Two countries, one goal, joint success!

Hungary-Romania Cross-Border Co-operation Programme 2007-2013

European Union
European Regional Development Fund

"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

APPLICATION OF GENETIC ALGORITHMS FOR STUDYING THE EVOLUTION OF BACTERIA POPULATION. SELECTION

Prof. dr. ing. Florin BLAGA

UNIVERSITY OF ORADEA
Romania

14 December 2012- Oradea

1. Introduction

- In selection the offspring producing individuals are chosen. The first step is fitness assignment.

- Each individual in the selection pool receives a reproduction probability depending on the own objective value and the objective value of all other individuals in the selection pool. This fitness is used for the actual selection step afterward

2. Rank-based fitness assignment

- In rank-based fitness assignment, the population is sorted according to the objective values.

- The fitness assigned to each individual depends only on its position in the individuals rank and not on the actual objective value.

2.1. Linear ranking

The fitness value for an individual:

$$Fitneß(Pos) = 2 - SP + 2 \cdot (SP - 1) \cdot \frac{(Pos - 1)}{(Nind - 1)}$$

- ***Nind*** **the number of individuals in the population;**
- ***Pos*** **the position of an individual in this population (least fit individual has *Pos* = 1, the fittest individual *Pos* = *Nind*) ;**
- ***SP*** **the selective pressure.**

2.2. Non-linear ranking

The fitness value for an individual:

$$Fitneß(Pos) = \frac{Nind \cdot X^{Pos-1}}{\sum_{i=1}^{Nind} X^{i-1}}$$

- **X is computed as the root of the polynomial:**

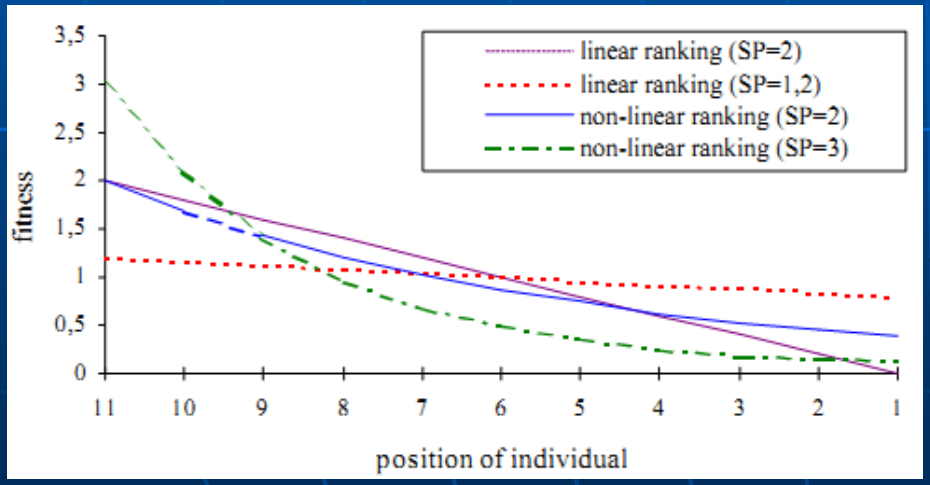
$$0 = (SP - Nind) \cdot X^{Nind-1} + SP \cdot X^{Nind-2} + \ldots + SP \cdot X + SP$$

- **Non-linear ranking allows values of selective pressure in [1, Nind - 2]**

"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

2.3. Comparison of linear and non-linear ranking

Fitness assignment for linear and non-linear ranking



Dependency of fitness value from selective pressure

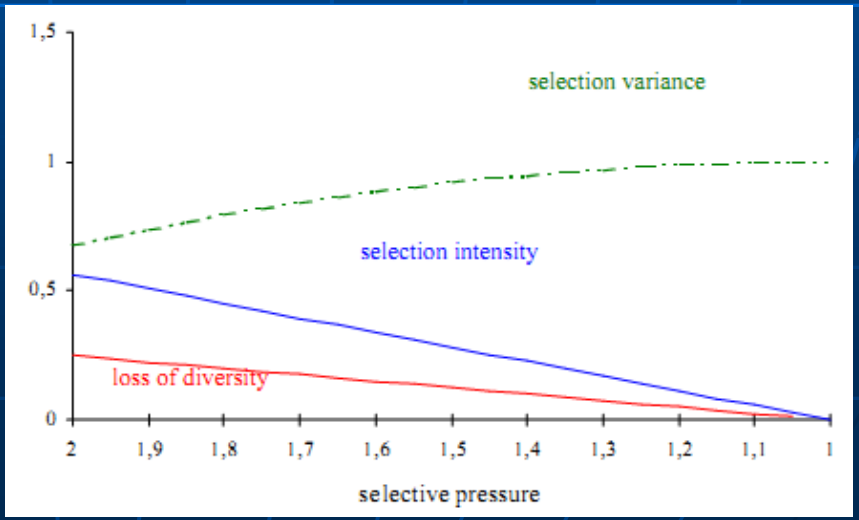
objective value	fitness value (parameter: selective pressure)				
	linear ranking		no ranking	non-linear ranking	
	2.0	1.1	1.0	3.0	2.0
1	2.0	1.10	1,0	3.00	2.00
3	1.8	1.08	1,0	2.21	1.69
4	1.6	1.06	1,0	1.62	1.43
7	1.4	1.04	1,0	1.99	1.21
8	1.2	1.02	1,0	0.88	1.03
9	1.0	1.00	1,0	0.65	0.87
10	0.8	0.98	1,0	0.48	0.74
15	0.6	0.96	1,0	0.35	0.62
20	0.4	0.94	1,0	0.26	0.53
30	0.2	0.92	1,0	0.19	0.45
95	0.0	0.90	1,0	0.14	0.38

"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

2.4. Analysis of linear ranking

- **Selection intensity:** $$SelInt_{LinRank}(SP) = (SP-1)\cdot\frac{1}{\sqrt{\pi}}$$
- **Loss of diversity :** $$LossDiv_{LinRank}(SP) = \frac{(SP-1)}{4}$$
- **Selection variance:** $$SelVar_{LinRank}(SP) = 1-\frac{(SP-1)^2}{\pi} = 1-SelInt_{LinRank}(SP)^2$$

- **Properties of linear ranking:**



3. Roulette wheel selection

- The simplest selection scheme is roulette-wheel selection, also called stochastic sampling with replacement

- This is a stochastic algorithm and involves the following technique:
  - The individuals are mapped to contiguous segments of a line, such that each individual's segment is equal in size to its fitness;
  - A random number is generated and the individual whose segment spans the random number is selected;
  - The process is repeated until the desired number of individuals is obtained (called mating population).

- This technique is analogous to a roulette wheel with each slice proportional in size to the fitness

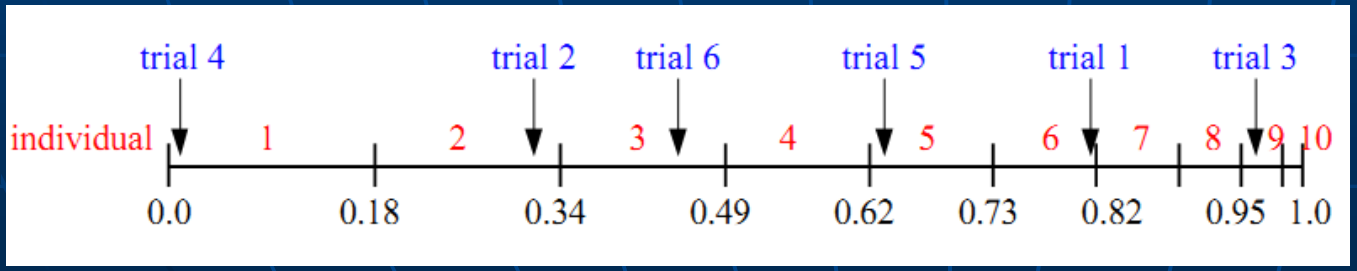
"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

- **The selection probability for 11 individuals, linear ranking and selective pres-sure of 2 together with the fitness value:**

Number of individual	1	2	3	4	5	6	7	8	9	10	11
fitness value	2.0	1.8	1.6	1.4	1.2	1.0	0.8	0.6	0.4	0.2	0.0
selection probability	0.18	0.16	0.15	0.13	0.11	0.09	0.07	0.06	0.03	0.02	0.0

- **For selecting the mating population the appropriate number of uniformly distributed random numbers (uniform distributed between 0.0 and 1.0) is independently generated.**
- **Sample of 6 random numbers: 0.81, 0.32, 0.96, 0.01, 0.65, 0.42.**

- **Roulette-wheel selection:**



1, 2, 3, 4, 6, 8

4. Local selection

- In local selection every individual resides inside a constrained environment called the local neighborhood.
- The neighborhood is defined by the structure in which the population is distributed.
- The neighborhood can be seen as the group of potential mating partners.
- The first step is the selection of the first half of the mating population uniform at random (or using one of the other mentioned selection algorithms, for example, stochastic universal sam-pling or truncation selection).
- Now a local neighborhood is defined for every selected individ-ual. Inside this neighborhood the mating partner is selected (best, fitness proportional, or uni-form at random)

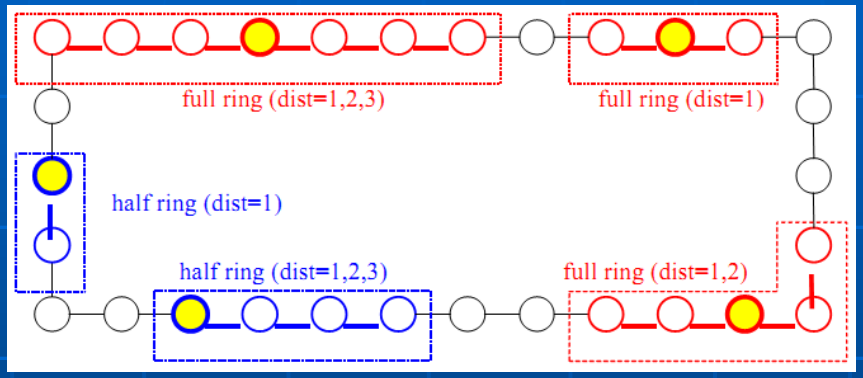
Two countries, one goal, joint success!

Hungary-Romania
Cross-Border Co-operation
Programme 2007-2013

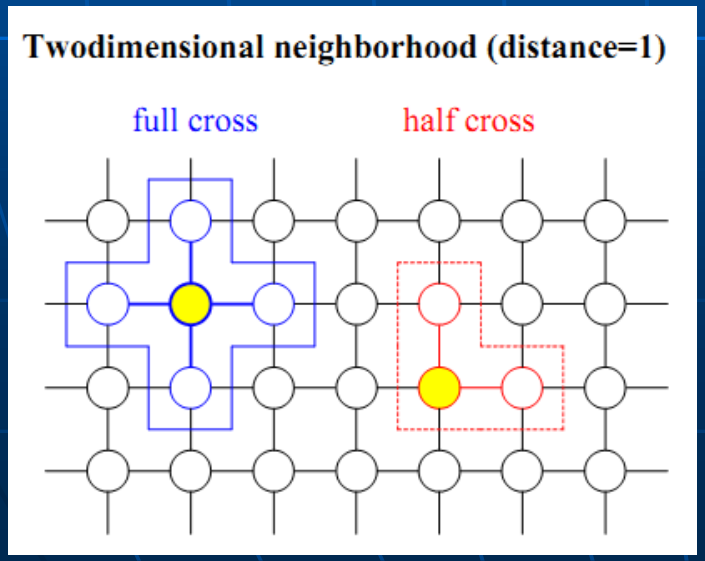
European Union
European Regional Development Fund

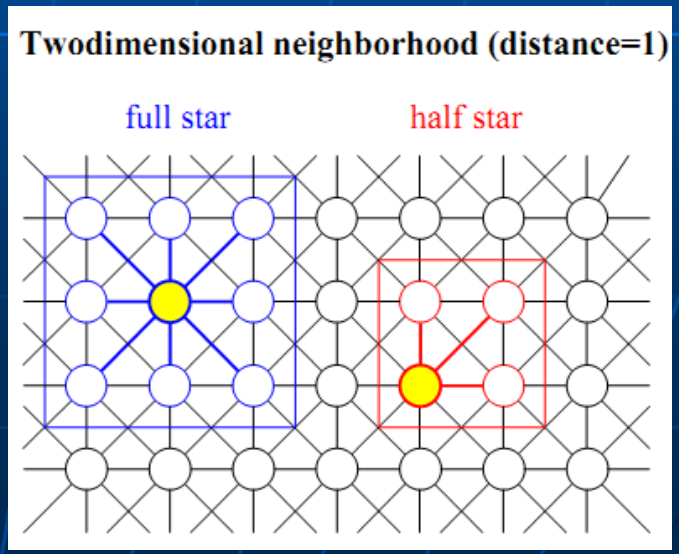
"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

Linear neighborhood: full and half ring



Two-dimensional neighborhood; left: full and half cross, right: full and half star





- **The size of the neighborhood determines the speed of propagation of information between the individuals of a population, thus deciding between rapid propagation or maintenance of a high diversity/variability in the population.**

Number of neighbors for local selection

structure of selection	distance	
	1	2
full ring	2	4
half ring	1	2
full cross	4	8 (12)
half cross	2	4 (5)
full star	8	24
half star	3	8

5. Truncation selection

- In truncation selection individuals are sorted according to their fitness. Only the best individu-als are selected for parents. These selected parents produce uniform at random offspring.
- The parameter for truncation selection is the truncation threshold Trunc. Trunc I ndicates the pro-portion of the population to be selected as parents and takes values ranging from 50%-10%.
- Individuals below the truncation threshold do not produce offspring.
- These selected parents produce uniform at random offspring.

The parameter for truncation selection is the truncation threshold *Trunc.* Trunc indicates the pro-portion of the population to be selected as parents and takes values ranging from 50%-10%.

"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

Relation between truncation threshold and selection intensity

truncation threshold	1%	10%	20%	40%	50%	80%
selection intensity	2.66	1.76	1.2	0.97	0.8	0.34

5.1 Analysis of truncation selection

Selection intensity:

$$SelInt_{Truncation}(Trunc) = \frac{1}{Trunc} \cdot \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{f_c^2}{2}}$$

Loss of diversity :

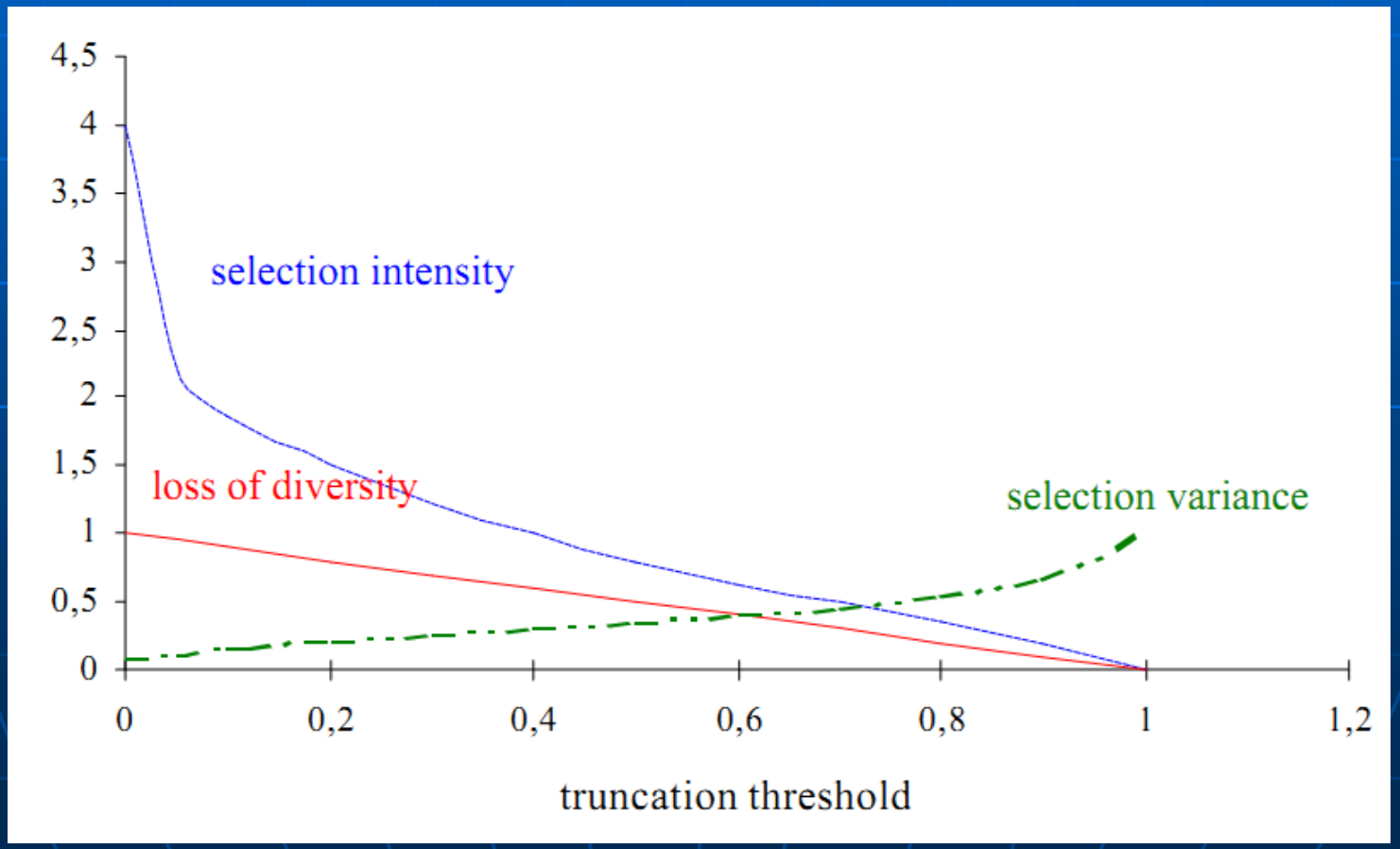
$$LossDiv_{Truncation}(Trunc) = 1 - Trunc$$

Selection variance :

$$SelVar_{Truncation}(Trunc) = 1 - SelInt_{Truncation}(Trunc) \cdot (SelInt_{Truncation}(Trunc) - f_c)$$

"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

Properties of truncation selection



6. Tournament selection

- In tournament selection a number Tour of individuals is chosen randomly from the population and the best individual from this group is selected as parent
- This process is repeated as often as individuals must be chosen
- These selected parents produce uniform at random offspring.
- The parameter for tournament selection is the tournament size Tour
- Tour takes values ranging from 2 to Nind (number of individuals in population)

"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

Relation between tournament size and selection intensity

tournament size	1	2	3	5	10	30
selection intensity	0	0.56	0.85	1.15	1.53	2.04

6.1 Analysis of tournament selection

Selection intensity

$$SelInt_{Turnier}(Tour) \approx \sqrt{2 \cdot \left(\ln(Tour) - \ln\left(\sqrt{4.14 \cdot \ln(Tour)}\right)\right)}$$

Loss of diversity

$$LossDiv_{Turnier}(Tour) = Tour^{\frac{-1}{Tour-1}} - Tour^{\frac{-Tour}{Tour-1}}$$

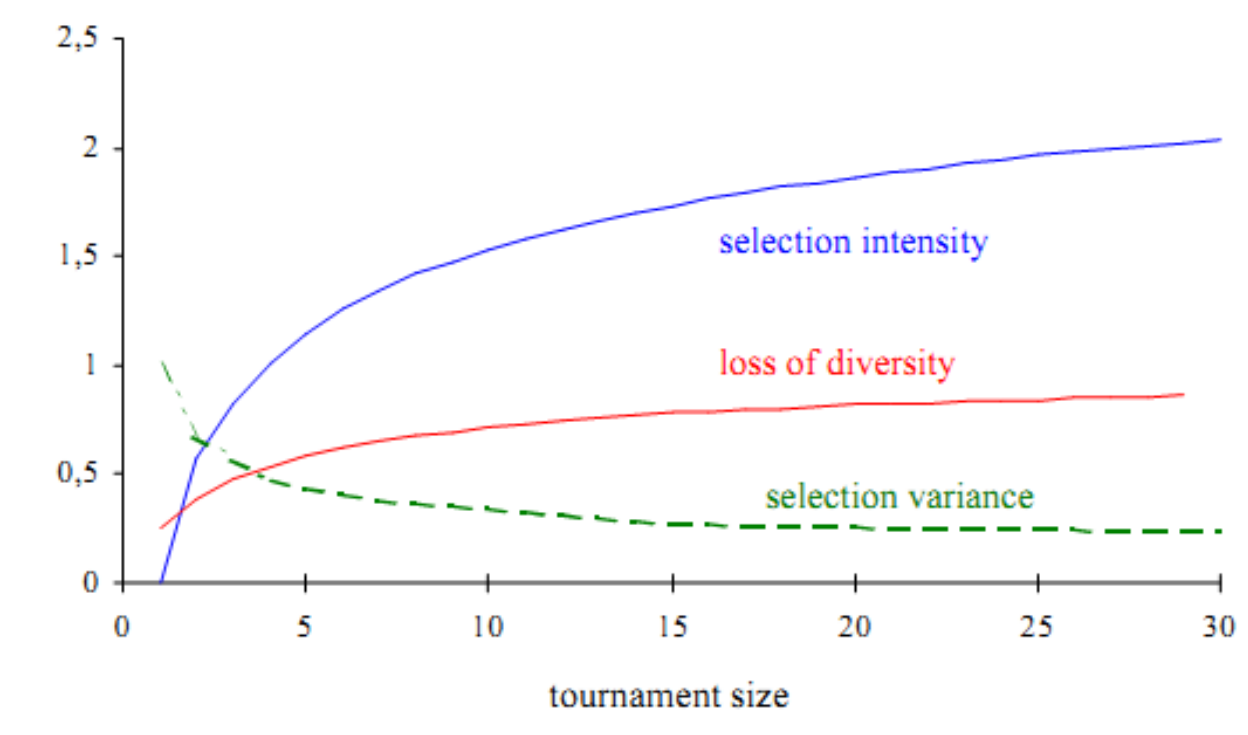
(About 50% of the population are lost at tournament size Tour=5).

Selection variance

$$SelVar_{Turnier}(Tour) \approx \frac{0.918}{\ln(1.186 + 1.328 \cdot Tour)}$$

"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

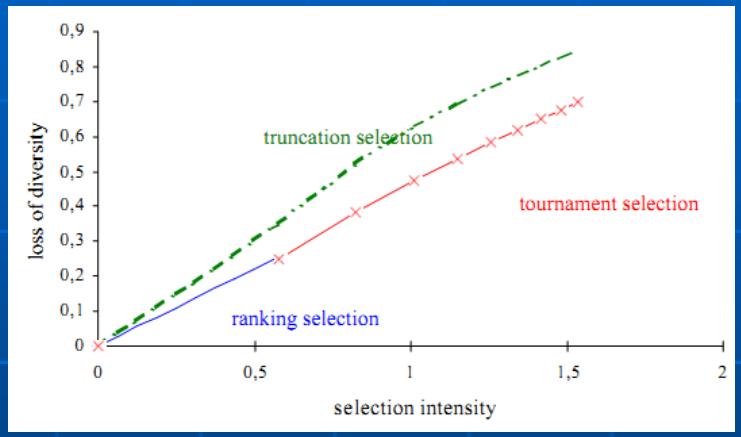
Properties of tournament selection

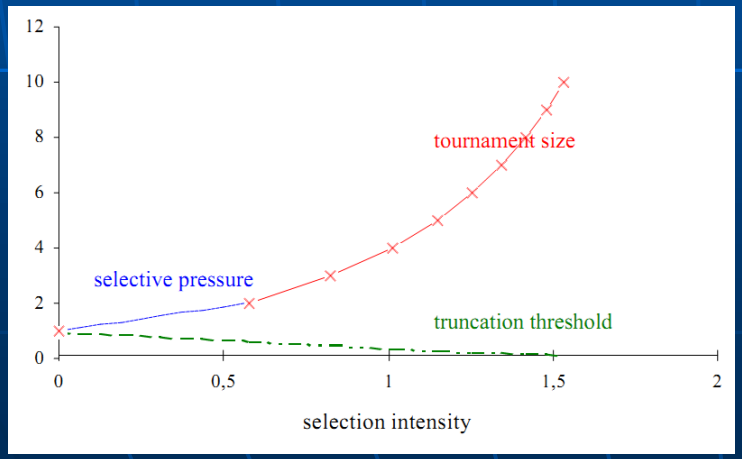


"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

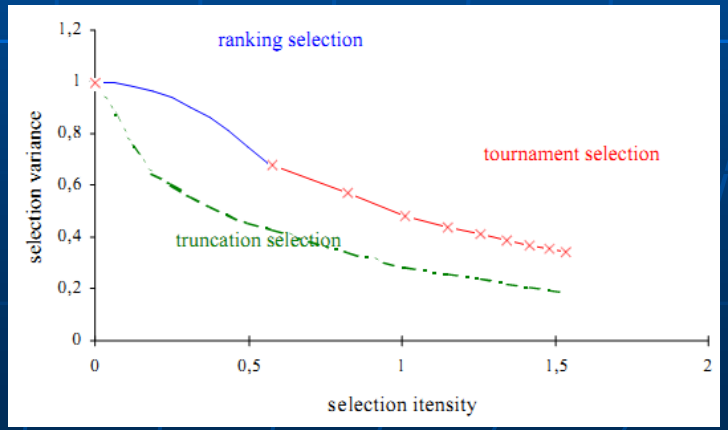
7. Comparison of selection schemes

Dependence of selection parameter on selection intensity





Dependence of loss of diversity on selection intensity



Dependence of selection variance on selection intensity

Two countries, one goal, joint success!

Hungary-Romania
Cross-Border Co-operation
Programme 2007-2013

European Union
European Regional Development Fund

"Realization of Hungarian-Romanian R&D Laboratory for the Development of Major Projects in Polluted Terrain Cleaning"
HURO/0802/100

Thank you !