Model-based Evolutionary Algorithms

Marcin Komarnicki

Agenda

- Optimization problem formulation
- Simple Genetic Algorithm
- Linkage Learning
- Dependency Structure Matrix
- Linkage Tree

Optimization

- Goal: best possible solution

- Definition:

$$
\begin{array}{ll}
\text{Optimize} & f(x) \\
\text{subject to} & x \in D \subseteq S \\
\text{where} & f : S \to R
\end{array}
$$

Binary optimization

- N-dimensional binary problem: $S = \{0, 1\}^N$
- Example for $N = 3$
  - $S = \{0, 1\}^3 = \{0, 1\} \times \{0, 1\} \times \{0, 1\}$
  - $x = [0, 1, 0]$

Deceptive trap function

- Benchmark problem
- Definition of k-order deceptive trap function:

$$t_k(u) = \begin{cases} k & ,u = k \\ n - u - 1 & ,u \neq k \end{cases}$$

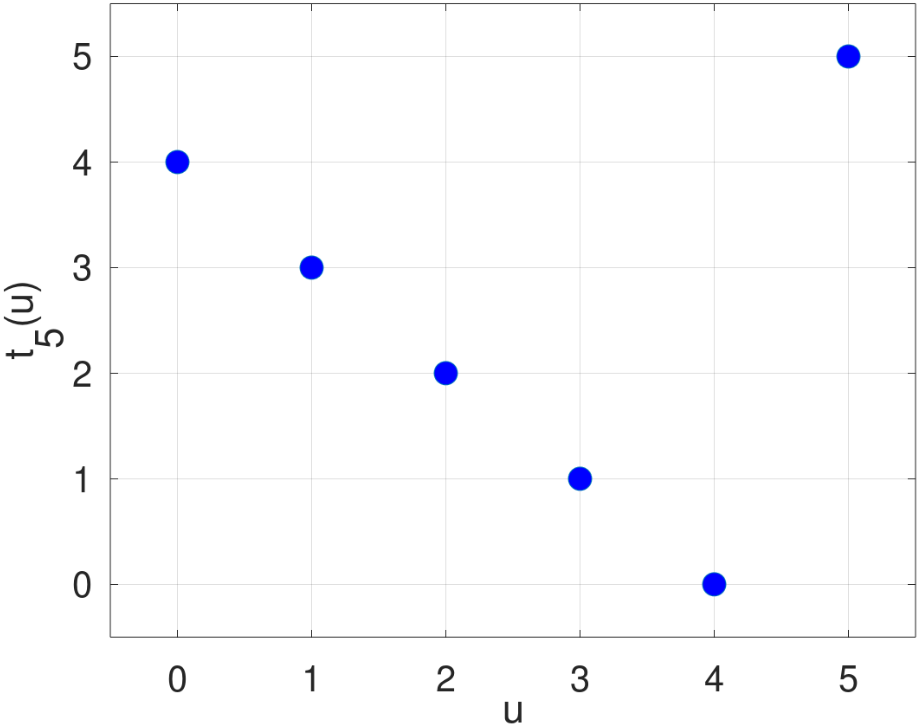
where u – unitation

- Unitation – number of $1s$:

$$x = [1, 0, 0, 1, 1] \rightarrow u = 3$$

5-order deceptive trap function

u	$t_5(u)$
0	4
1	3
2	2
3	1
4	0
5	5



$f([0,0,0,0,0]) = t_5(0) = 4$ (local optimum)

$f([1,1,1,1,1]) = t_5(5) = 5$ (global optimum)

Concatenated deceptive trap functions

- Sum of m decetpive trap functions
- E.g. sum of three 5-order deceptive trap functions:

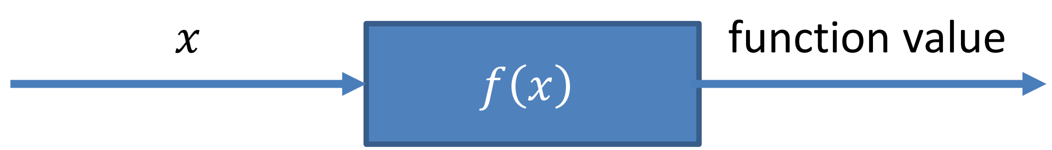
$$x = [1, 1, 1, 1, 1, \quad 0, 1, 0, 1, 1, \quad 0, 0, 0, 1, 1]$$

$$f(x) = t_5(5) + t_5(3) + t_5(2) = 5 + 1 + 2 = 8$$

- Fully separable problem

Simple Genetic Algorithm

- Black-box optimization method



- Population-based

I1:	0	1	1	0	0	1	1	0	0
I2:	1	0	0	1	0	0	1	1	0
I3:	1	0	0	1	1	0	0	0	1
I4:	0	1	1	0	0	1	1	1	0
I5:	0	1	0	1	0	1	1	1	1
I6:	0	1	0	0	0	0	1	1	0

Simple Genetic Algorithm

- Bit-flipping mutation

Individual before:	0	1	0	1	0	1	1	1	1
Mask:	-	-	-	-	+	-	+	-	+
Individual after:	0	1	0	1	1	1	0	1	0

Simple Genetic Algorithm

- Uniform crossover

P1:	1	1	1	0	0	0	1	1	1
P2:	0	0	0	1	1	1	0	0	0
M:	1	2	1	2	2	1	1	1	2

C1:	1	0	1	1	1	0	1	1	0
C2:	0	1	0	0	0	1	0	0	1
M:	1	2	1	2	2	1	1	1	2

Linkage Learning

- Information about gene dependencies
- Previous example:

  Genes **blue**, **gold**, and **green** are linked!

- Linked genes should be exchanged together

Dependency Structure Matrix (DSM)

- Square matrix
- Dependencies between some components
- $[d_{i,j}]_{1\leq i\leq n, 1\leq j\leq n}$ – relationship between the i^{th} and j^{th} components
- Greater value – greater dependency

X	$d_{1,2}$	$d_{1,3}$	$d_{1,4}$
$d_{2,1}$	X	$d_{2,3}$	$d_{2,4}$
$d_{3,1}$	$d_{3,2}$	X	$d_{3,4}$
$d_{4,1}$	$d_{4,2}$	$d_{4,3}$	X

Mutual information

- Definition

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \ln \frac{p(x,y)}{p(x)p(y)} \geq 0$$

where X and Y are random variables

- Minimum when X and Y are independent

$$p(x,y) = p(x)p(y) \Rightarrow \ln \frac{p(x,y)}{p(x)(y)} = \ln 1 = 0$$

DSM in EAs

- Rows and columns – genes
- Mutual information
  - Definition

$$I(G_i; G_j) = \sum_{g_i \in G_i} \sum_{g_j \in G_j} p_{i,j}(g_i, g_j) \ln \frac{p_{i,j}(g_i, g_j)}{p_i(g_i) p_j(g_j)}$$

    where G_i and G_j – the i^{th} and j^{th} gene

  - E.g. binary search space

$$G_i = G_j = \{0, 1\}$$

- Goal: find groups of dependent genes

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

- Rows – individuals
- Columns – genes
- 8 individuals and 9 genes

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$
\begin{aligned}
d_{2,3} &= I(G_2, G_3) \\
&= p_{2,3}(0,0)\ln\frac{p_{2,3}(0,0)}{p_2(0)p_3(0)} \\
&+ p_{2,3}(0,1)\ln\frac{p_{2,3}(0,1)}{p_2(0)p_3(1)} \\
&+ p_{2,3}(1,0)\ln\frac{p_{2,3}(1,0)}{p_2(1)p_3(0)} \\
&+ p_{2,3}(1,1)\ln\frac{p_{2,3}(1,1)}{p_2(1)p_3(1)}
\end{aligned}
$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$p_{2,3}(0,0) = \frac{2}{8} = 0.25$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$p_{2,3}(0,0) = \frac{2}{8} = 0.25$$

$$p_{2,3}(0,1) = \frac{2}{8} = 0.25$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$p_{2,3}(0,0) = \frac{2}{8} = 0.25$$

$$p_{2,3}(0,1) = \frac{2}{8} = 0.25$$

$$p_{2,3}(1,0) = \frac{1}{8} = 0.125$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$p_{2,3}(0,0) = \frac{2}{8} = 0.25$$

$$p_{2,3}(0,1) = \frac{2}{8} = 0.25$$

$$p_{2,3}(1,0) = \frac{1}{8} = 0.125$$

$$p_{2,3}(1,1) = \frac{3}{8} = 0.375$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$p_2(0) = \frac{4}{8} = 0.5$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$p_2(0) = \frac{4}{8} = 0.5$$

$$p_2(1) = \frac{4}{8} = 0.5$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$p_2(0) = \frac{4}{8} = 0.5$$

$$p_2(1) = \frac{4}{8} = 0.5$$

$$p_3(0) = \frac{3}{8} = 0.375$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$p_2(0) = \frac{4}{8} = 0.5$$

$$p_2(1) = \frac{4}{8} = 0.5$$

$$p_3(0) = \frac{3}{8} = 0.375$$

$$p_3(1) = \frac{5}{8} = 0.625$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

$$d_{2,3} = I(G_2, G_3)$$
$$= p_{2,3}(0,0) \ln \frac{p_{2,3}(0,0)}{p_2(0)p_3(0)}$$
$$+ p_{2,3}(0,1) \ln \frac{p_{2,3}(0,1)}{p_2(0)p_3(1)}$$
$$+ p_{2,3}(1,0) \ln \frac{p_{2,3}(1,0)}{p_2(1)p_3(0)}$$
$$+ p_{2,3}(1,1) \ln \frac{p_{2,3}(1,1)}{p_2(1)p_3(1)}$$
$$= 0.03$$

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

DSM								
X	0	0.03	0.13	0.13	0.13	0	0	0
0	X	0.03	0.13	0.13	0.13	0	0	0
0.03	0.03	X	0.38	0.38	0.38	0	0	0
0.13	0.13	0.38	X	0.69	0.69	0	0	0
0.13	0.13	0.38	0.69	X	0.69	0	0	0
0.13	0.13	0.38	0.69	0.69	X	0	0	0
0	0	0	0	0	0	X	0	0
0	0	0	0	0	0	0	X	0
0	0	0	0	0	0	0	0	X

DSM creation

Population

1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

DSM

X	0	0.03	0.13	0.13	0.13	0	0	0
0	X	0.03	0.13	0.13	0.13	0	0	0
0.03	0.03	X	0.38	0.38	0.38	0	0	0
0.13	0.13	0.38	X	0.69	0.69	0	0	0
0.13	0.13	0.38	0.69	X	0.69	0	0	0
0.13	0.13	0.38	0.69	0.69	X	0	0	0
0	0	0	0	0	0	X	0	0
0	0	0	0	0	0	0	X	0
0	0	0	0	0	0	0	0	X

DSM creation

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

DSM								
X	0	0.03	0.13	0.13	0.13	0	0	0
0	X	0.03	0.13	0.13	0.13	0	0	0
0.03	0.03	X	0.38	0.38	0.38	0	0	0
0.13	0.13	0.38	X	0.69	0.69	0	0	0
0.13	0.13	0.38	0.69	X	0.69	0	0	0
0.13	0.13	0.38	0.69	0.69	X	0	0	0
0	0	0	0	0	0	X	0	0
0	0	0	0	0	0	0	X	0
0	0	0	0	0	0	0	0	X

DSM creation

Population

1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

DSM

X	0	0.03	0.13	0.13	0.13	0	0	0
0	X	0.03	0.13	0.13	0.13	0	0	0
0.03	0.03	X	0.38	0.38	0.38	0	0	0
0.13	0.13	0.38	X	0.69	0.69	0	0	0
0.13	0.13	0.38	0.69	X	0.69	0	0	0
0.13	0.13	0.38	0.69	0.69	X	0	0	0
0	0	0	0	0	0	X	0	0
0	0	0	0	0	0	0	X	0
0	0	0	0	0	0	0	0	X

Distance measure

- Required by clustering algorithm
- Definition

$$D(G_i;G_j) = \frac{H(G_i;G_j) - I(G_i;G_j)}{H(G_i;G_j)}; \quad 0 \leq D(G_i;G_j) \leq 1$$

$$H(G_i;G_j) = -\sum_{g_i \in G_i} \sum_{g_j \in G_j} p_{i,j}(g_i, g_j) \ln p_{i,j}(g_i, g_j)$$

- If $H(G_i;G_j) = 0$ then $D(G_i;G_j) = 0$

Distance measure

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

Distance measure values								
X	1	0.97	0.9	0.9	0.9	1	1	1
1	X	0.97	0.9	0.9	0.9	1	1	1
0.97	0.97	X	0.61	0.61	0.61	1	1	1
0.9	0.9	0.61	X	0	0	1	1	1
0.9	0.9	0.61	0	X	0	1	1	1
0.9	0.9	0.61	0	0	X	1	1	1
1	1	1	1	1	1	X	0	0
1	1	1	1	1	1	0	X	0
1	1	1	1	1	1	0	0	X

Distance measure

Population

1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

Distance measure values

X	1	0.97	0.9	0.9	0.9	1	1	1
1	X	0.97	0.9	0.9	0.9	1	1	1
0.97	0.97	X	0.61	0.61	0.61	1	1	1
0.9	0.9	0.61	X	0	0	1	1	1
0.9	0.9	0.61	0	X	0	1	1	1
0.9	0.9	0.61	0	0	X	1	1	1
1	1	1	1	1	1	X	0	0
1	1	1	1	1	1	0	X	0
1	1	1	1	1	1	0	0	X

Distance measure

Population

1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

Distance measure values

X	1	0.97	0.9	0.9	0.9	1	1	1
1	X	0.97	0.9	0.9	0.9	1	1	1
0.97	0.97	X	0.61	0.61	0.61	1	1	1
0.9	0.9	0.61	X	0	0	1	1	1
0.9	0.9	0.61	0	X	0	1	1	1
0.9	0.9	0.61	0	0	X	1	1	1
1	1	1	1	1	1	X	0	0
1	1	1	1	1	1	0	X	0
1	1	1	1	1	1	0	0	X

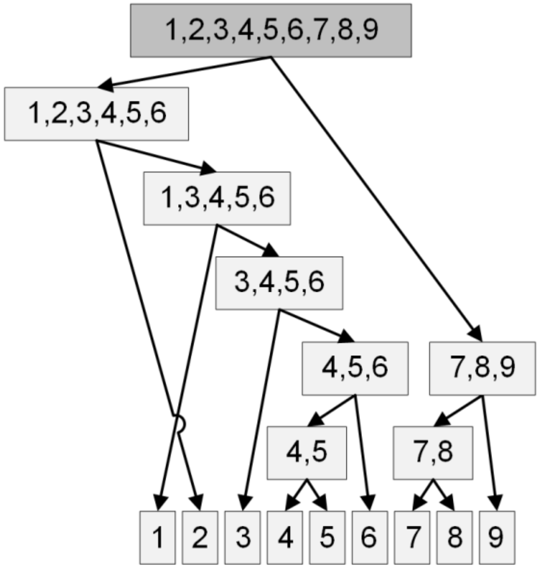
Distance measure

Population								
1	1	1	0	0	0	1	1	1
0	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	1	1	1	1	1	1
1	0	1	0	0	0	1	1	1
0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	0	0	1	1	1	1	1	1

Distance measure values								
X	1	0.97	0.9	0.9	0.9	1	1	1
1	X	0.97	0.9	0.9	0.9	1	1	1
0.97	0.97	X	0.61	0.61	0.61	1	1	1
0.9	0.9	0.61	X	0	0	1	1	1
0.9	0.9	0.61	0	X	0	1	1	1
0.9	0.9	0.61	0	0	X	1	1	1
1	1	1	1	1	1	X	0	0
1	1	1	1	1	1	0	X	0
1	1	1	1	1	1	0	0	X

Linkage tree

- Leaves = all gene indexes
- Nodes = clusters containing dependent gene indexes
- Hierarchical clustering algorithm



Hierarchical clustering algorithm

- Merge nearest clusters
- Distances between clusters
  - Clusters of size 1 – take value from table

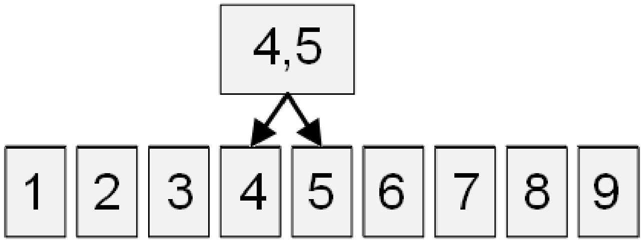
Distance measure values

X	1	0.97	0.9	0.9	0.9	1	1	1
1	X	0.97	0.9	0.9	0.9	1	1	1
0.97	0.97	X	0.61	0.61	0.61	1	1	1
0.9	0.9	0.61	X	0	0	1	1	1
0.9	0.9	0.61	0	X	0	1	1	1
0.9	0.9	0.61	0	0	X	1	1	1
1	1	1	1	1	1	X	0	0
1	1	1	1	1	1	0	X	0
1	1	1	1	1	1	0	0	X

  - Larger clusters – use reduction formula

$$D\left(C_k;(C_i \cup C_j)\right) = \frac{|C_i|}{|C_i| + |C_j|} D(C_k;C_i) + \frac{|C_j|}{|C_i| + |C_j|} D(C_k;C_j)$$

Linkage tree creation



Reduction formula

- Definition

$$D\left(C_k;\left(C_i \cup C_j\right)\right) = \frac{|C_i|}{|C_i| + \left|C_j\right|} D(C_k; C_i) + \frac{\left|C_j\right|}{|C_i| + \left|C_j\right|} D\left(C_k; C_j\right)$$

- Example

$$C_k = \{1\}, C_i = \{4\}, C_j = \{5\}$$

$$D(\{1\}; \{4\}) = 1, D(\{1\}; \{5\}) = 1$$

$$D(\{1\}; \{4, 5\}) = D(\{1\}; \{4\} \cup \{5\}) = \frac{1}{1+1} \cdot 1 + \frac{1}{1+1} \cdot 1 = 1$$

Linkage tree creation