

- Evolutionary algorithms: Origins
- Evolutionary algorithms: Components
- Evolutionary algorithms: Applications





# Heuristic Search and Evolutionary Algorithms

# Lecture 6: Evolutionary Algorithms – Representation, Mutation and Recombination

# Chao Qian (钱超)

Associate Professor, Nanjing University, China

Email: qianc@nju.edu.cn Homepage: http://www.lamda.nju.edu.cn/qianc/

# Representation and variation operators

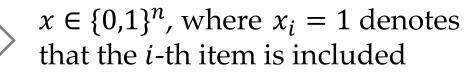
- The first stage of applying evolutionary algorithms is to decide a right representation for the problem
  - Binary representation
  - Integer representation
  - Real-valued representation
  - Permutation representation
  - Tree representation
- Variation operators depend on the chosen representation
  - Mutation
  - Recombination

# Binary representation

• Genotype space:  $\{0,1\}^n$ 

Knapsack problem with *n* items





$$\arg \max_{x \in \{0,1,\dots,15\}} x^2$$

 $x \in \{0,1\}^4$ , and the corresponding integer is  $\sum_{i=1}^4 x_i 2^{4-i}$ 

# Binary representation

• Genotype space:  $\{0,1\}^n$ 

$$\arg \max_{x \in \{0,1,\dots,15\}} x^2$$



Standard binary coding 7 is represented by 0111 8 is represented by 1000 9 is represented by 1001

#### $x \in \{0,1\}^4$ , and the corresponding integer is $\sum_{i=1}^4 x_i 2^{i-1}$

The Hamming distance between consecutive integers can be very large

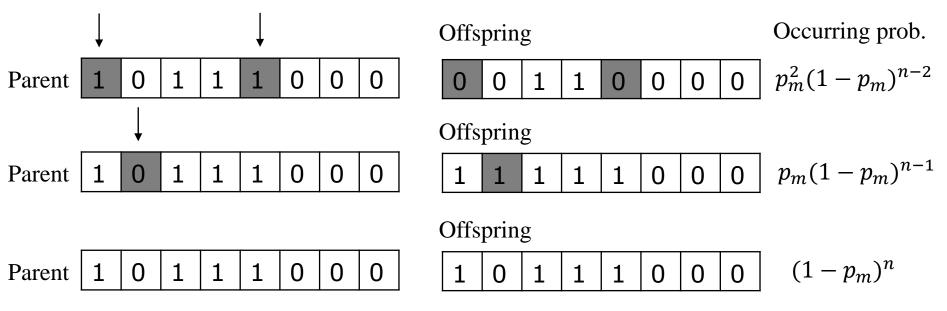
#### Gray coding

0	0000	4	0110	8	1100	12	1010
1	0001	5	0111	9	1101	13	1011
2	0011	6	0101	10	1111	14	1001
3	0010	7	0100	11	1110	15	1000

The Hamming distance between consecutive integers is always 1

# Binary representation: Mutation

• Bit-wise mutation: flip each bit independently with prob.  $p_m$ 



The common setting of  $p_m$  is 1/n, where *n* is the length of the binary string

• One-bit mutation: flip a randomly chosen bit

For the above mutation behaviors, the occurring prob. are 0, 1/n and 0.

# Binary representation: Mutation

• Bit-wise mutation: flip each bit independently with prob.  $p_m$ 

The number of flipped bits is a random variable, denoted by *X* 

$$\mathbf{P}(X=k) = \binom{n}{k} p_m^k (1-p_m)^{n-k}$$

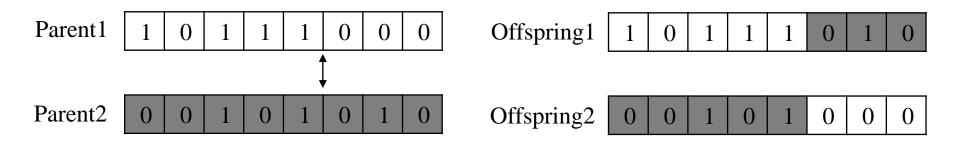
*X* satisfies the binomial distribution  $B(n, p_m)$ 

The expected number of flipped bits is  $E[X] = np_m$ , which is 1 for  $p_m = 1/n$ 

• One-bit mutation: flip a randomly chosen bit The number of flipped bits is 1

# Binary representation: Recombination

- One-point crossover
  - ≻ Choose a random number  $r \in \{1, 2, ..., n 1\}$
  - Split both parents at this point
  - Create two offspring by exchanging the tails

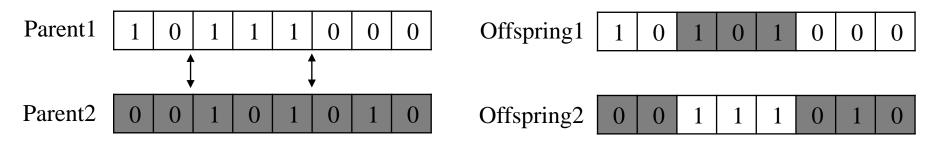


The occurring probability is 1/(n-1)

# Binary representation: Recombination

- *m*-point crossover
  - > Choose *m* crossover points from  $\{1, 2, ..., n 1\}$
  - Split both parents along these points
  - Create two offspring by taking alternative segments

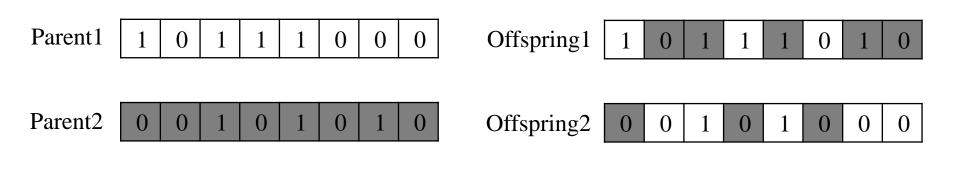
#### 2-point crossover



The occurring probability is 2/((n-1)(n-2))

# Binary representation: Recombination

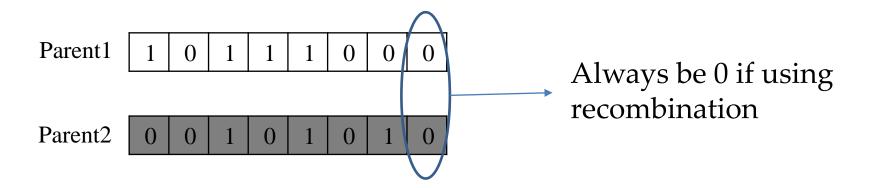
- Uniform crossover
  - For the first offspring, each gene is inherited from the first parent with probability *p* independently; otherwise from the second parent
  - The second offspring is created using the inverse mapping
    The common setting of p is 1/2



The occurring probability is  $p^3(1-p)^{n-3}$ 

# Binary representation: Mutation and recombination

• Only mutation can introduce new information



- Only recombination can combine information from two parents
  - One-point and *m*-point crossover: more likely to keep together genes that are near each other
  - Uniform crossover: no positional bias

- Genotype space: N<sup>*n*</sup>, where N denotes the integer
- Mutation: mutate each gene independently with probability *p<sub>m</sub>*

  - Creep mutation: add a small value, which is sampled from a distribution
    Ordinal

attributes

• **Recombination:** same as for binary representation

# Real-valued representation: Mutation

- Genotype space:  $\mathbb{R}^n$ , where  $\mathbb{R}$  denotes the real-number
- Mutation

$$\mathbf{x} = (x_1, \dots, x_n) \rightarrow \mathbf{x}' = (x_1', \dots, x_n')$$
, where  $x_i, x_i' \in [lb_i, ub_i]$ 

Uniform mutation: for each x<sub>i</sub>, with prob. p<sub>m</sub>, change it to a value drawn uniformly randomly from [lb<sub>i</sub>, ub<sub>i</sub>]

$$x'_{i} = \begin{cases} x_{i} & \text{with prob. } 1 - p_{m} \\ U(lb_{i}, ub_{i}) & \text{otherwise} \end{cases}$$

Nonuniform mutation: for each  $x_i$ , add a value drawn randomly from a Gaussian distribution  $N(0, \sigma^2)$ 

 $\sigma$ : mutation step size

$$x'_i = x_i + \delta$$
, where  $\delta \sim N(0, \sigma^2)$ 

### Real-valued representation: Nonuniform mutation

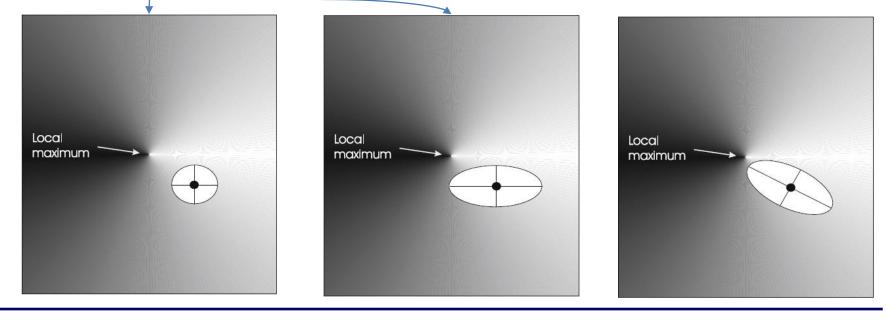
• Uncorrelated mutation with one step size  $\sigma$ :

$$x_i' = x_i + \sigma \cdot N_i(0,1)$$

• Uncorrelated mutation with *n* step sizes:

$$C x_i' = x_i + \sigma_i \cdot N_i(0,1)$$





### Real-valued representation: Nonuniform mutation

• Correlated mutation:

$$\delta \sim N(\mathbf{0}, \mathbf{I}) \xrightarrow{\text{scaling}} \mathbf{S}\delta = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \delta \xrightarrow{\text{rotation}} \mathbf{R}S\delta = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \mathbf{S}\delta$$
  
bivariate standard  
normal distribution  
$$\mathbf{Cov}(\delta') = \mathbf{R}S \operatorname{Cov}(\delta) (\mathbf{R}S)^{\mathrm{T}} = \begin{bmatrix} a^{2}\cos^{2}(\alpha) + b^{2}\sin^{2}(\alpha) & (a^{2} - b^{2})\sin(2\alpha)/2 \\ (a^{2} - b^{2})\sin(2\alpha)/2 & a^{2}\sin^{2}(\alpha) + b^{2}\cos^{2}(\alpha) \end{bmatrix}$$
$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$
$$\tan(2\alpha) = 2c_{12}/(c_{11} - c_{22})$$

Generalization: 
$$c_{ii} = \sigma_i^2$$
,  $c_{ij,i\neq j} = (\sigma_i^2 - \sigma_j^2) \tan(2\alpha_{ij})/2$ 

rotation angle for the *i*-th and *j*-th dimension

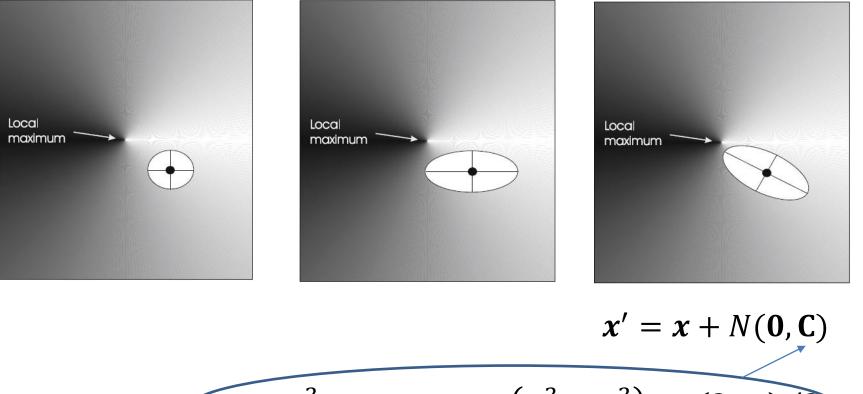
 $\checkmark$ 

# Real-valued representation: Nonuniform mutation

# Uncorrelated mutation with one step size $\sigma$

Uncorrelated mutation with *n* step sizes

Correlated mutation



$$c_{ii} = \sigma_i^2$$
,  $c_{ij,i\neq j} = (\sigma_i^2 - \sigma_j^2) \tan(2\alpha_{ij})/2$ 

- Self-adaptive mutation: Mutation step size  $\sigma$  is not set by user but coevolves with solution
- For example, the genotype is now  $(x_1, ..., x_n, \sigma)$ Self-adaptation  $\sigma \to \sigma'$  $x'_i = x_i + N_i(0, \sigma')$ 
  - The fitness of x' can be used to measure the goodness of both the offspring x' and the mutation step size  $\sigma'$
  - Why? Under different circumstances, different step sizes will behave differently

• Uncorrelated mutation with one step size  $\sigma$ :

$$(x_{1}, \dots, x_{n}, \sigma) \rightarrow (x'_{1}, \dots, x'_{n}, \sigma')$$
Self-adaptation  $\sigma' = \sigma \cdot e^{\tau \cdot N(0,1)}$ 

$$x'_{i} = x_{i} + \sigma' \cdot N_{i}(0,1)$$

$$\begin{bmatrix} \text{Typically } \tau \propto 1/\sqrt{n} \\ \sigma' < \epsilon_{0} \Rightarrow \sigma' = \epsilon_{0} \end{bmatrix}$$

- Satisfy these requirements:
  - Smaller modifications occur more often than larger ones
  - Greater than 0
  - The median is 1
  - Neural on average: equal prob. of drawing a value and its reciprocal

• Uncorrelated mutation with *n* step sizes:

$$(x_{1}, \dots, x_{n}, \sigma_{1}, \dots, \sigma_{n}) \rightarrow (x'_{1}, \dots, x'_{n}, \sigma'_{1}, \dots, \sigma'_{n})$$
Self-adaptation  $\sigma'_{i} = \sigma_{i} \cdot e^{\tau' \cdot N(0,1) + \tau \cdot N_{i}(0,1)}$ 

$$\begin{cases} \tau' \propto \frac{1}{\sqrt{2n}}, \tau \propto \frac{1}{\sqrt{2\sqrt{n}}} \\ \sigma'_{i} < \epsilon_{0} \Rightarrow \sigma'_{i} = \epsilon_{0} \end{cases}$$

*e<sup>τ'·N(0,1)</sup>*: overall change; *e<sup>τ·N<sub>i</sub>(0,1)</sup>*: coordinate-wise change
 The distribution is still lognormal, satisfying those requirements

• Correlated mutation:

$$(x_1, \dots, x_n, \sigma_1, \dots, \sigma_n, \alpha_1, \dots, \alpha_{n(n-1)/2})$$
  
  $\rightarrow (x'_1, \dots, x'_n, \sigma'_1, \dots, \sigma'_n, \alpha'_1, \dots, \alpha'_{n(n-1)/2})$ 

Self-adaptation  

$$\begin{aligned}
\sigma_i' &= \sigma_i \cdot e^{\tau' \cdot N(0,1) + \tau \cdot N_i(0,1)} \\
\alpha_j' &= \alpha_j + \beta \cdot N_j(0,1)
\end{aligned}$$

$$\begin{aligned}
\tau' &\propto \frac{1}{\sqrt{2n}}, \tau &\propto \frac{1}{\sqrt{2\sqrt{n}}} \\
\beta &\approx 5^o \\
\sigma_i' &< \epsilon_0 \Rightarrow \sigma_i' = \epsilon_0 \\
|\alpha_j'| &> \pi \Rightarrow \alpha_j' = \\
\alpha_j' - 2\pi \cdot sign(\alpha_j')
\end{aligned}$$

# Real-valued representation: Recombination

- **Discrete recombination:** Same as for binary representation, e.g., *m*-point crossover and uniform crossover
- Arithmetic recombination: Create offspring "between" parents

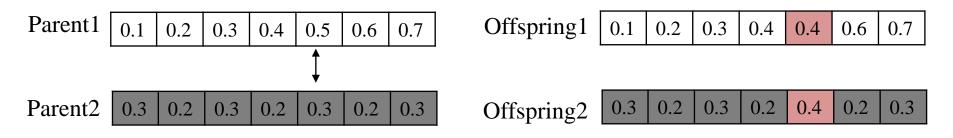
$$z_i = (1 - \alpha)x_i + \alpha y_i$$
, where  $\alpha \in [0, 1]$ 

• Blend recombination: Create offspring in a larger region  $z_i = (1 - \gamma)x_i + \gamma y_i$ , where  $\gamma = (1 + 2\alpha)u - \alpha$ ,  $u \in [0,1]$  Real-valued representation: Arithmetic recombination

• Single arithmetic recombination:

$$(x_1, \dots, x_n) \qquad (y_1, \dots, y_n)$$
  
Select a random k  
$$(x_1, \dots, x_{k-1}, \alpha y_k + (1 - \alpha) x_k, x_{k+1}, \dots, x_n)$$
  
$$(y_1, \dots, y_{k-1}, \alpha x_k + (1 - \alpha) y_k, y_{k+1}, \dots, y_n)$$

Example with  $\alpha = 0.5$ 

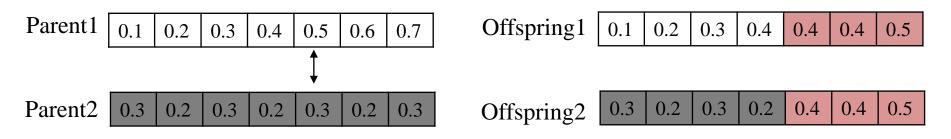


Real-valued representation: Arithmetic recombination

• Simple arithmetic recombination:

$$(x_1, \dots, x_n) \qquad (y_1, \dots, y_n)$$
  
Select a random k  
$$(x_1, \dots, x_{k-1}, \alpha y_k + (1 - \alpha) x_k, \dots, \alpha y_n + (1 - \alpha) x_n)$$
  
$$(y_1, \dots, y_{k-1}, \alpha x_k + (1 - \alpha) y_k, \dots, \alpha x_n + (1 - \alpha) y_n)$$

Example with  $\alpha = 0.5$ 



Real-valued representation: Arithmetic recombination

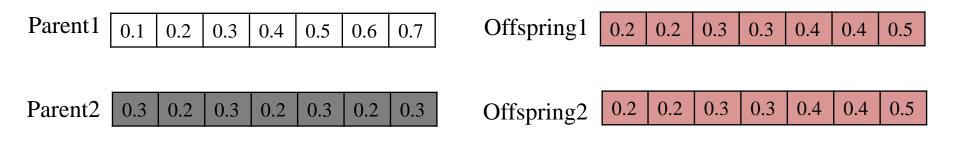
• Whole arithmetic recombination:

$$(x_{1}, ..., x_{n}) \qquad (y_{1}, ..., y_{n})$$

$$(\alpha y_{1} + (1 - \alpha)x_{1}, ..., \alpha y_{n} + (1 - \alpha)x_{n})$$

$$(\alpha x_{1} + (1 - \alpha)y_{1}, ..., \alpha x_{n} + (1 - \alpha)y_{n})$$

Example with  $\alpha = 0.5$ 



### Real-valued representation: Blend recombination

• Blend recombination: Create offspring in a larger region

$$z_{i} = (1 - \gamma)x_{i} + \gamma y_{i}, \text{ where } \gamma = (1 + 2\alpha)u - \alpha, u \in [0,1]$$

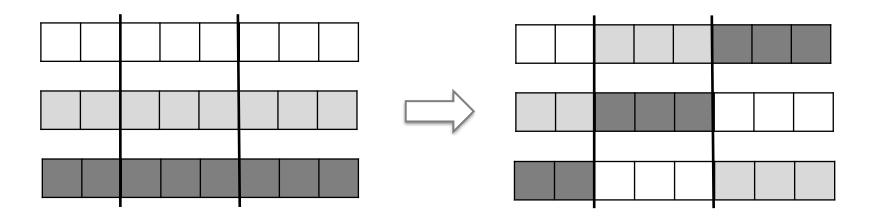
$$Q_{i} = x_{i} + (y_{i} - x_{i})(1 + 2\alpha)u - \alpha(y_{i} - x_{i})$$

$$Q_{i}$$

$$u \sim U(0,1) \Rightarrow z_{i} \sim U(x_{i} - \alpha d_{i}, y_{i} + \alpha d_{i})$$
How about  $z_{i} = (1 - \gamma)y_{i} + \gamma x_{i}$ ?

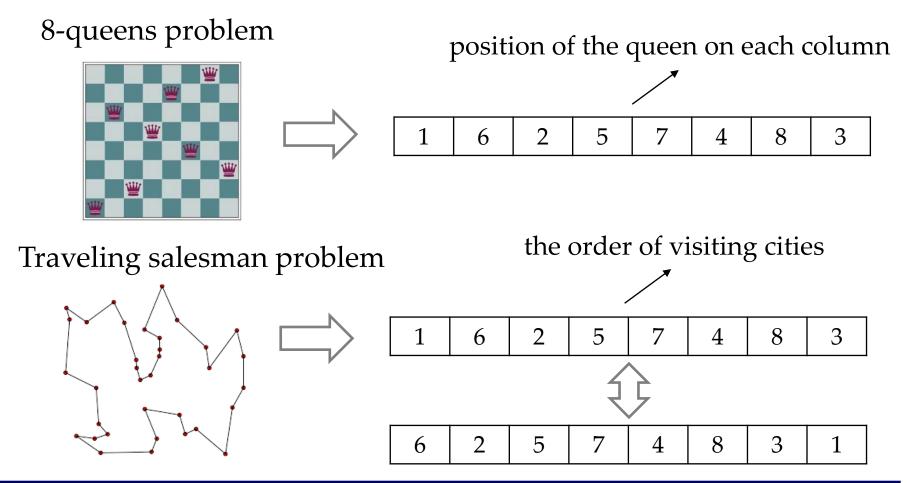
### Multi-parent recombination

- For example, diagonal crossover for *m* parents:
  - > Choose m 1 crossover points randomly
  - Compose *m* offspring from the segments of the parents in along a "diagonal", wrapping around



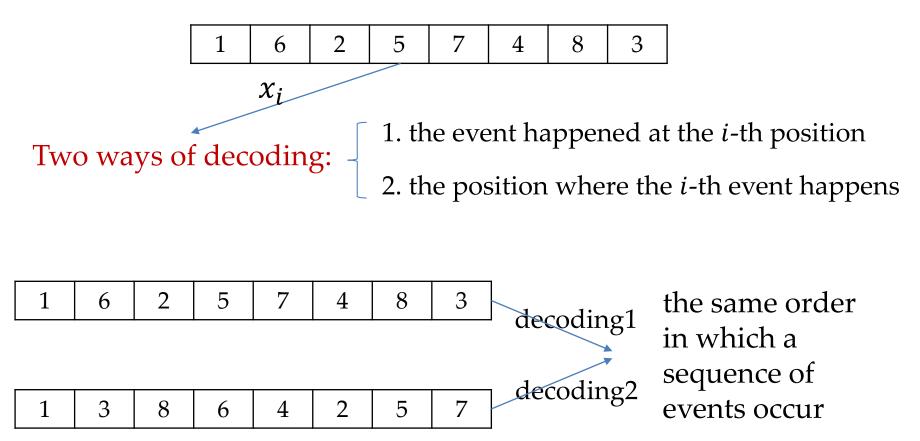
# Permutation representation

• Genotype space: a permutation of a fixed set of values



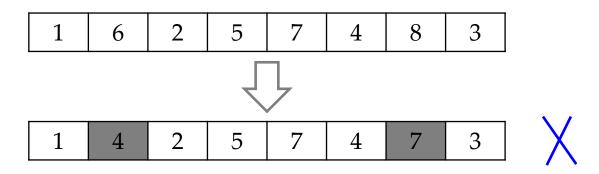
# Permutation representation

• Genotype space: a permutation of a fixed set of values



# Permutation representation: Mutation

• Mutation operators for binary, integer and real-valued representation will lead to inadmissible solutions

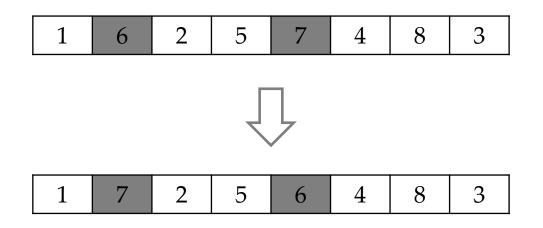


- Common mutation for permutation representation
  - Swap mutation
  - Insert mutation
  - Scramble mutation
  - Inversion mutation

The mutation probability now reflects the probability of applying mutation, rather than altering a single gene

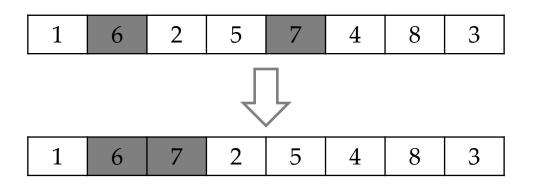
# Permutation representation: Swap mutation

- Swap mutation:
  - Select two positions randomly
  - Swap their values



# Permutation representation: Insert mutation

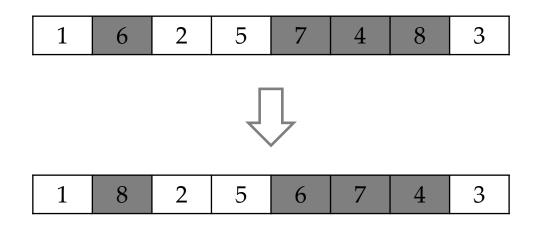
- Insert mutation:
  - Select two positions randomly
  - Move the second next to the first
  - Shift the rest along to accommodate



Preserve most of the order and the adjacency information

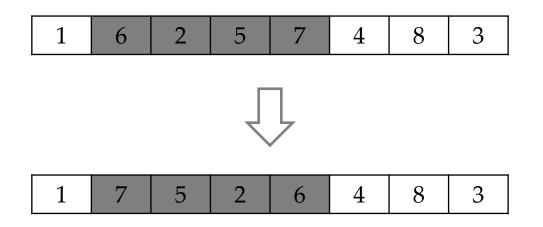
### Permutation representation: Scramble mutation

- Scramble mutation:
  - Select a subset of positions randomly
  - Rearrange their values randomly



### Permutation representation: Inversion mutation

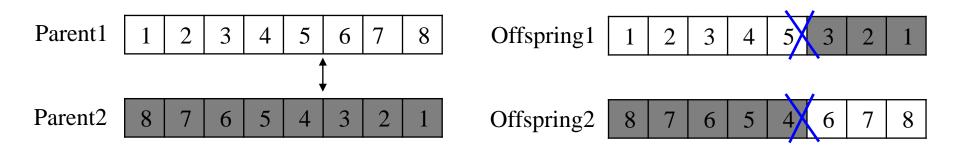
- Inversion mutation:
  - Select two positions randomly
  - Invert the values between them



Preserve most adjacency information (only breaks two links), but disruptive of order information

# Permutation representation: Recombination

• Recombination operators for binary, integer and real-valued representation will lead to inadmissible solutions



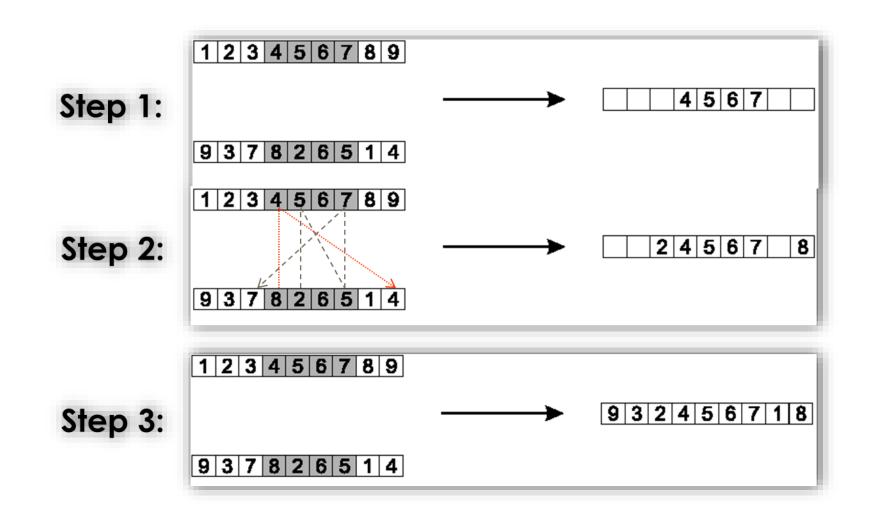
- Common recombination for permutation representation
  - Partially mapped crossover
  - Edge crossover
  - Order crossover
  - Cycle crossover

combine order or

 adjacency information from the two parents Permutation representation: Partially mapped crossover

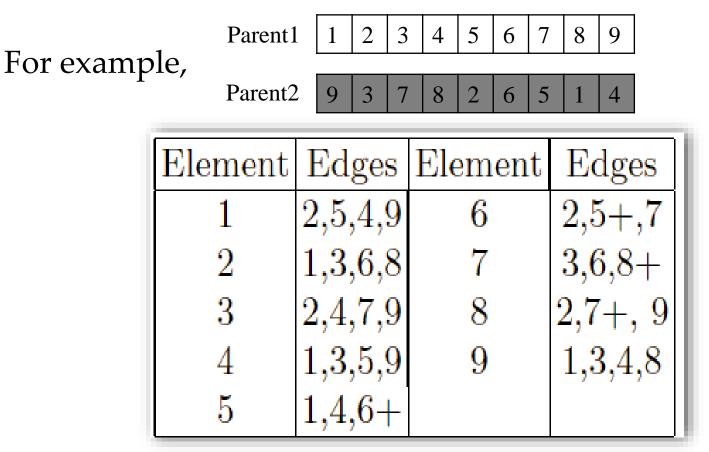
- Partially mapped crossover:
  - 1. Choose two crossover points randomly, and copy the segment between them from parent P1 into the first offspring
  - 2. Starting from the first crossover point, look for elements in that segment of P2 that have not been copied
  - 3. For each of these *i*, look in the offspring to see what element *j* has been copied in its place from P1
  - 4. Place *i* into the position occupied by *j* in P2, since we know that we will not be putting *j* there (as is already in offspring)
  - 5. If the place occupied by *j* in P2 has already been filled in the offspring by *k*, put *i* in the position occupied by *k* in P2
  - 6. Having dealt with the elements from the crossover segment, the rest of the first offspring can be filled from P2
  - 7. Create the second offspring analogously with parental roles reversed

Permutation representation: Partially mapped crossover



## Permutation representation: Edge crossover

• First step, construct a table listing which edges are present in the two parents; if an edge is common, mark with a +



## Permutation representation: Edge crossover

- Edge crossover: After constructing the edge table,
  - 1. Pick an initial element, entry, at random and put it in the offspring
  - 2. Set the variable current element = entry
  - 3. Remove all references to current element from the table
  - 4. Examine the list for current element:
    - ➢ If there is a common edge, pick that to be next element
    - Otherwise, pick the entry in the list which itself has the shortest list
    - ➢ Ties are split at random
  - 5. In the case of reaching an empty list: a new element is chosen at random

### Permutation representation: Edge crossover

Element	Edges	Element	Edges
1	$2,\!5,\!4,\!9$	6	2,5+,7
2	$1,\!3,\!6,\!8$	7	3,6,8+
3	$2,\!4,\!7,\!9$	8	2,7+,9
4	$1,\!3,\!5,\!9$	9	1,3,4,8
5	1,4,6+		

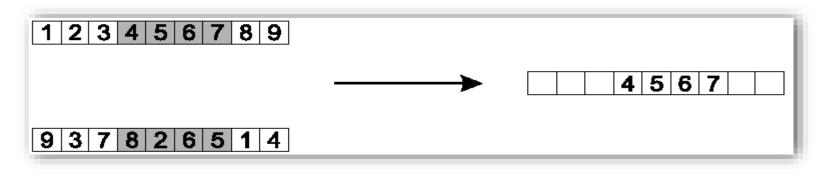
Choices	Element	Reason	Partial
	selected		result
All	1	Random	[1]
2,5,4,9	<b>5</b>	Shortest list	[1 5]
4,6	6	Common edge	[1 5 6]
2,7	2	Random choice (both have two items in list)	[1 5 6 2]
3,8	8	Shortest list	[1 5 6 2 8]
7,9	7	Common edge	[1 5 6 2 8 7]
3	3	Only item in list	[1 5 6 2 8 7 3]
4,9	9	Random choice	[15628739]
4	4	Last element	[156287394]

### Permutation representation: Order crossover

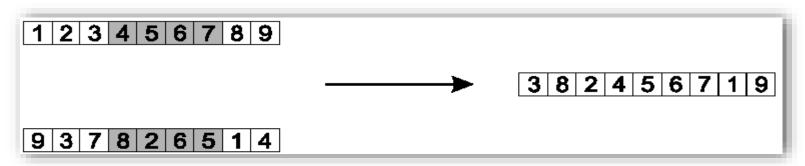
- Order crossover:
  - 1. Choose two crossover points randomly
  - 2. Copy the segment between them from the first parent to the first offspring
  - 3. Copy the numbers that are not in the segment, to the first offspring:
    - starting right from the second crossover point
    - using the order of the second parent
    - > and wrapping around at the end
  - 4. Create the second offspring in an analogous manner, with the parent roles inversed

### Permutation representation: Order crossover

1. Copy a randomly selected segment from the first parent



2. Copy the remaining numbers into the offspring in the order that they appear in the second parent



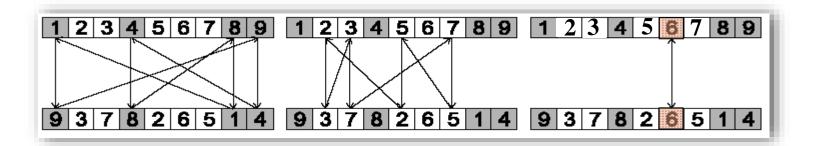
## Permutation representation: Cycle crossover

### • Cycle crossover:

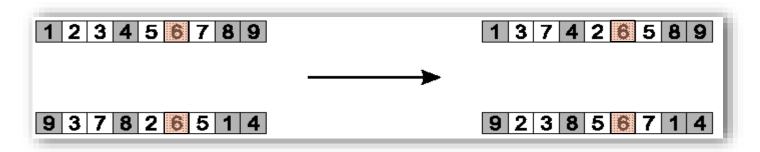
- 1. Divide the alleles into cycles
  - Start with the first unused position and allele of P1
  - Look at the allele in the same position in P2
  - ➢ Go to the position with the same allele in P1
  - Add this allele to the cycle
  - Repeat steps 2-4 until arriving at the first allele of P1
- 2. Create the offspring by selecting alternate cycles from each parent

### Permutation representation: Cycle crossover

1. Identify cycles



2. Copy alternate cycles into offspring

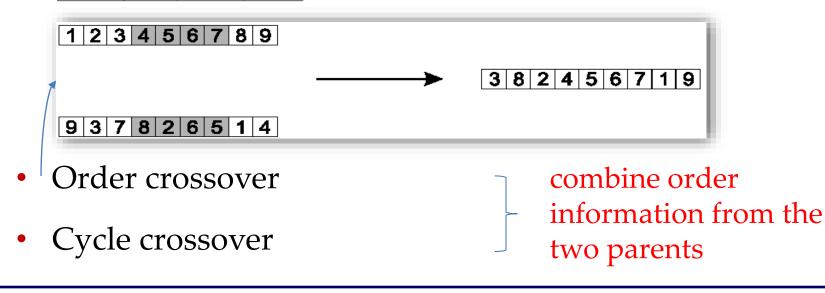


## Permutation representation: Recombination

- Partially mapped crossover
- Edge crossover

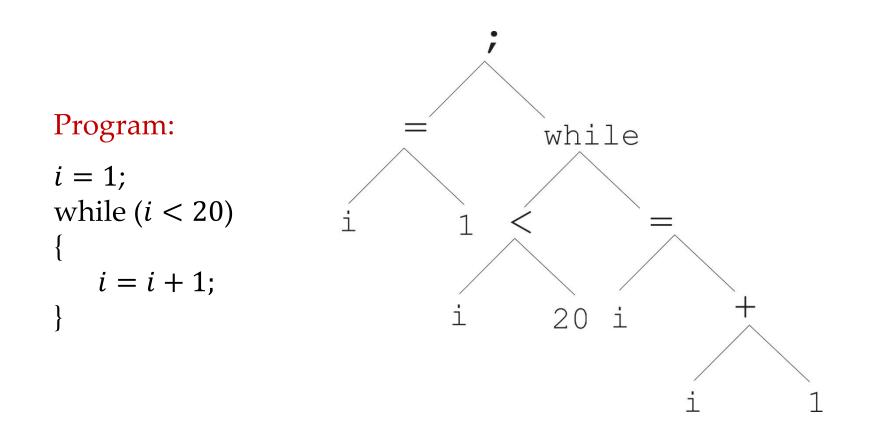
	Element	Edges	Element	Edges
	1	2,5,4,9	6	2,5+,7
1	2	1,3,6,8	7	3,6,8+
	3	2,4,7,9	8	2,7+,9
	4	$1,\!3,\!5,\!9$	9	1,3,4,8
	5	1,4,6+		

combine adjacency information from the two parents



## Tree representation

• Genotype space: a tree

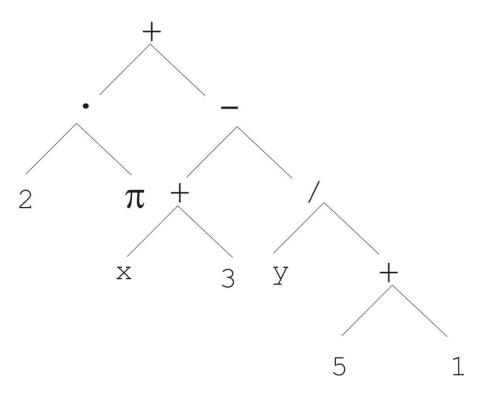


## Tree representation

• Genotype space: a tree

Arithmetic formula:

$$2 \cdot \pi + \left( (x+3) - \frac{y}{5+1} \right)$$



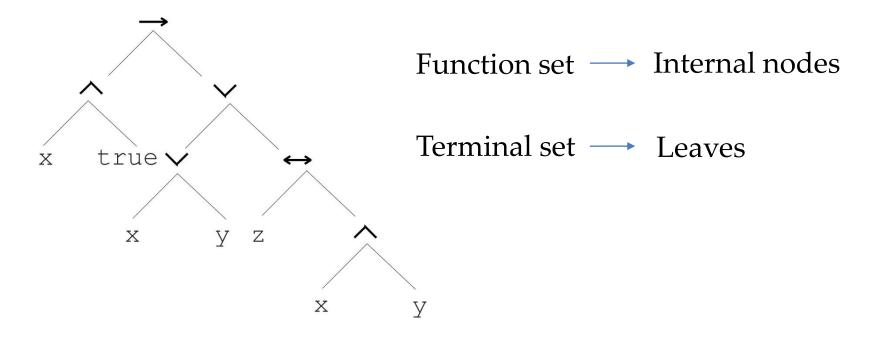
# Tree representation

• Genotype space: a tree

Logical formula:

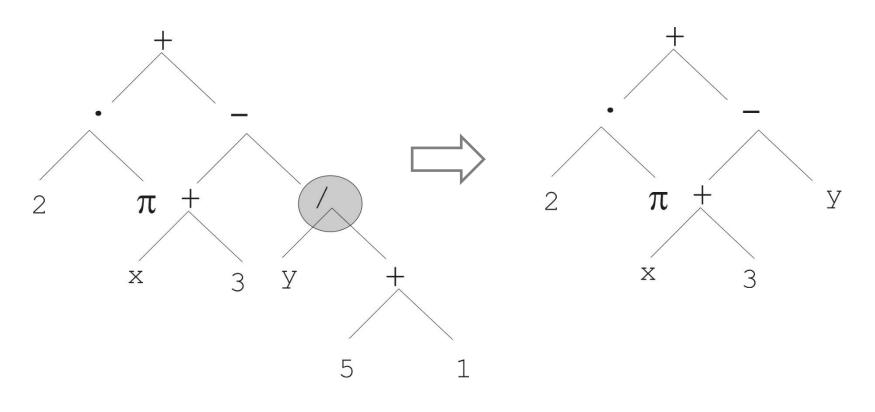
Variable structure

 $(x \land true) \rightarrow ((x \lor y) \lor (z \leftrightarrow (x \land y)))$ 



## Tree representation: Mutation

• Mutation: replace randomly chosen subtree by randomly generated tree



## Tree representation: Recombination

 Recombination: exchange two randomly chosen subtrees among the parents

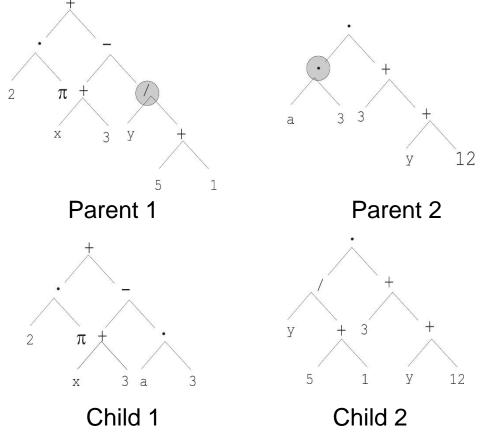
Assume uniform selection within each parent for exchanging

1

9

13

The probability:



- Decade long debate: which one is better
  - Evolutionary programming: originally without recombination
  - Genetic programming: originally without mutation
- Recombination is explorative, which can make a big jump to an area somewhere "in between" two (parent) areas
- Mutation is exploitative, which creates random small diversions, thereby staying near (in the area of ) the parent
- Now, it is good to have both in general

## Summary

- Binary representation
- Integer representation
- Real-valued representation
- Permutation representation
- Tree representation

RepresentationMutationRecombination

• A. E. Eiben and J. E. Smith. Introduction to Evolutionary Computing. Chapter 4.