The Plackett-Luce Ranking Model on Permutation-based Optimization Problems

Josu Ceberio, Alexander Mendiburu, Jose A. Lozano

Intelligent Systems Group
Department of Computer Science and Artificial Intelligence
University of the Basque Country UPV/EHU

2013 IEEE Congress on Evolutionary Computation (CEC 2013)
Cancun, Mexico, 22-23 April 2013

Outline

1. Permutation-based Optimization Problems
2. Estimation of Distribution Algorithms
3. Previous approaches
4. The Plackett-Luce Ranking Model
5. Some Experiments
6. Discussion

Outline

Permutation-based Optimization Problems

Definition

A specific subset of NP-Hard optimization problems.

Permutation-based Optimization Problems

Definition

Problems whose solution can be naturally represented as a **permutation**.

Permutation-based Optimization Problems

Examples

Travelling Salesman Problem

Permutation-based Optimization Problems

Examples

Travelling Salesman Problem

Quadratic Assignment Problem

Permutation-based Optimization Problems

Examples

Travelling Salesman Problem

Quadratic Assignment Problem

Linear Ordering Problem

Permutation-based Optimization Problems

Examples

Travelling Salesman Problem

Quadratic Assignment Problem

Linear Ordering Problem

Kemeny Ranking Problem

Permutation-based Optimization Problems

Examples

Travelling Salesman Problem

Quadratic Assignment Problem

Linear Ordering Problem

Kemeny Ranking Problem

Flowshop Scheduling Problem

Permutation-based Optimization Problems

Examples

Travelling Salesman Problem

Quadratic Assignment Problem

Linear Ordering Problem

Kemeny Ranking Problem

Flowshop Scheduling Problem

Permutation-based Optimization Problems

The Linear Ordering Problem (LOP)

Let an $n \times n$ matrix $H = [h_{ij}]$ be given.

Permutation-based Optimization Problems

The Linear Ordering Problem (LOP)

Let an $n \times n$ matrix $H = [h_{ij}]$ be given.

Determine a simultaneous permutation of rows and columns of H such that the sum of the entries above the main diagonal is maximized.

$$f(\pi) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} h_{\pi(i)\pi(j)}$$

Permutation-based Optimization Problems

The Linear Ordering Problem (LOP)

	1	2	3	4	5
1	0	16	11	15	7
2	21	0	14	15	9
3	26	23	0	26	12
4	22	22	11	0	13
5	30	28	25	24	0

(1, 2, 3, 4, 5) f= 138

Permutation-based Optimization Problems

The Linear Ordering Problem (LOP)

	1	2	3	4	5
1	0	16	11	15	7
2	21	0	14	15	9
3	26	23	0	26	12
4	22	22	11	0	13
5	30	28	25	24	0

(1, 2, 3, 4, 5) f= 138

Permutation-based Optimization Problems

The Linear Ordering Problem (LOP)

	1	2	3	4	5
1	0	16	11	15	7
2	21	0	14	15	9
3	26	23	0	26	12
4	22	22	11	0	13
5	30	28	25	24	0

(1, 2, 3, 4, 5) f= 138

	2	3	1	4	5
2	0	14	21	15	9
3	23	0	26	26	12
1	16	11	0	15	7
4	22	11	22	0	13
5	28	25	30	24	0

(2, 3, 1, 4, 5) f= 158

Permutation-based Optimization Problems

The Linear Ordering Problem (LOP)

	1	2	3	4	5
1	0	16	11	15	7
2	21	0	14	15	9
3	26	23	0	26	12
4	22	22	11	0	13
5	30	28	25	24	0

(1, 2, 3, 4, 5)   f= 138

Permutation-based Optimization Problems

The Linear Ordering Problem (LOP)

	1	2	3	4	5
1	0	16	11	15	7
2	21	0	14	15	9
3	26	23	0	26	12
4	22	22	11	0	13
5	30	28	25	24	0

(1, 2, 3, 4, 5) f= 138

	2	3	1	4	5
2	0	14	21	15	9
3	23	0	26	26	12
1	16	11	0	15	7
4	22	11	22	0	13
5	28	25	30	24	0

(2, 3, 1, 4, 5) f= 158

Permutation-based Optimization Problems

The Flowshop Scheduling Problem (FSP)

Given a set of n jobs and m machines
and known processing times p_{ij}

Given a set of n jobs and m machines
and known processing times p_{ij}

Find the sequence for scheduling jobs optimally.

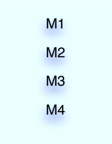
Permutation-based Optimization Problems

The Flowshop Scheduling Problem (FSP)

Given a set of n jobs and m machines
and known processing times p_{ij}

Find the sequence for scheduling jobs optimally.

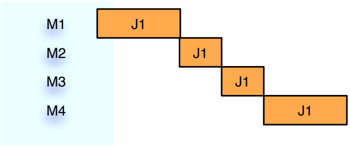
Example n=5 & m=4



Given a set of n jobs and m machines
and known processing times p_{ij}

Find the sequence for scheduling jobs optimally.

Example n=5 & m=4



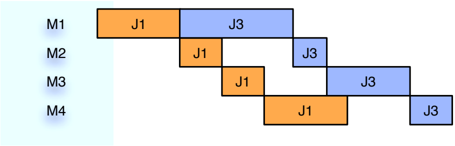
Permutation-based Optimization Problems

The Flowshop Scheduling Problem (FSP)

Given a set of n jobs and m machines
and known processing times p_{ij}

Find the sequence for scheduling jobs optimally.

Example n=5 & m=4



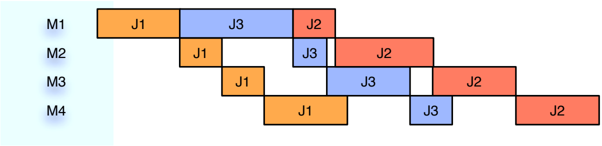
Permutation-based Optimization Problems

The Flowshop Scheduling Problem (FSP)

Given a set of n jobs and m machines
and known processing times p_{ij}

Find the sequence for scheduling jobs optimally.

Example n=5 & m=4



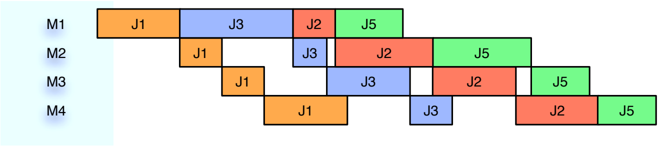
Permutation-based Optimization Problems

The Flowshop Scheduling Problem (FSP)

Given a set of n jobs and m machines
and known processing times p_{ij}

Find the sequence for scheduling jobs optimally.

Example n=5 & m=4



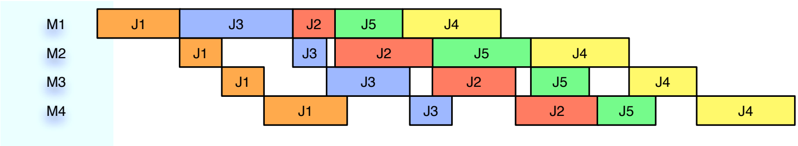
Permutation-based Optimization Problems

The Flowshop Scheduling Problem (FSP)

Given a set of n jobs and m machines
and known processing times p_{ij}

Find the sequence for scheduling jobs optimally.

Example n=5 & m=4



$$\pi = (1, 3, 2, 5, 4)$$

Outline

Evolutionary Algorithms

Estimation of Distribution Algorithms

Definition

Similar to Genetic Algorithms

Estimation of Distribution Algorithms

Definition

Given set of candidate individuals

Given set of candidate individuals

Learn a probability distribution

Estimation of Distribution Algorithms

Definition

Given set of candidate individuals

Learn a probability distribution

Sample the probability distribution to obtain the new population

Extensively used for a wide variety of problems.

Extensively used for a wide variety of problems.

But not so much on permutation-based optimization problems.

Outline

Standard representations are not effective due to the mutual exclusivity constraints associated with permutations.

Standard representations are not effective due to the mutual exclusivity constraints associated with permutations.

EDAs for discrete domains

Learn a probability distribution over a set $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$ where $\Omega_i = \{1, \ldots, r_i\}$ and $r_i \in \mathbb{N}, i = 1, \ldots, n$.

The sampling does not provide a permutation individual, but an individual in Ω.

Standard representations are not effective due to the mutual exclusivity constraints associated with permutations.

EDAs for the real domain

The sampled real vectors are transformed into permutations with the Random Keys strategy.

$$\begin{array}{c} (0.30, 0.10, 0.40, 0.20) \\ (0.25, 0.14, 0.35, 0.16) \\ (0.60, 0.20, 0.80, 0.40) \\ (0.33, 0.05, 0.35, 0.29) \\ (0.27, 0.15, 0.31, 0.20) \end{array} \quad \rightarrow \quad (3, 1, 4, 2)$$

High redundancy in the codification.

A distance-based exponential model.

A distance-based exponential model.

Analogous over permutations to the Gaussian distribution.

Existing approaches

Probability Models on Rankings - The Mallows model

A distance-based exponential model.

Analogous over permutations to the Gaussian distribution.

Two parameters:

Consensus ranking σ_0

Spread parameter θ

Existing approaches

Probability Models on Rankings - The Mallows model

A distance-based exponential model.

Analogous over permutations to the Gaussian distribution.

Two parameters:

- Consensus ranking σ_0
- Spread parameter θ

Kendall-τ distance: counts the number of pairwise disagreements.

Existing approaches

Probability Models on Rankings - The Mallows model

A distance-based exponential model.

Analogous over permutations to the Gaussian distribution.

Two parameters:

Consensus ranking σ_0

Spread parameter θ

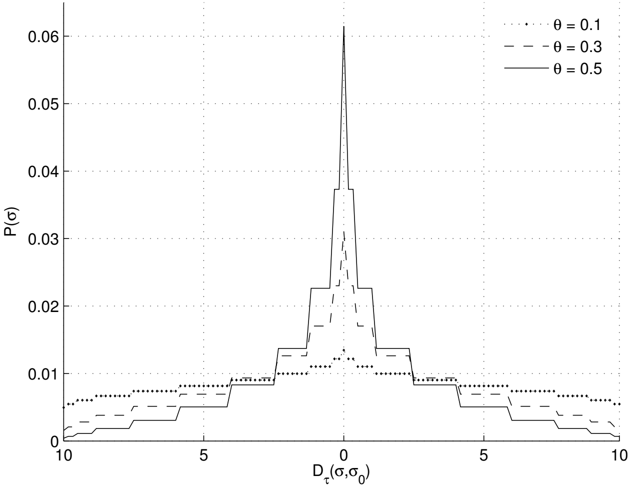
Kendall-τ distance: counts the number of pairwise disagreements.

Probability distribution:

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

Existing approaches

Probability Models on Rankings - The Mallows model



If the distance D can be written as

$$D(\pi, \sigma) = \sum_{j=1}^{n-1} V_j(\pi, \sigma)$$

and the V_j are independent under the uniform distribution

If the distance D can be written as

$$D(\pi, \sigma) = \sum_{j=1}^{n-1} V_j(\pi, \sigma)$$

and the V_j are independent under the uniform distribution, then the Mallows model can be generalized to a n-parameters model

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

If the distance D can be written as

$$D(\pi, \sigma) = \sum_{j=1}^{n-1} V_j(\pi, \sigma)$$

and the V_j are independent under the uniform distribution, then the Mallows model can be generalized to a n-parameters model

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

The Generalized Mallows model.

Outline

The Plackett-Luce ranking model is a multistage ranking model.

A vector of scores $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ defines the preference of each item to be ranked in top rank.

The Plackett-Luce Ranking Model

Definition

The probability that item i is chosen as the most preferred item among those in B is

$$P_B(i) = \frac{w_i}{\sum_{j \in B} w_j}$$

The Plackett-Luce Ranking Model

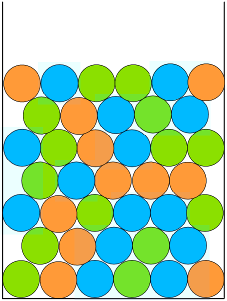
Definition

For every ranking π the Plackett-Luce model is given by

$$P(\pi|\boldsymbol{w}) = \prod_{i=1}^{n-1} \frac{w_{\pi\langle i\rangle}}{\sum_{j=i}^{n} w_{\pi\langle j\rangle}}$$

The Plackett-Luce Ranking Model

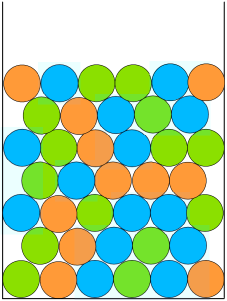
Vase model interpretation



A vase with infinite coloured balls.

The Plackett-Luce Ranking Model

Vase model interpretation

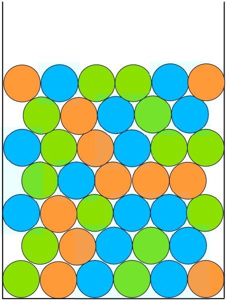


A vase with infinite coloured balls.

With known proportions of each colour.
(w_r, w_g, w_b).

The Plackett-Luce Ranking Model

Vase model interpretation



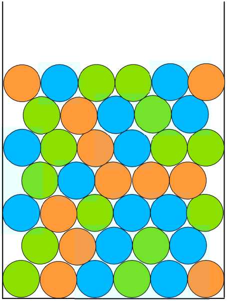
A vase with infinite coloured balls.

With known proportions of each colour.
(w_r, w_g, w_b).

Draw balls from the vase until an ordering of coloured balls is obtained.

Probability Models on Rankings

The Plackett-Luce model - Vase model interpretation



Stage 1

We draw a ball. And it is red.

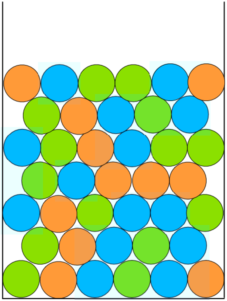
The probability to extract a red ball at this stage is:

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

$\pi =$ 

Probability Models on Rankings

The Plackett-Luce model - Vase model interpretation



Stage 1

We draw a ball. And it is red.

The probability to extract a red ball at this stage is:

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

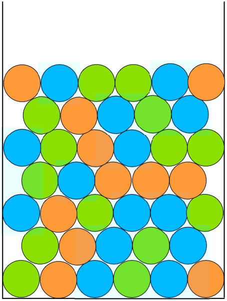
$\pi =$ 

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

Probability Models on Rankings

The Plackett-Luce model - Vase model interpretation

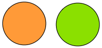
Stage 2



We draw another ball. And it is green.

The probability to extract a green ball from the remaining balls is:

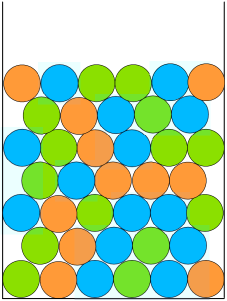
$$\frac{w_g}{w_g + w_b}$$

$\pi =$ 

Probability Models on Rankings

The Plackett-Luce model - Vase model interpretation

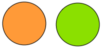
Stage 2



We draw another ball. And it is green.

The probability to extract a green ball from the remaining balls is:

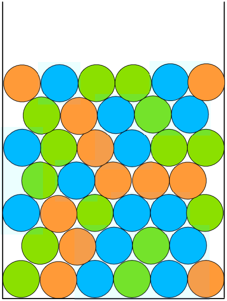
$$\frac{w_g}{w_g + w_b}$$

$\pi =$ 

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

Probability Models on Rankings

The Plackett-Luce model - Vase model interpretation



Stage 2

We draw the blue ball.

The probability to extract a blue ball is:

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

$\pi =$ 

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

Assuming a set $\boldsymbol{\pi} = \{\pi_1, \ldots, \pi_N\}$ and given $\boldsymbol{w} = (w_1, \ldots, w_n)$, the probability to observe the set,

$$P(\boldsymbol{\pi}|\boldsymbol{w}) = \prod_{k=1}^{N} \prod_{i=1}^{n-1} \frac{w_{\pi^k\langle i\rangle}}{\sum_{j=i}^{n} w_{\pi^k\langle j\rangle}}$$

The MLE of $\boldsymbol{w}$ is then given by those parameters that maximize the equation above, or equivalently the log-likelihood function

$$\ell(\boldsymbol{w}) = \sum_{k=1}^{N} \sum_{i=1}^{n-1} \left(\ln w_{\pi^k\langle i\rangle} - \ln \sum_{j=i}^{n} w_{\pi^k\langle j\rangle} \right)$$

Assuming a set $\boldsymbol{\pi} = \{\pi_1, \ldots, \pi_N\}$ and given $\boldsymbol{w} = (w_1, \ldots, w_n)$, the probability to observe the set,

$$P(\boldsymbol{\pi}|\boldsymbol{w}) = \prod_{k=1}^{N} \prod_{i=1}^{n-1} \frac{w_{\pi^k\langle i\rangle}}{\sum_{j=i}^{n} w_{\pi^k\langle j\rangle}}$$

The MLE of $\boldsymbol{w}$ is then given by those parameters that maximize the equation above, or equivalently the log-likelihood function

$$\ell(\boldsymbol{w}) = \sum_{k=1}^{N} \sum_{i=1}^{n-1} \left(\ln w_{\pi^k\langle i\rangle} - \ln \sum_{j=i}^{n} w_{\pi^k\langle j\rangle} \right)$$

Minorization-Maximization (MM) algorithm (Hunter 2004)

Outline

Algorithms

- Histogram-based EDAs: NHBSA and EHBSA.
- Mallows EDA (MaEDA) and Generalized Mallows EDA (GMEDA)
- Our proposal: Plackett-Luce EDA (PLEDA)

Problems

- LOP: 14 instances of size 150 and 14 instances of size 250.
- FSP: 10 instances of configurations 100×20, 200×20 and 250×20.

Average Relative Percentage Deviation (ARPD) of 20 repetitions of each of the algorithms is calculated.

Size	Instance	PLEDA	MaEDA	GMEDA	NHBSA	EHBSA
150	t65b11xx	0.056	0.072	0.065	0.064	0.107
	t65d11xx	0.064	0.081	0.074	0.078	0.123
	t65f11xx	0.062	0.081	0.074	0.077	0.121
	t65l11xx	0.041	0.058	0.052	0.041	0.082
	t65n11xx	0.052	0.070	0.065	0.060	0.113
	t70b11xx	0.058	0.074	0.068	0.065	0.109
	t70d11xx	0.067	0.084	0.076	0.076	0.117
	t70f11xx	0.069	0.088	0.078	0.082	0.123
	t70l11xx	0.052	0.074	0.069	0.054	0.102
	t70n11xx	0.056	0.073	0.066	0.061	0.112
	t75d11xx	0.064	0.078	0.073	0.071	0.115
	t75e11xx	0.048	0.069	0.062	0.300	0.300
	t75k11xx	0.055	0.074	0.064	0.292	0.291
	t75n11xx	0.060	0.078	0.071	0.302	0.299
250	t65b11xx	0.074	0.086	0.080	0.088	0.146
	t65d11xx	0.083	0.099	0.090	0.105	0.161
	t65f11xx	0.078	0.094	0.086	0.097	0.159
	t65l11xx	0.061	0.075	0.071	0.068	0.132
	t65n11xx	0.075	0.092	0.084	0.094	0.162
	t70b11xx	0.074	0.088	0.081	0.092	0.157
	t70d11xx	0.080	0.095	0.086	0.097	0.157
	t70f11xx	0.087	0.101	0.094	0.109	0.162
	t70l11xx	0.066	0.082	0.076	0.078	0.150
	t70n11xx	0.074	0.091	0.081	0.094	0.162
	t75d11xx	0.075	0.088	0.080	0.090	0.155
	t75e11xx	0.079	0.094	0.086	0.078	0.143
	t75k11xx	0.079	0.096	0.087	0.082	0.142
	t75n11xx	0.075	0.090	0.082	0.074	0.140

Table: ARPD results of 20 repetitions of the EDAs for the LOP instances.

Some Experiments

The Linear Ordering Problem

Size	Instance	PLEDA	MaEDA	GMEDA	NHBSA	EHBSA
150	t65b11xx	**0.056**	0.072	0.065	0.064	0.107
	t65d11xx	**0.064**	0.081	0.074	0.078	0.123
	t65f11xx	**0.062**	0.081	0.074	0.077	0.121
	t65l11xx	**0.041**	0.058	0.052	**0.041**	0.082
	t65n11xx	**0.052**	0.070	0.065	0.060	0.113
	t70b11xx	**0.058**	0.074	0.068	0.065	0.109
	t70d11xx	**0.067**	0.084	0.076	0.076	0.117
	t70f11xx	**0.069**	0.088	0.078	0.082	0.123
	t70l11xx	**0.052**	0.074	0.069	0.054	0.102
	t70n11xx	**0.056**	0.073	0.066	0.061	0.112
	t75d11xx	**0.064**	0.078	0.073	0.071	0.115
	t75e11xx	**0.048**	0.069	0.062	0.300	0.300
	t75k11xx	**0.055**	0.074	0.064	0.292	0.291
	t75n11xx	**0.060**	0.078	0.071	0.302	0.299
250	t65b11xx	**0.074**	0.086	0.080	0.088	0.146
	t65d11xx	**0.083**	0.099	0.090	0.105	0.161
	t65f11xx	**0.078**	0.094	0.086	0.097	0.159
	t65l11xx	**0.061**	0.075	0.071	0.068	0.132
	t65n11xx	**0.075**	0.092	0.084	0.094	0.162
	t70b11xx	**0.074**	0.088	0.081	0.092	0.157
	t70d11xx	**0.080**	0.095	0.086	0.097	0.157
	t70f11xx	**0.087**	0.101	0.094	0.109	0.162
	t70l11xx	**0.066**	0.082	0.076	0.078	0.150
	t70n11xx	**0.074**	0.091	0.081	0.094	0.162
	t75d11xx	**0.075**	0.088	0.080	0.090	0.155
	t75e11xx	0.079	0.094	0.086	**0.078**	0.143
	t75k11xx	**0.079**	0.096	0.087	0.082	0.142
	t75n11xx	0.075	0.090	0.082	**0.074**	0.140

Table: ARPD results of 20 repetitions of the EDAs for the LOP instances.

Some Experiments

The Linear Ordering Problem

Size	Instance	PLEDA	MaEDA	GMEDA	NHBSA	EHBSA
150	t65b11xx	**0.056**	0.072	0.065	**0.064**	0.107
	t65d11xx	**0.064**	0.081	**0.074**	0.078	0.123
	t65f11xx	**0.062**	0.081	**0.074**	0.077	0.121
	t65l11xx	**0.041**	0.058	**0.052**	**0.041**	0.082
	t65n11xx	**0.052**	0.070	0.065	**0.060**	0.113
	t70b11xx	**0.058**	0.074	0.068	**0.065**	0.109
	t70d11xx	**0.067**	0.084	**0.076**	**0.076**	0.117
	t70f11xx	**0.069**	0.088	**0.078**	0.082	0.123
	t70l11xx	**0.052**	0.074	0.069	**0.054**	0.102
	t70n11xx	**0.056**	0.073	0.066	**0.061**	0.112
	t75d11xx	**0.064**	0.078	0.073	**0.071**	0.115
	t75e11xx	**0.048**	0.069	**0.062**	0.300	0.300
	t75k11xx	**0.055**	0.074	**0.064**	0.292	0.291
	t75n11xx	**0.060**	0.078	**0.071**	0.302	0.299
250	t65b11xx	**0.074**	0.086	**0.080**	0.088	0.146
	t65d11xx	**0.083**	0.099	**0.090**	0.105	0.161
	t65f11xx	**0.078**	0.094	**0.086**	0.097	0.159
	t65l11xx	**0.061**	0.075	0.071	**0.068**	0.132
	t65n11xx	**0.075**	0.092	**0.084**	0.094	0.162
	t70b11xx	**0.074**	0.088	**0.081**	0.092	0.157
	t70d11xx	**0.080**	0.095	**0.086**	0.097	0.157
	t70f11xx	**0.087**	0.101	**0.094**	0.109	0.162
	t70l11xx	**0.066**	0.082	**0.076**	0.078	0.150
	t70n11xx	**0.074**	0.091	**0.081**	0.094	0.162
	t75d11xx	**0.075**	0.088	**0.080**	0.090	0.155
	t75e11xx	**0.079**	0.094	0.086	**0.078**	0.143
	t75k11xx	**0.079**	0.096	0.087	**0.082**	0.142
	t75n11xx	**0.075**	0.090	0.082	**0.074**	0.140

Table: ARPD results of 20 repetitions of the EDAs for the LOP instances.

Config.	Instance	PLEDA	MaEDA	GMEDA	NHBSA	EHBSA
100×20	1	0.074	0.045	0.026	0.060	0.062
	2	0.076	0.044	0.026	0.059	0.057
	3	0.075	0.047	0.026	0.057	0.057
	4	0.075	0.048	0.027	0.061	0.060
	5	0.077	0.048	0.026	0.061	0.060
	6	0.080	0.053	0.029	0.061	0.061
	7	0.076	0.044	0.027	0.061	0.062
	8	0.074	0.047	0.025	0.058	0.057
	9	0.071	0.042	0.025	0.058	0.055
	10	0.078	0.052	0.030	0.061	0.059
200×20	1	0.083	0.070	0.027	0.071	0.076
	2	0.085	0.073	0.023	0.074	0.082
	3	0.085	0.081	0.054	0.070	0.070
	4	0.087	0.077	0.024	0.076	0.082
	5	0.083	0.066	0.022	0.075	0.085
	6	0.079	0.050	0.020	0.073	0.084
	7	0.080	0.051	0.019	0.074	0.085
	8	0.085	0.073	0.029	0.074	0.081
	9	0.085	0.062	0.022	0.078	0.086
	10	0.082	0.074	0.022	0.073	0.077
250×20	1	0.082	0.076	0.041	0.075	0.085
	2	0.085	0.081	0.067	0.075	0.082
	3	0.077	0.065	0.024	0.072	0.082
	4	0.084	0.076	0.037	0.076	0.089
	5	0.090	0.085	0.075	0.079	0.086
	6	0.080	0.074	0.037	0.074	0.086
	7	0.078	0.074	0.041	0.070	0.077
	8	0.088	0.080	0.032	0.078	0.085
	9	0.080	0.074	0.023	0.074	0.083
	10	0.085	0.079	0.029	0.076	0.084

Table: ARPD results of 20 repetitions of the EDAs for the FSP instances.

Config.	Instance	PLEDA	MaEDA	GMEDA	NHBSA	EHBSA
100×20	1	0.074	0.045	**0.026**	0.060	0.062
	2	0.076	0.044	**0.026**	0.059	0.057
	3	0.075	0.047	**0.026**	0.057	0.057
	4	0.075	0.048	**0.027**	0.061	0.060
	5	0.077	0.048	**0.026**	0.061	0.060
	6	0.080	0.053	**0.029**	0.061	0.061
	7	0.076	0.044	**0.027**	0.061	0.062
	8	0.074	0.047	**0.025**	0.058	0.057
	9	0.071	0.042	**0.025**	0.058	0.055
	10	0.078	0.052	**0.030**	0.061	0.059
200×20	1	0.083	0.070	**0.027**	0.071	0.076
	2	0.085	0.073	**0.023**	0.074	0.082
	3	0.085	0.081	**0.054**	0.070	0.070
	4	0.087	0.077	**0.024**	0.076	0.082
	5	0.083	0.066	**0.022**	0.075	0.085
	6	0.079	0.050	**0.020**	0.073	0.084
	7	0.080	0.051	**0.019**	0.074	0.085
	8	0.085	0.073	**0.029**	0.074	0.081
	9	0.085	0.062	**0.022**	0.078	0.086
	10	0.082	0.074	**0.022**	0.073	0.077
250×20	1	0.082	0.076	**0.041**	0.075	0.085
	2	0.085	0.081	**0.067**	0.075	0.082
	3	0.077	0.065	**0.024**	0.072	0.082
	4	0.084	0.076	**0.037**	0.076	0.089
	5	0.090	0.085	**0.075**	0.079	0.086
	6	0.080	0.074	**0.037**	0.074	0.086
	7	0.078	0.074	**0.041**	0.070	0.077
	8	0.088	0.080	**0.032**	0.078	0.085
	9	0.080	0.074	**0.023**	0.074	0.083
	10	0.085	0.079	**0.029**	0.076	0.084

Table: ARPD results of 20 repetitions of the EDAs for the FSP instances.

Config.	Instance	PLEDA	MaEDA	GMEDA	NHBSA	EHBSA
100×20	1	0.074	**0.045**	**0.026**	0.060	0.062
	2	0.076	**0.044**	**0.026**	0.059	0.057
	3	0.075	**0.047**	**0.026**	0.057	0.057
	4	0.075	**0.048**	**0.027**	0.061	0.060
	5	0.077	**0.048**	**0.026**	0.061	0.060
	6	0.080	**0.053**	**0.029**	0.061	0.061
	7	0.076	**0.044**	**0.027**	0.061	0.062
	8	0.074	**0.047**	**0.025**	0.058	0.057
	9	0.071	**0.042**	**0.025**	0.058	0.055
	10	0.078	**0.052**	**0.030**	0.061	0.059
200×20	1	0.083	**0.070**	**0.027**	0.071	0.076
	2	0.085	**0.073**	**0.023**	0.074	0.082
	3	0.085	0.081	**0.054**	**0.070**	0.070
	4	0.087	0.077	**0.024**	**0.076**	0.082
	5	0.083	**0.066**	**0.022**	0.075	0.085
	6	0.079	**0.050**	**0.020**	0.073	0.084
	7	0.080	**0.051**	**0.019**	0.074	0.085
	8	0.085	**0.073**	**0.029**	0.074	0.081
	9	0.085	**0.062**	**0.022**	0.078	0.086
	10	0.082	0.074	**0.022**	**0.073**	0.077
250×20	1	0.082	0.076	**0.041**	**0.075**	0.085
	2	0.085	0.081	**0.067**	**0.075**	0.082
	3	0.077	**0.065**	**0.024**	0.072	0.082
	4	0.084	**0.076**	**0.037**	**0.076**	0.089
	5	0.090	0.085	**0.075**	**0.079**	0.086
	6	0.080	**0.074**	**0.037**	**0.074**	0.086
	7	0.078	0.074	**0.041**	**0.070**	0.077
	8	0.088	0.080	**0.032**	**0.078**	0.085
	9	0.080	**0.074**	**0.023**	**0.074**	0.083
	10	0.085	0.079	**0.029**	**0.076**	0.084

Table: ARPD results of 20 repetitions of the EDAs for the FSP instances.

The PLEDA performs the best for the LOP,
but is not competitive when solving the FSP.

The PLEDA performs the best for the LOP,
but is not competitive when solving the FSP.

The GMEDA performs the best for the FSP,
and is still competitive for LOP.

Outline

Discussion

Understanding the results...

How do the properties of the models determine the behavior of the EDAs when solving each problem?

In the LOP...

...the contribution of an index $\pi(i)$ to the objective function depends on the previous and posterior indices to it, but not on their relative ordering.

In the LOP...

...the contribution of an index $\pi(i)$ to the objective function depends on the previous and posterior indices to it, but not on their relative ordering.

In the Plackett-Luce

...the choice probability at the j^{th} stage depends only on the set of items remaining at that stage, and not on the relative ordering of the $j-1$ items previously selected.

In the LOP...

...the contribution of an index $\pi(i)$ to the objective function depends on the previous and posterior indices to it, but not on their relative ordering.

The Generalized Mallows

The choice probability made at stage j is defined as

$$P(V_j(\pi\pi_0^{-1}) = r_j) = \frac{e^{-\theta_j r_j}}{\psi_j(\theta_j)}$$

where r_j denotes the number of mistakes made at that stage with respect to a central ranking π_0.

In the FSP...

...the contribution of an index $\pi(i)$ to the objective function depends on the relative ordering of the previous $i-1$ indices.

In the FSP...

...the contribution of an index $\pi(i)$ to the objective function depends on the relative ordering of the previous $i-1$ indices.

The Plackett-Luce

The bad performance of the PLEDA is induced by the inappropriate characteristics of the model.

In the FSP...

...the contribution of an index $\pi(i)$ to the objective function depends on the relative ordering of the previous $i - 1$ indices.

The Generalized Mallows

The probability of a solution is calculated according to its resemblance with a consensus ranking.

In the FSP...

...the contribution of an index $\pi(i)$ to the objective function depends on the relative ordering of the previous $i-1$ indices.

The Generalized Mallows

The probability of a solution is calculated according to its resemblance with a consensus ranking.

Kendall's-τ distance

High expressiveness!!

The Plackett-Luce Ranking Model on Permutation-based Optimization Problems

Josu Ceberio, Alexander Mendiburu, Jose A. Lozano

Intelligent Systems Group
Department of Computer Science and Artificial Intelligence
University of the Basque Country UPV/EHU

2013 IEEE Congress on Evolutionary Computation (CEC 2013)
Cancun, Mexico, 22-23 April 2013

Parameters Settings

Population size	$10n$
Selection size	n
Offspring size	$10n - 1$
Max. Evaluations	$1000n$

n denotes the size of the problem.