eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Intelligent Systems Group

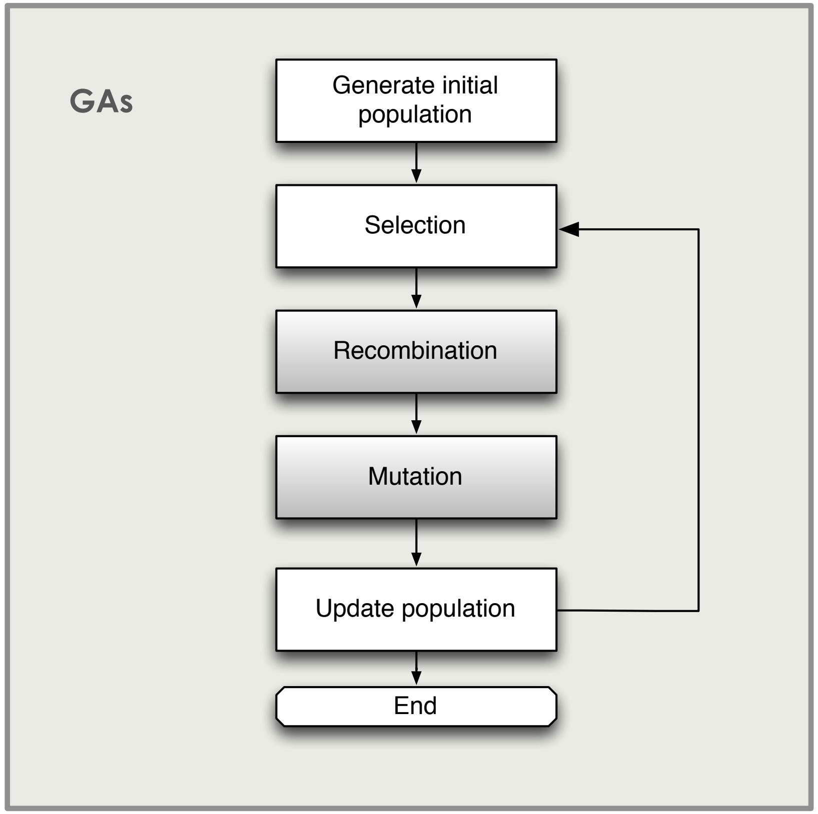
Permutation-based Optimization Problems and Estimation of Distribution Algorithms

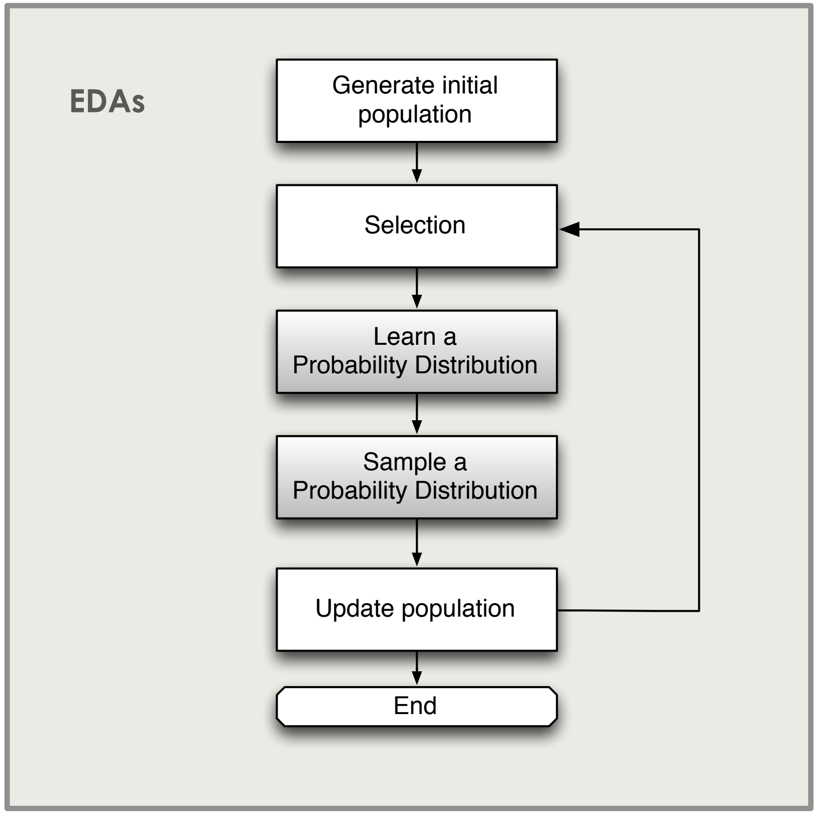
Josu Ceberio

Supervisors: Jose A. Lozano, Alexander Mendiburu

Estimation of Distribution Algorithms

- Estimation of Distribution Algorithms (EDAs)
  - Use probabilistic models to represent relevant information about the promising solutions.





Permutation-based Problems

- Combinatorial optimization problems
  - Minimize or maximize a given function

- Solutions are codified as permutations.
  - A bijection of the set $\{1, \ldots, n\}$ into itself, $\sigma = (\sigma(1), \ldots, \sigma(n))$

- Examples
  - Travelling Salesman Problem
  - Linear Ordering Problem
  - Quadratic Assignment Problem
  - Flowshop Scheduling Problem
  - …

Previously…

- Adaptation of existing EDAs
  - EDAs for **integer** domains.
  - EDAs for **real-valued** domains.

POOR PERFORMANCE

- Few efficient designs for permutation-based problems
  - **Edge** Histogram Based Sampling Algorithm (Tsutsui et al.)
  - **Node** Histogram Based Sampling Algorithm (Tsutsui et al.)

- Problems arise when we increase the size of the problem.

J. Ceberio, E. Irurozki, A. Mendiburu, J.A. Lozano. A review on Estimation of Distribution Algorithms in Permutation-based Combinatorial Optimization Problems. *Progress in Artificial Intelligence*. 2012

Probability Models on Rankings

- Models based on Low-order marginals
- Parametric Models
  - Thurstone Order Statistic model
  - Plackett-Luce model
  - Ranking induced by pair-comparisons:
    Mallows-Bradley-Terry models
  - Distance-based ranking models
  - Multi-stage ranking models
- Non-Parametric Models
  - Based on Kernels
  - Based on Fourier Transforms

Probability Models on Rankings

- Models based on Low-order marginals
- Parametric Models
  - Thurstone Order Statistic model
  - **Plackett-Luce model**
  - Ranking induced by pair-comparisons:
    Mallows-Bradley-Terry models
  - **Distance-based ranking models: Mallows model**
  - **Multi-stage ranking models: Generalized Mallows model**
- Non-Parametric Models
  - Based on Kernels
  - Based on Fourier Transforms

Distance-based ranking models

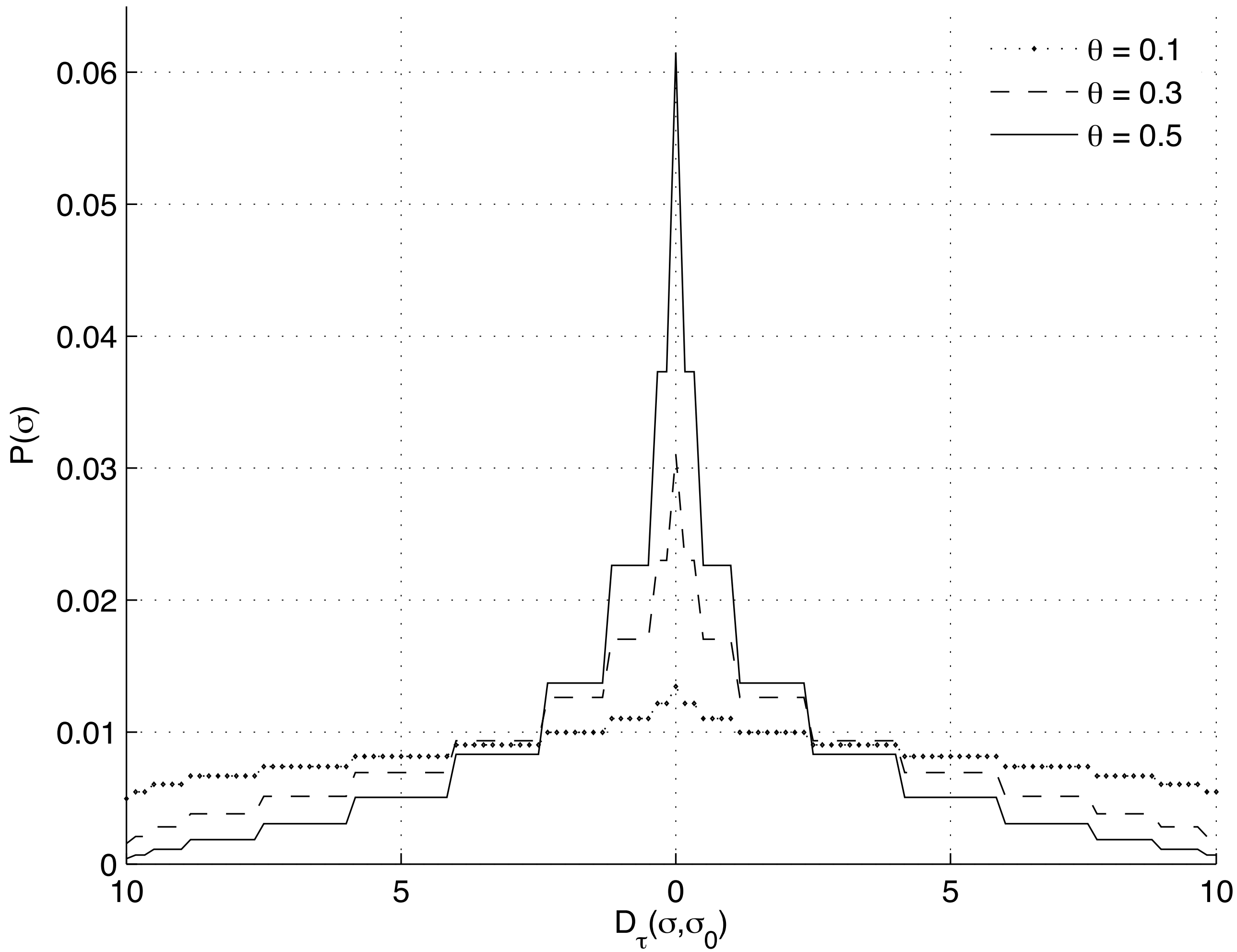
- The Mallows model is a distance-based exponential unimodal model.
- Analogous over permutations to the Gaussian distribution.
- Two parameters
  - Consensus ranking, σ_0
  - Spread parameter, θ
- The Mallows model

$$P(\sigma) = \frac{1}{\psi(\theta)} e^{-\theta D_\tau(\sigma,\sigma_0)}$$

The Mallows model

Definition

$$P(\sigma) = \frac{1}{\psi(\theta)} e^{-\theta D_\tau(\sigma,\sigma_0)}$$



σ_0 : Consensus ranking

θ : Spread parameter

$D_\tau(\sigma,\sigma_0)$: distance

$\psi(\theta)$: Partition function

Mallows EDA for PFSP

Good preliminary results of Mallows EDA for the PFSP-Makespan

Probability models on rankings look that have room in the framework of EDAs

Drawbacks:

- Unimodality
- Permutations at the same distance from σ_0 have the same probability

J. Ceberio, A. Mendiburu, J.A. Lozano. Introducing the Mallows Model on Estimation of Distribution Algorithms. *In Proceedings of International Conference International Conference on Neural Information Processing* 2011 *(ICONIP'11)*, pp461-470.

Multi-stage ranking models

Fligner and Verducci formulated that if the distance d can be written as

$$d(\pi, \sigma) = \sum_{i=1}^{n} S_i(\sigma, \pi)$$

and the S_i are independent under the uniform distribution,

then the Mallows model is factorizable. Even more it can be generalized to a n-parameters model:

$$P(\sigma) \propto exp(-\sum_{j=1}^{n} \theta_j S_j(\sigma, \sigma_0))$$

Where a different spread parameter θ_j is associated with each position in the permutation.

Multi-stage ranking models

The Kendall distance is additively decomposable,

thus, the Mallows model can be generalized to the Generalized Mallows model

We propose the Generalized Mallows EDA

Hybrid Generalized Mallows EDA for PFSP

Hybrid Generalized Mallows EDA is a efficient algorithm for solving the PFSP-TFT. Succeed in 152/220 instances.

The participation of the GM-EDA is essential for its success.

J. Ceberio, E. Irurozki, A. Mendiburu, J.A. Lozano. A Distance-based Ranking Model EDA for the Permutation Flowshop Scheduling Problem. *IEEE Transactions on Evolutionary Computation*. **Under Review.**

Plackett-Luce model

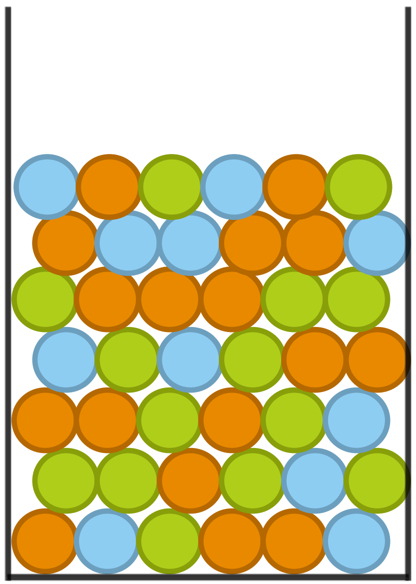
- The Plackett-Luce model

$$P(\pi) = \prod_{i=1}^{n-1} \frac{w_{\pi^{-1}(i)}}{\sum_{j=i}^{n} w_{\pi^{-1}(j)}}$$

- We have a vector of parameters $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$
- w_i is the probability of chosen object i
- The larger the parameter w_i in comparison to w_j, the higher the probability that the item i appears on a top rank

Plackett-Luce model

- Vase model interpretation



- A vase with infinite colored balls
- With known proportions of each color

  w_r

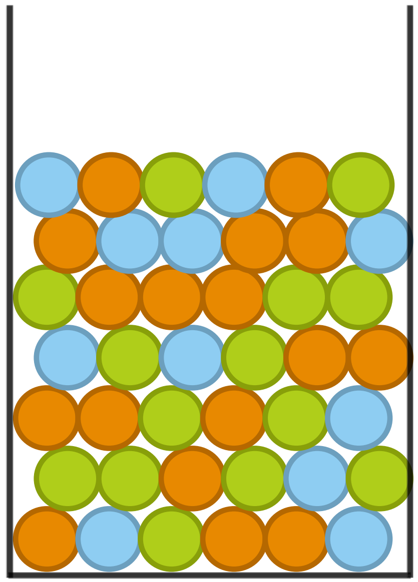
  w_g

  w_b

- Now draw balls from the vase until an ordering of coloured balls is obtained (a "permutation" of balls)

Plackett-Luce model

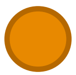
- Vase model interpretation



- Stage 1.
  - We draw a ball. And it is red.
  - The probability to extract a red ball at the first stage is:

$$\frac{w_r}{w_r + w_g + w_b}$$

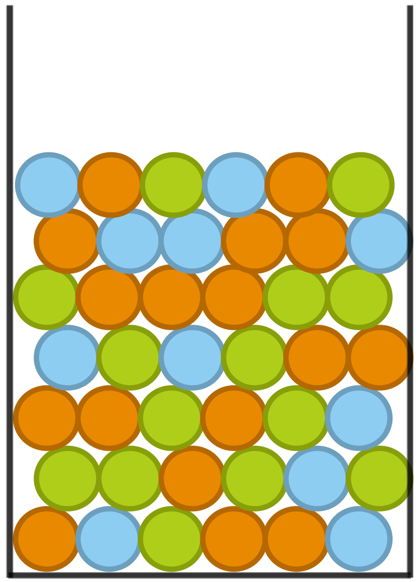
- Probability of the partial ranking

$\pi =$ 

$$P(\pi) = \frac{w_r}{w_r + w_g + w_b}$$

Plackett-Luce model

- Vase model interpretation



- Stage 2.
  - We draw another ball. And it is green.
  - The probability to extract a green ball from the remaining balls is:

$$\frac{w_g}{w_g + w_b}$$

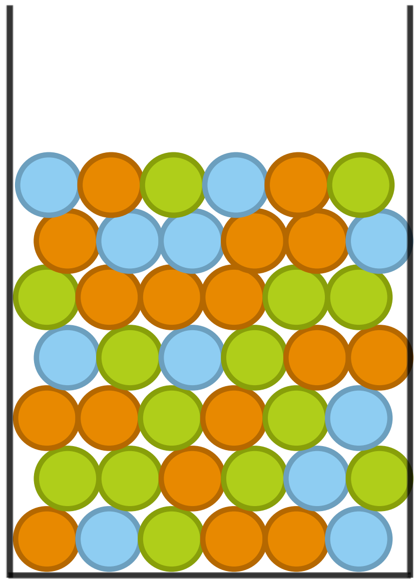
- Probability of the partial ranking

$\pi =$ 

$$P(\pi) = \frac{w_r}{w_r + w_g + w_b} \times \frac{w_g}{w_g + w_b}$$

Plackett-Luce model

- Vase model interpretation



- Stage 3.
  - We draw a last ball. It must be blue.
  - The probability to extract a blue when only we consider blue balls is:

$$\frac{w_b}{w_b}$$

$\pi =$ 

- Probability of the ranking

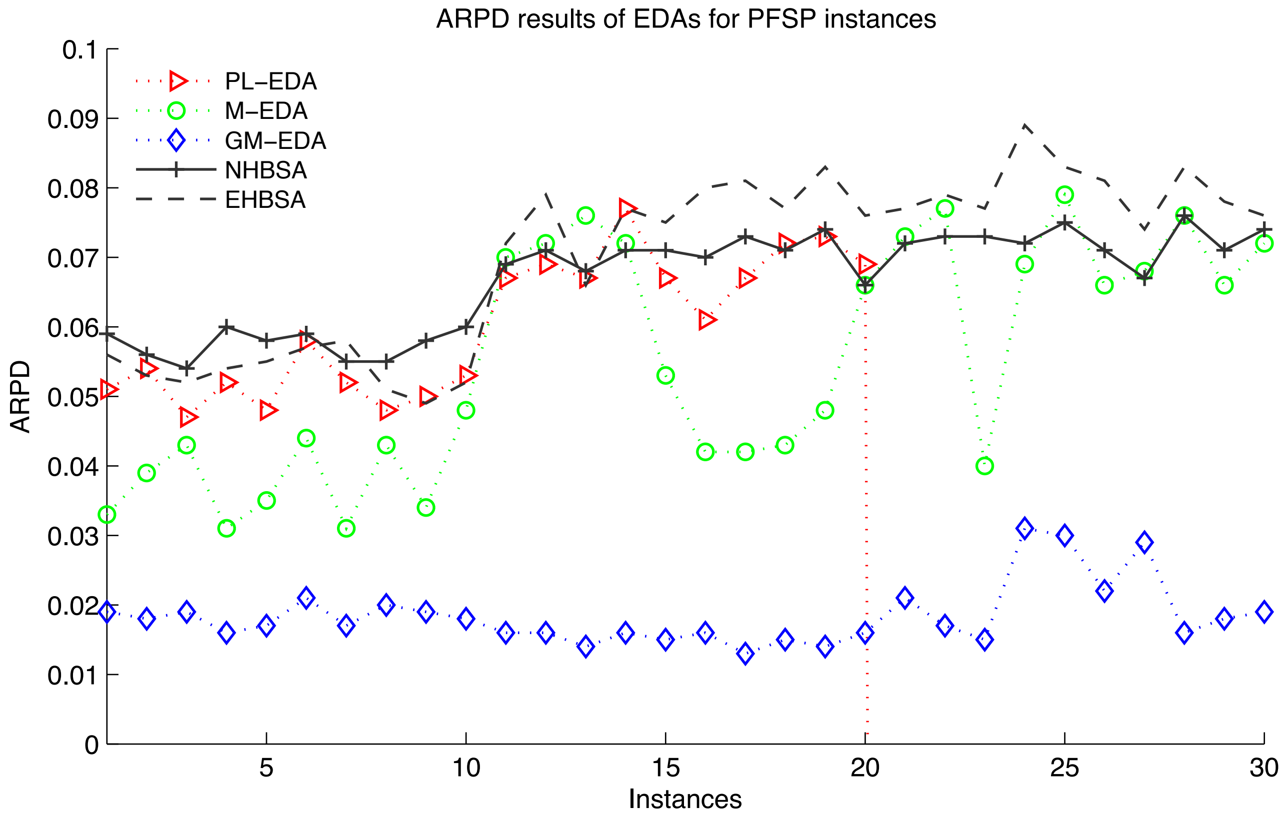
$$P(\pi) = \frac{w_r}{w_r + w_g + w_b} \times \frac{w_g}{w_g + w_b} \times \frac{w_b}{w_b}$$

Plackett-Luce model

We propose the Plackett-Luce EDA

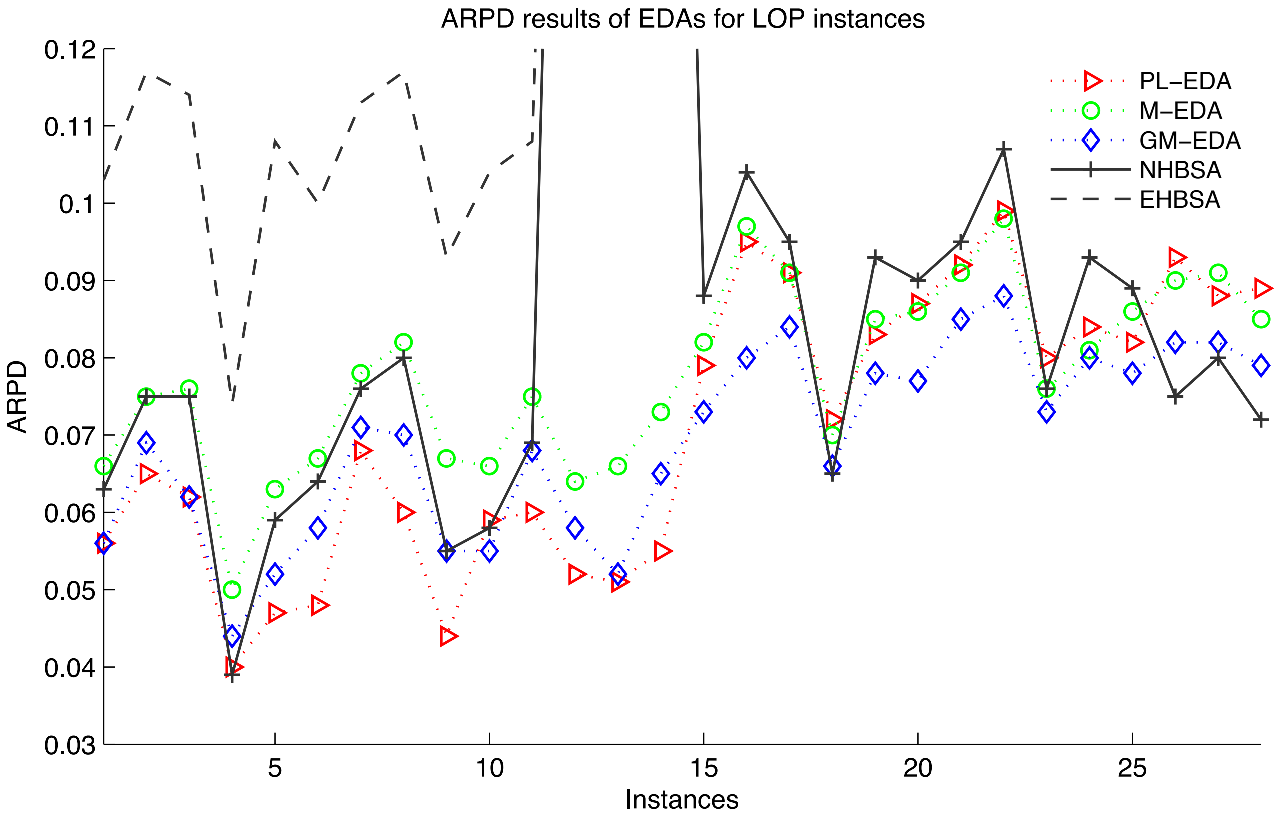
The L-decomposability property of the model suggests that it may perform very well for the Linear Ordering Problem

Some experiments for the PFSP



GM-EDA is by far the best algorithm.
Then M-EDA and next PL-EDA.

Some experiments for the LOP



PL-EDA is the best for n=150 instances.
GM-EDA is the best for n=250 instances.

Conclusions & Future work

- Which are the reasons for those performances?
  - LOP: L-decomposable models perform the best i.e. PL-EDA and GM-EDA.
  - PFSP: Distance-based models as GM-EDA and M-EDA are the most efficient EDAs.

- Future work
  - Extend the work to more problems and more models.
  - Study the elementary landscape decomposition of the problems for some neighborhoods.
  - Fourier decomposition of the objective function.

eman ta zabal zazu

Universidad del País Vasco Euskal Herriko Unibertsitatea

Intelligent Systems Group

Permutation-based Optimization Problems and Estimation of Distribution Algorithms

Josu Ceberio