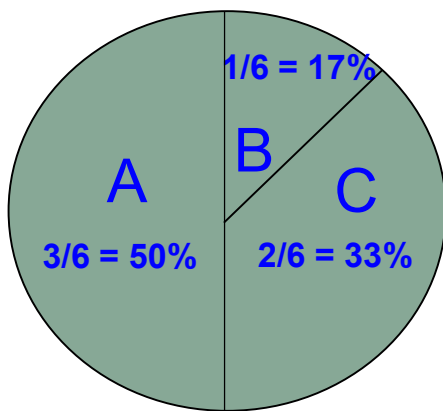
- The purpose is to bias the mating pool (those who can pass on their traits to the next generation) with fitter individuals
- Selection must NOT be too strong or too weak
- Selection Strategies
  - Random Selection
  - Roulette Wheel Selection
  - Rank Selection
  - Tournament selection etc.

Dr. Monidipa Das, Department of CDS, IISER Kolkata

## Roulette wheel selection



- Assign to each individual a part of the roulette wheel
- Spin the wheel  $n$  times to select  $n$  individuals



$$\text{fitness}(A) = 3$$

$$\text{fitness}(B) = 1$$

$$\text{fitness}(C) = 2$$

Dr. Monidipa Das, Department of CDS, IISER Kolkata

## Rank selection



- In Linear Rank selection, individuals are assigned subjective fitness based on the rank within the population:
  - $sf_i = (P - r_i)(max - min)/(P - 1) + min$
  - Where  $r_i$  is the rank of individual  $i$ ,
  - $P$  is the population size,
  - $max$  represents the fitness to assign to the best individual,
  - $min$  represents the fitness to assign to the worst individual.
- $p_i = sf_i / \sum sf_j$ . Roulette Wheel Selection can be performed using the subjective fitnesses.
- One disadvantage associated with linear rank selection is that the population must be sorted on each cycle.

Dr. Monidipa Das, Department of CDS, IISER Kolkata

## Tournament selection



- In tournament selection several tournaments are played among a few individuals.
- The individuals are chosen at random from the population.
- The winner of each tournament is selected for next generation.
- Selection pressure can be adjusted by changing the tournament size.
- Weak individuals have a smaller chance to be selected if tournament size is large.

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# Genetic Algorithm: An Overview



- **Step-1:** Start with a randomly generated population of individuals.  
The population size is  $n$ .
- **Step-2:** Calculate the fitness of each individual in the population.
- **Step-3:** Repeat the following steps until  $n$  children or offspring are created.
  - Select pair of parents from the population
  - With probability  $P_c$  perform crossover on these selected parents to generate two offspring
  - Mutate offspring with probability  $P_m$
- **Step-4:** Replace the old population with this new population
- **Step-5:** Go to step 2 until the termination condition is reached.

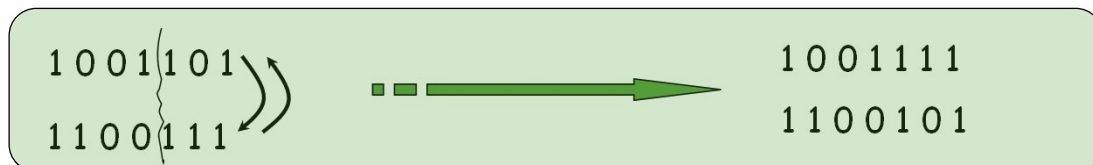
Dr. Monidipa Das, Department of CDS, IISER Kolkata

## Crossover



Crossover is a critical feature of genetic algorithms:

- It greatly accelerates search early in evolution of a population
- Choose a random point on the two parents
- Split parents at this crossover point
- Create children by exchanging tails
- $P_c$  typically in range (0.6, 0.9)



Dr. Monidipa Das, Department of CDS, IISER Kolkata

# Crossover



- One-point crossover
- Two-point crossover
- N-point crossover
- Uniform crossover

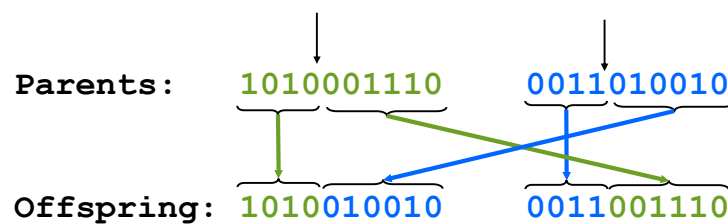
Dr. Monidipa Das, Department of CDS, IISER Kolkata

## One-point crossover



- Randomly one position in the chromosomes is chosen
- Child 1 is head of chromosome of parent 1 with tail of chromosome of parent 2
- Child 2 is head of 2 with tail of 1

Randomly chosen position



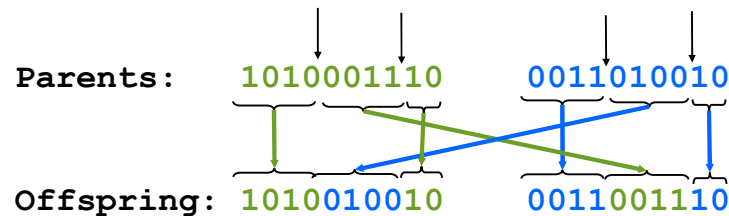
Dr. Monidipa Das, Department of CDS, IISER Kolkata

## Two-point crossover



- Randomly two positions in the chromosomes are chosen
- Avoids that genes at the head and genes at the tail of a chromosome

### Randomly chosen position

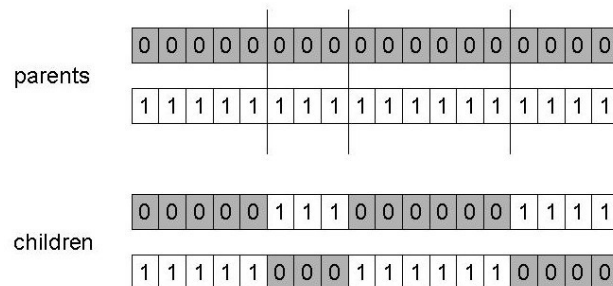


Dr. Monidipa Das, Department of CDS, IISER Kolkata

## n-point crossover



- Choose n random crossover points
- Split along those points
- Glue parts, alternating between parents
- Generalisation of 1 point (still some positional bias)



Dr. Monidipa Das, Department of CDS, IISER Kolkata

## Uniform crossover



- A random mask is generated
- The mask determines which bits are copied from one parent and which from the other parent
- Bit density in mask determines how much material is taken from the other parent (takeover parameter)

**Mask:**            0110011000            (Randomly generated)

**Parents:**        1010001110            0011010010

**Offspring:**      0011001010            1010010110

Dr. Monidipa Das, Department of CDS, IISER Kolkata

## Crossover



- **Cross-over for Continuous Parameters**
  - A cross-over operator that linearly combines two parent chromosome vectors to produce two new offspring according to the following equations:
    - $\text{Offspring1} = a * \text{Parent1} + (1-a) * \text{Parent2}$
    - $\text{Offspring2} = (1-a) * \text{Parent1} + a * \text{Parent2}$
  - Consider the following two parents (each consisting of four floating genes) which have been selected for cross-over:
    - Parent1: (0.3 1.4 0.2 7.4)
    - Parent2: (0.5 4.5 0.1 5.6)
- If **a = 0.7**, the following two offspring would be produced:
  - O1=(0.36 2.33 0.17 6.86)
  - O2=(0.44 3.57 0.13 6.14)

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# Problems with crossover



- Depending on coding, simple crossovers can have high chance to produce illegal offspring
  - E.g. in TSP with simple binary or path coding, most offspring will be illegal because not all cities will be in the offspring and some cities will be there more than once
- Uniform crossover can often be modified to avoid this problem
  - E.g. in TSP with simple path coding:
    - Where mask is 1, copy cities from one parent
    - Where mask is 0, choose the remaining cities in the order of the other parent

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# Genetic Algorithm: An Overview



- **Step-1:** Start with a randomly generated population of individuals.  
The population size is  $n$ .
- **Step-2:** Calculate the fitness of each individual in the population.
- **Step-3:** Repeat the following steps until  $n$  children or offspring are created.
  - Select pair of parents from the population
  - With probability  $P_c$  perform crossover on these selected parents to generate two offspring
  - Mutate offspring with probability  $P_m$
- **Step-4:** Replace the old population with this new population
- **Step-5:** Go to step 2 until the termination condition is reached.

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# Mutation Operator



- Mutation represents a change in the gene.
- Mutation is a background operator. Its role is to provide a guarantee that the search algorithm is not trapped on a local optimum.
- The mutation operator flips a randomly selected gene in a chromosome.
- The mutation probability is quite small in nature, and is kept low for GAs, typically in the range between 0.001 and 0.01.

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# Mutation



- Alter each gene independently with a probability  $P_m$
- $P_m$  is called the mutation rate
- Typically between  $1/pop\_size$  and  $1/chromosome\_length$



parent

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

child

0	1	0	0	1	0	1	1	0	0	0	1	0	1	1	0	0	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# Example



- Simple problem: maximize  $x^2$  over  $\{0, 1, \dots, 31\}$
- **GA approach:**
  - Representation: binary code, e.g.  $01101 \leftrightarrow 13$
  - Population size: 4
  - Random initialization
  - Roulette wheel selection
  - 1-point crossover, bitwise mutation

Dr. Monidipa Das, Department of CDS, IISER Kolkata

## $x^2$ example: selection



String no.	Initial population	$x$ Value	Fitness $f(x) = x^2$	$Prob_i$
1	0 1 1 0 1	13	169	0.14
2	1 1 0 0 0	24	576	0.49
3	0 1 0 0 0	8	64	0.06
4	1 0 0 1 1	19	361	0.31
Sum			1170	1.00
Average			293	0.25
Max			576	0.49

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# $x^2$ example: selection



String no.	Mating pool	Crossover point	Offspring after xover	$x$ Value	Fitness $f(x) = x^2$
1	0 1 1 0   1	4	0 1 1 0 0	12	144
2	1 1 0 0   0	4	1 1 0 0 1	25	625
2	1 1   0 0 0	2	1 1 0 1 1	27	729
4	1 0   0 1 1	2	1 0 0 0 0	16	256
Sum					1754
Average					439
Max					729

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# $x^2$ example: selection



String no.	Offspring after xover	Offspring after mutation	$x$ Value	Fitness $f(x) = x^2$
1	0 1 1 0 0	1 1 1 0 0	26	676
2	1 1 0 0 1	1 1 0 0 1	25	625
2	1 1 0 1 1	1 1 0 1 1	27	729
4	1 0 0 0 0	1 0 1 0 0	18	324
Sum				2354
Average				588.5
Max				729

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# Practical Example - TSP

## TSP Encoding



- **Integer representation**
  - Tour 1-3-2: (1 3 2)
- **Binary representation**
  - Tour 1-3-2 is represented as: ( 00 10 01 )
- **Character/Alphabet representation**
  - Tour CityA-CityB-CityC: (A B C)

# TSP – Crossover operator



## • Order Based crossover (OX2)

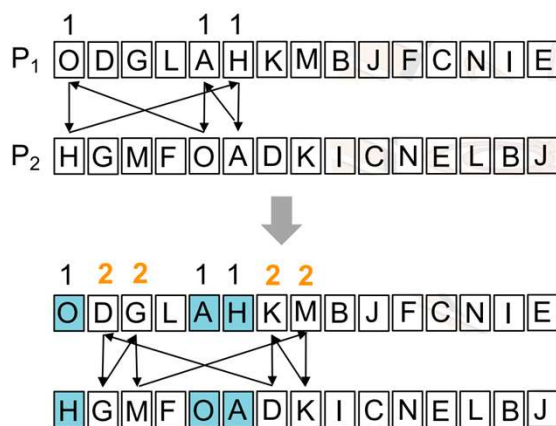
- Selects at random several positions in the parent tour
- Imposes the order of nodes in selected positions of one parent on the other parent
- Parents: (1 2 3 4 5 6 7 8) and (2 4 6 8 7 5 3 1)
- Selected positions, 2<sup>nd</sup>, 3<sup>rd</sup> and 6<sup>th</sup>
- Impose order on (2 4 6 8 7 5 3 1) & (1 2 3 4 5 6 7 8)
- Children are (2 4 3 8 7 5 6 1) and (1 2 3 4 6 5 7 8)

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# TSP – Crossover operator

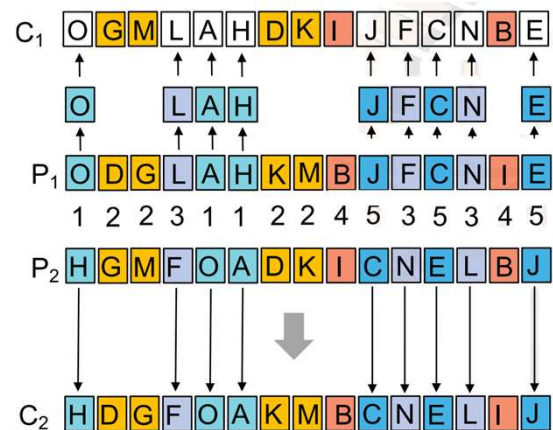


## Cycle Crossover



Adapted from NPTEL Course: Prof. Deepak Khemani, IIT Madras

C1 gets even numbered cycles from P2 and odd numbered cycles from P1



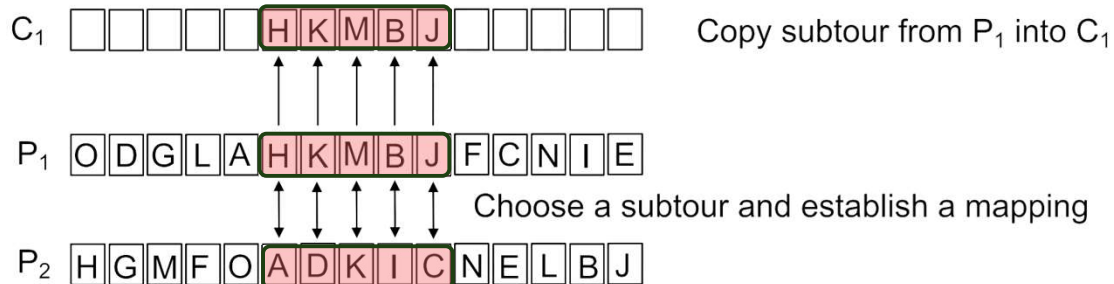
C2 gets even numbered cycles from P1 and odd numbered cycles from P2

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# TSP – Crossover operator



## Partially Mapped Crossover (PMX)



Would like to copy remaining cities from  $P_2$   
but  
the locations for cities A, D, I, C are occupied  
by cities H, K, B, J respectively

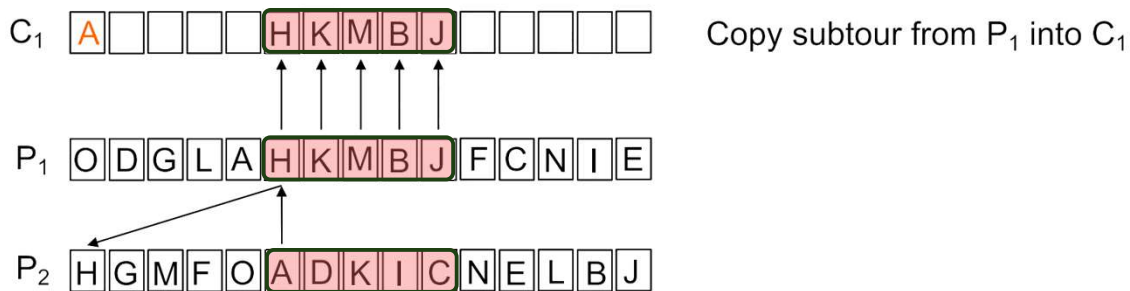
Adapted from NPTEL Course: Prof. Deepak Khemani, IIT Madras

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# TSP – Crossover operator



## Partially Mapped Crossover (PMX)



Where should city A be in  $C_1$ ?

Follow the partial map...

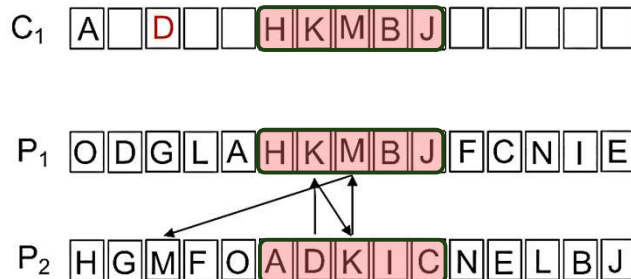
Adapted from NPTEL Course: Prof. Deepak Khemani, IIT Madras

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# TSP – Crossover operator



## Partially Mapped Crossover (PMX)



Where should city D be in  $C_1$ ?

Follow the partial map...

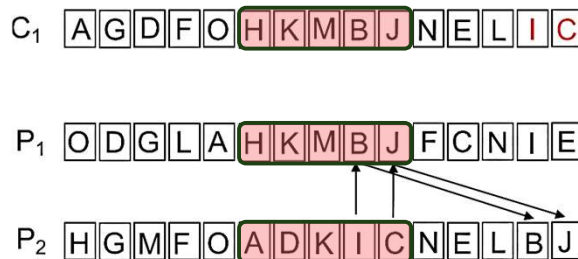
Adapted from NPTEL Course: Prof. Deepak Khemani, IIT Madras

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# TSP – Crossover operator



## Partially Mapped Crossover (PMX)



Likewise for cities I and C

Remember that city K is already in  $C_1$ ...

Copy the remaining cities directly from  $P_2$

The second child  $C_2$  is constructed in a similar manner, first copying the subtour from  $P_2$

### Homework

**What will be the chromosome representation of child2 ( $C_2$ ) as per PMX?**

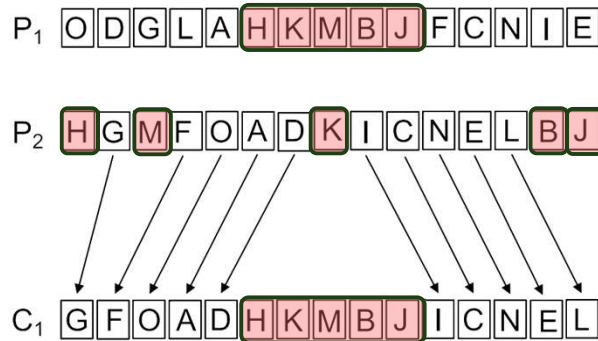
Adapted from NPTEL Course: Prof. Deepak Khemani, IIT Madras

Dr. Monidipa Das, Department of CDS, IISER Kolkata

# TSP – Crossover operator



## Order Crossover



Copy a subtour from  $P_1$  into  $C_1$  and the remaining from  $P_2$  in the order they occur in  $P_2$ .

### Homework

What will be the chromosome representation of child2 ( $C_2$ ) as per this scheme?

Adapted from NPTEL Course: Prof. Deepak Khemani, IIT Madras

Dr. Monidipa Das, Department of CDS, IISER Kolkata



# Questions?

Dr. Monidipa Das, Department of CDS, IISER Kolkata