Dealing with constraints in estimation of distribution algorithms: a different approach

Josu Ceberio, Alexander Mendiburu, Jose A. Lozano

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

Intelligent Systems Group

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea

FACULTY OF COMPUTER SCIENCE
UNIVERSITY OF THE BASQUE COUNTRY

Estimation of distribution algorithms

Introduction

METAHEURISTIC ALGORITHMS

SIMILAR TO GENETIC ALGORITHMS

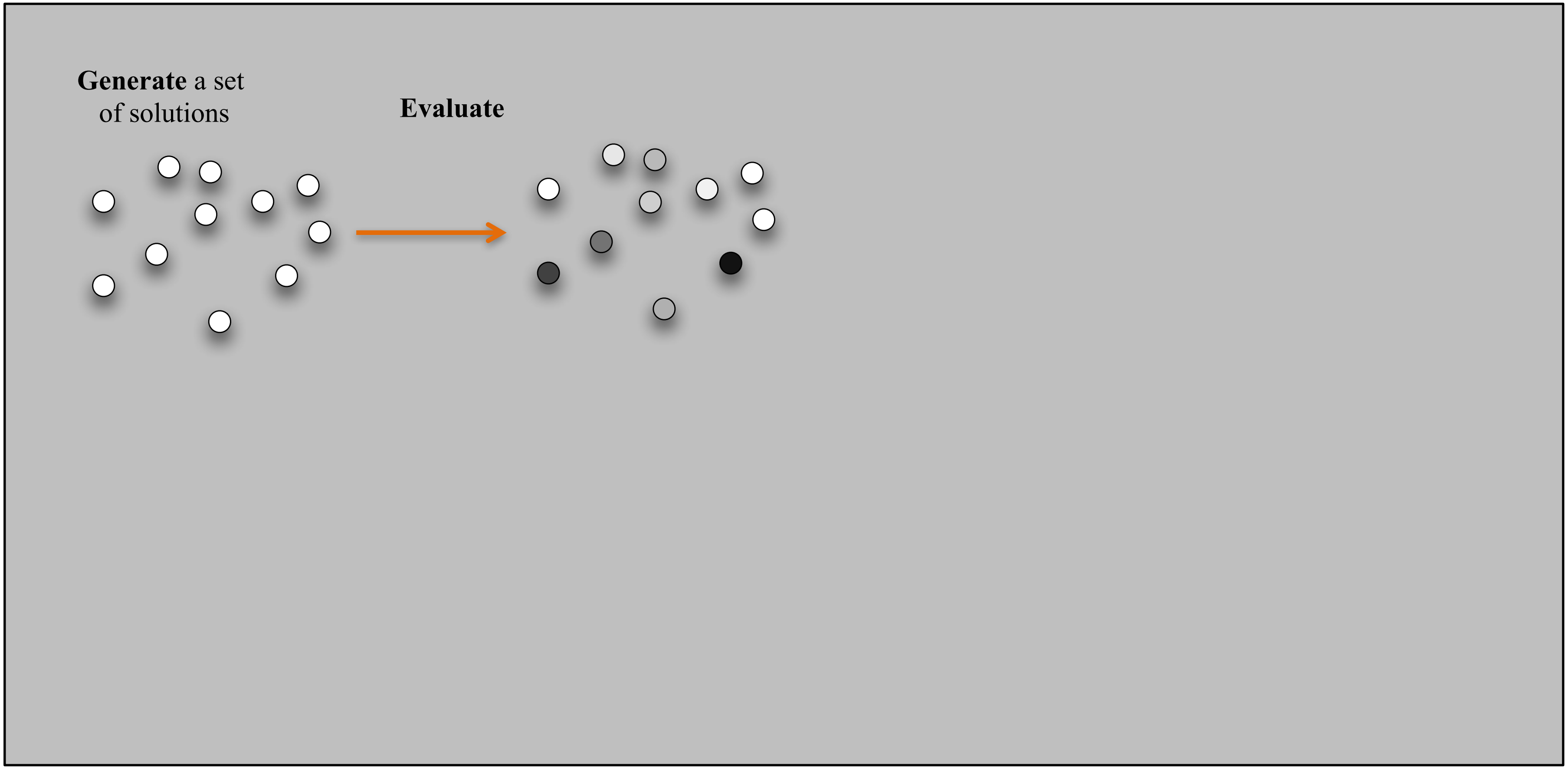
LEARN AND SAMPLE A PROBABILITY DISTRIBUTION

Estimation of distribution algorithms

Scheme

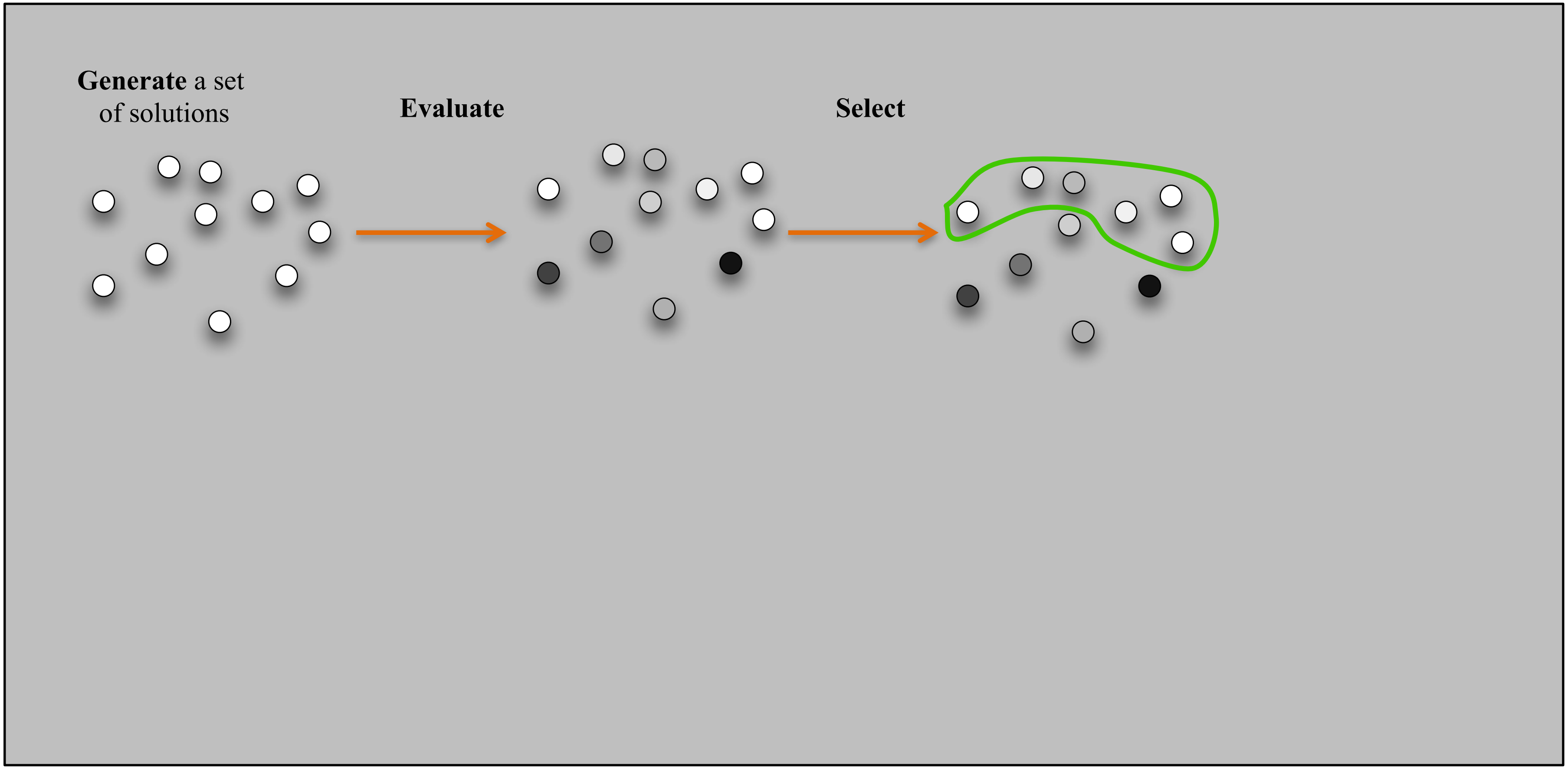
Estimation of distribution algorithms

Scheme



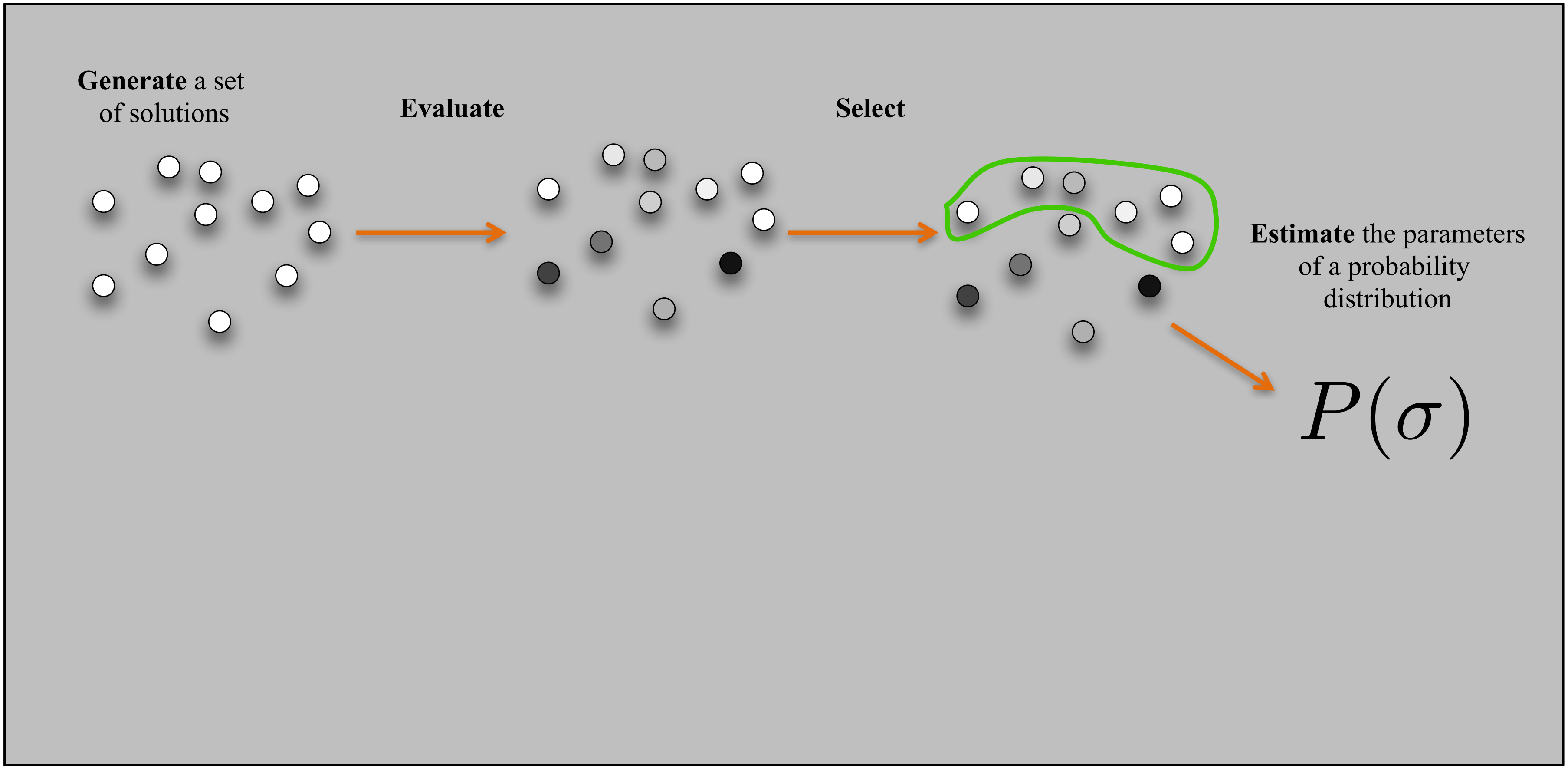
Estimation of distribution algorithms

Scheme



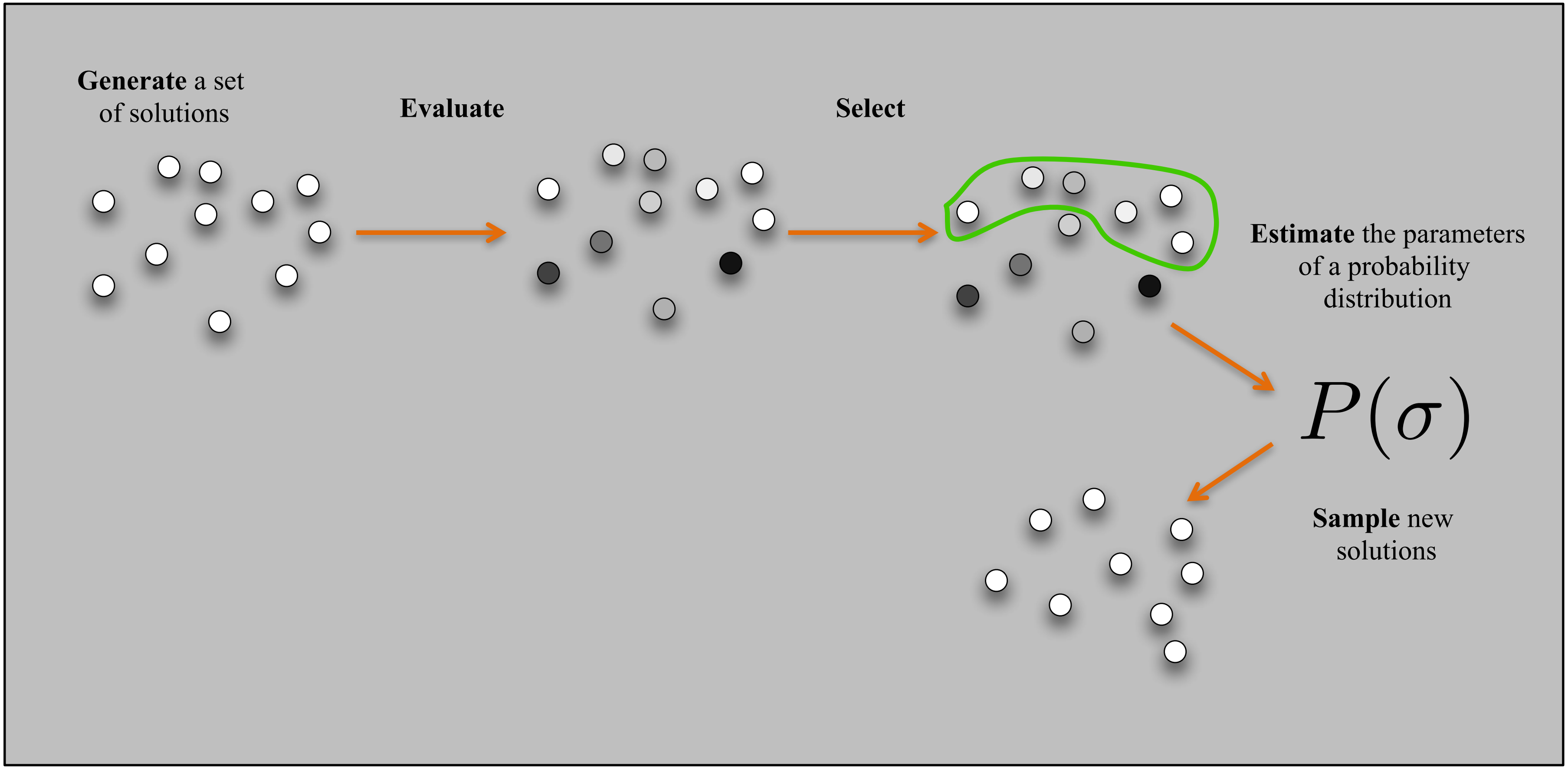
Estimation of distribution algorithms

Scheme



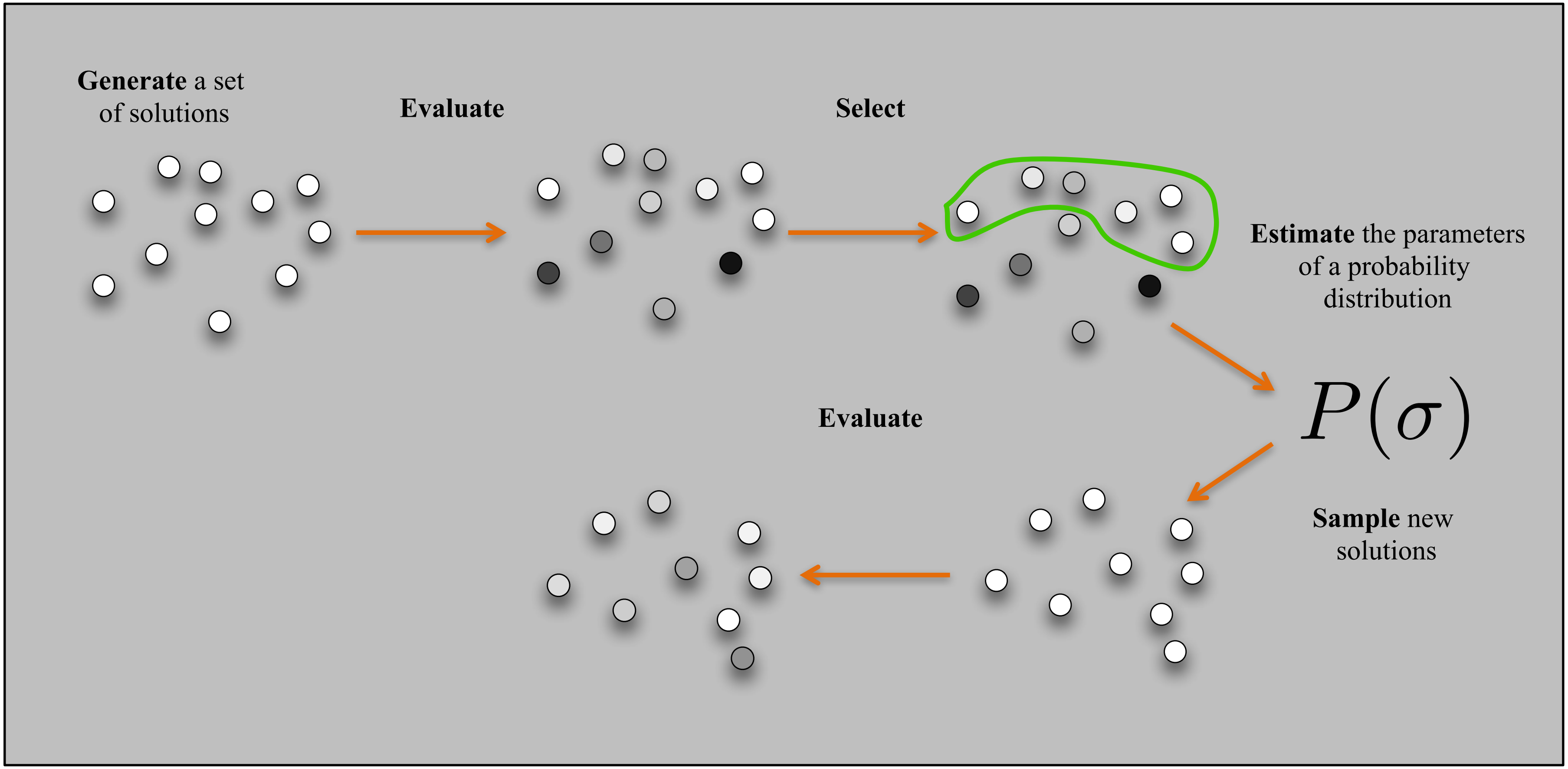
Estimation of distribution algorithms

Scheme



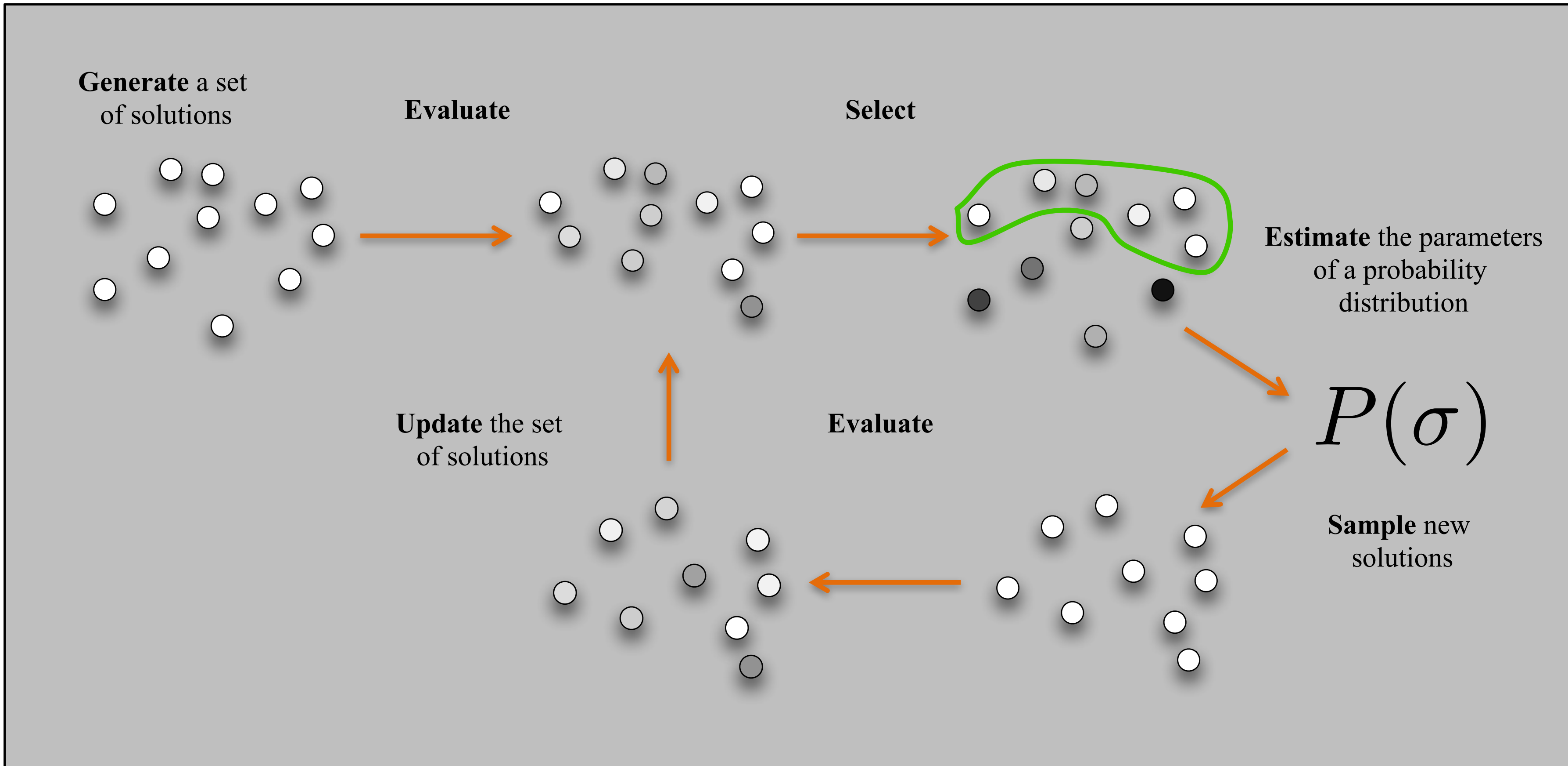
Estimation of distribution algorithms

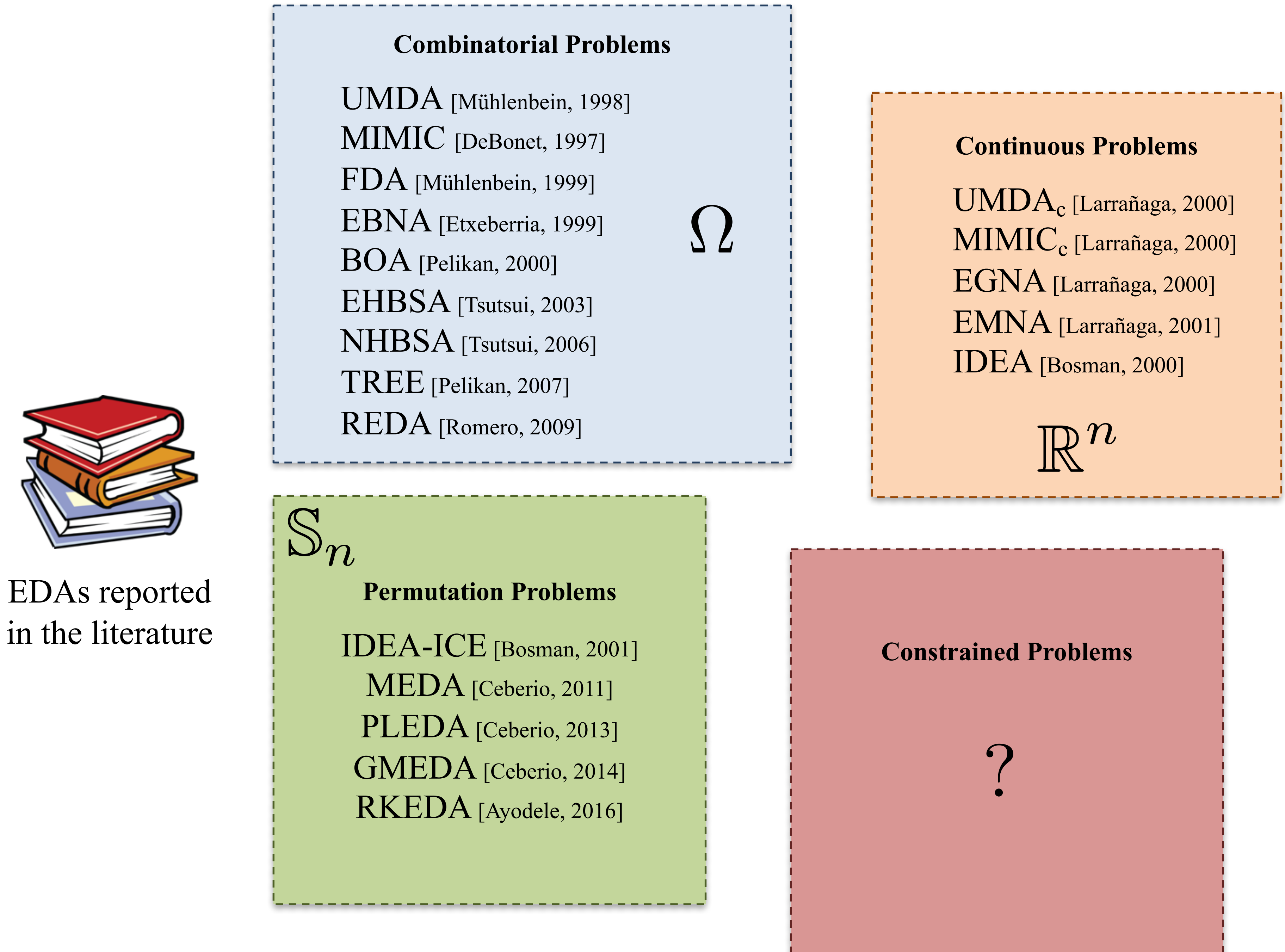
Scheme



Estimation of distribution algorithms

Scheme





EDAs reported in the literature
Combinatorial Problems
UMDA [Mühlenbein, 1998]
MIMIC [DeBonet, 1997]
FDA [Mühlenbein, 1999]
EBNA [Etxeberria, 1999]
BOA [Pelikan, 2000]
EHBSA [Tsutsui, 2003]
NHBSA [Tsutsui, 2006]
TREE [Pelikan, 2007]
REDA [Romero, 2009]
Ω
Continuous Problems
$UMDA_c$ [Larrañaga, 2000]
$MIMIC_c$ [Larrañaga, 2000]
EGNA [Larrañaga, 2000]
EMNA [Larrañaga, 2001]
IDEA [Bosman, 2000]
$\mathbb{R}^n$
$\mathbb{S}_n$
Permutation Problems
IDEA-ICE [Bosman, 2001]
MEDA [Ceberio, 2011]
PLEDA [Ceberio, 2013]
GMEDA [Ceberio, 2014]
RKEDA [Ayodele, 2016]
Constrained Problems
?

Constrained Optimization Problems

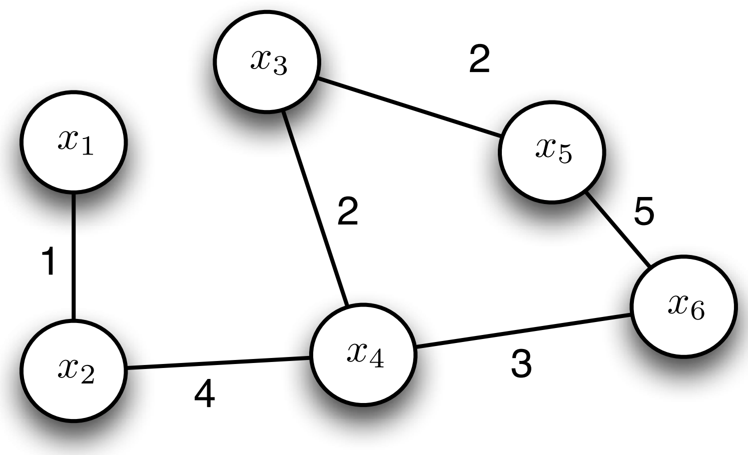
Definition

$$
\begin{aligned}
&\text{minimising} && f(\mathbf{x}), && \mathbf{x} = (x_1, \ldots, x_n) \\
&\text{subject to,} && g_i(\mathbf{x}) \leq 0, && i = 1, \ldots, r \\
& && h_j(\mathbf{x}) = 0, && j = r+1, \ldots, m
\end{aligned}
$$

Some examples

- → Knapsack Problem
- → Graph Colouring Problem
- → Maximum Satisfiability Problem
- → Capacitated Arc Routing Problem
- → …

Graph Partitioning Problem



Find a k-partition of vertices minimising the weight of edges between sets: the cut size

We considered the balanced *2*-partition GPP.

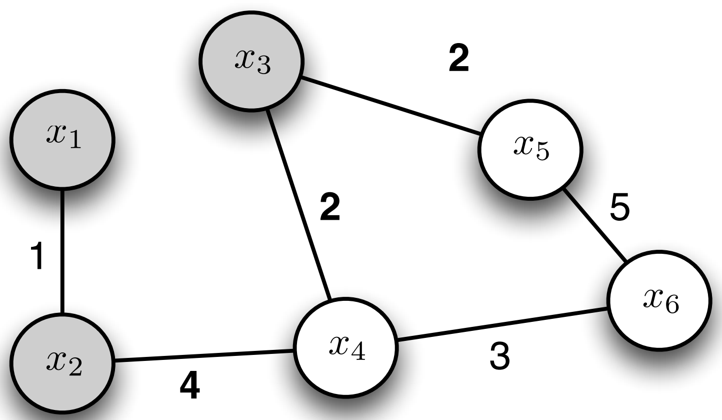
Solutions are codified as…

$$\mathbf{x} \in \{0, 1\}^n$$

$$f(\mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i(1 - x_j) w_{ij}$$

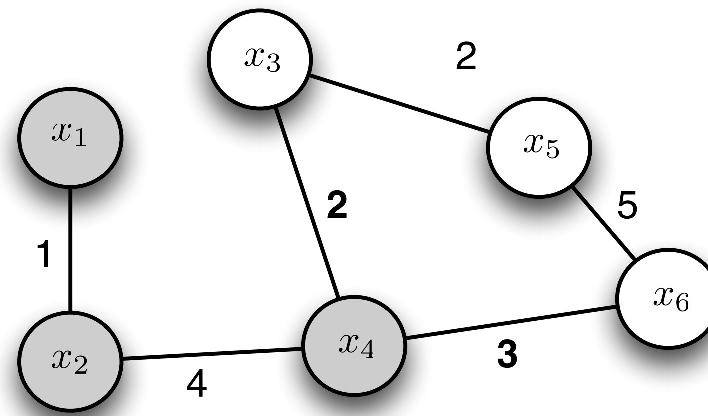
Objective Function

Graph Partitioning Problem



	x_1	x_2	x_3	x_4	x_5	x_6
$\mathbf{x}^1$	0	0	0	1	1	1

$f(\mathbf{x}^1) = 8$



	x_1	x_2	x_3	x_4	x_5	x_6
$\mathbf{x}^2$	0	0	1	0	1	1

$f(\mathbf{x}^2) = 5$

The constraint: equal number of zeros as ones

$$\sum_{i=1}^{n} x_i = n/2$$

Constrained Optimization Problems

Why are they challenging?

The search space of solutions induced by the codification is…

000000	001000	010000	011000	100000	101000	110000	111000
000001	001001	010001	011001	100001	101001	110001	111001
000010	001010	010010	011010	100010	101010	110010	111010
000011	001011	010011	011011	100011	101011	110011	111011
000100	001100	010100	011100	100100	101100	110100	111100
000101	001101	010101	011101	100101	101101	110101	111101
000110	001110	010110	011110	100110	101110	110110	111110
000111	001111	010111	011111	100111	101111	110111	111111

Constrained Optimization Problems

Why are they challenging?

The search space of solutions induced by the codification is…

000000	001000	010000	011000	100000	101000	110000	111000
000001	001001	010001	011001	100001	101001	110001	111001
000010	001010	010010	011010	100010	101010	110010	111010
000011	001011	010011	011011	100011	101011	110011	111011
000100	001100	010100	011100	100100	101100	110100	111100
000101	001101	010101	011101	100101	101101	110101	111101
000110	001110	010110	011110	100110	101110	110110	111110
000111	001111	010111	011111	100111	101111	110111	111111

The majority of the solutions are not feasible !

What happens if we run a UMDA?

Univariate Marginal Distribution Algorithm (UMDA)

$$P(x) = \prod_{i=1}^{n} P(x_i)$$

First order marginals
No dependencies are considered

What happens if we run a UMDA?

	x_1	x_2	x_3	x_4	x_5	x_6
0	0.6	0.5	0.25	0.66	0.9	0.5
1	0.4	0.5	0.75	0.33	0.1	0.5

↓

0 0 1 1 0 0

1 1 0 0 0 0

0 0 1 1 0 1

0 1 0 1 0 1

0 1 0 1 1 1

What happens if we run a UMDA?

	x_1	x_2	x_3	x_4	x_5	x_6
0	0.6	0.5	0.25	0.66	0.9	0.5
1	0.4	0.5	0.75	0.33	0.1	0.5

↓

0 0 1 1 0 0

1 1 0 0 0 0

0 0 1 1 0 1

0 1 0 1 0 1

0 1 0 1 1 1

Unfeasible solutions are generated…

Different approaches

Literature review

1. Repair solutions

- Modify solutions to hold the constraints

2. Penalty functions

- Punish solutions to be discarded when selection

3. Guarantee feasibility when sampling

- In EDAs, adapt the sampling to create feasible solutions

The role of the probability model is somehow denaturalized

The Idea

Conduct the optimisation entirely on the set of feasible solutions…

4. Use probability distributions defined on this set

Motivation

Permutation-based Problems

Combinatorial Optimization Problems

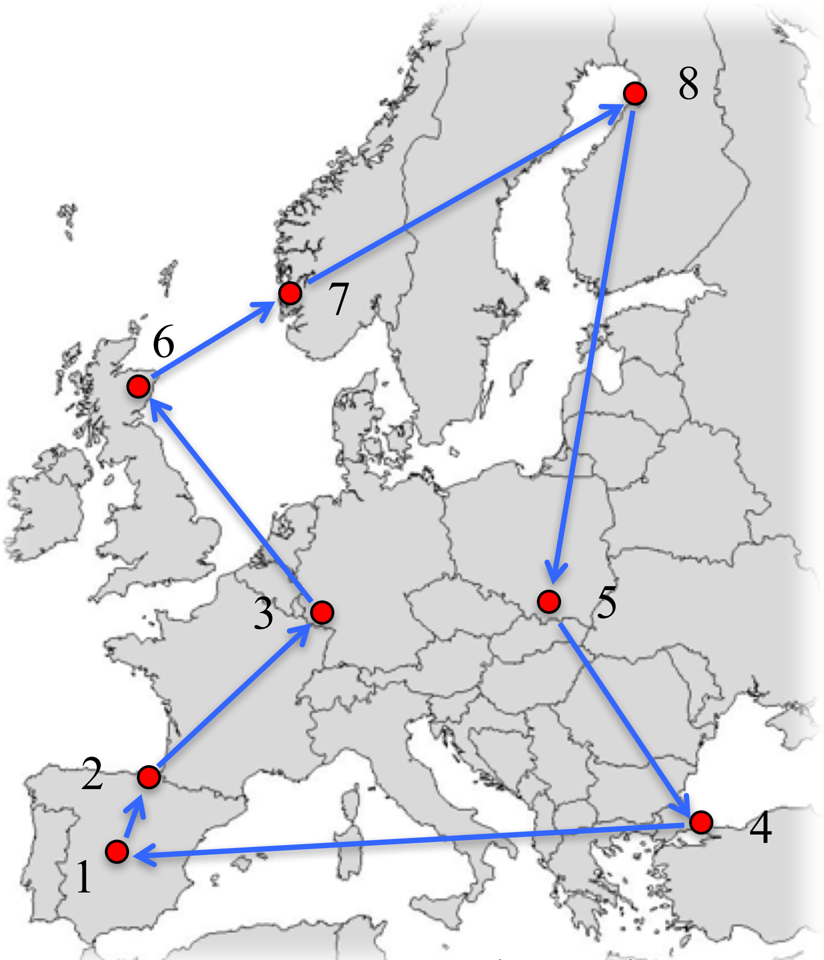
Whose solutions are represented as permutations

The search space consist of $n!$ solutions

$$8! = 40320$$

$$20! = 2.43 \times 10^{18}$$

Travelling Salesman Problem (TSP)



$$\sigma = 12367854$$

Motivation

Permutation-based Problems

The space of permutations can be seen as a constrained space of the integers space

$n = 3$

111	211	311
112	212	312
113	213	313
121	221	321
122	222	322
123	223	323
131	231	331
132	232	332
133	233	333

Motivation

Permutation-based Problems

The space of permutations can be seen as a constrained space of the integers space

$n = 3$

111	211	311
112	212	312
113	213	313
121	221	321
122	222	322
123	223	323
131	231	331
132	232	332
133	233	333

Motivation

Probability Models on Rankings

Bibliography

- M. A. Fligner and J. S. Verducci (1998), Multistage Ranking Models, *Journal of the American Statistical Association,* vol. 83, no. 403, pp. 892-901.
- D. E. Critchlow, M. A. Fligner, and J. S. Verducci (1991), Probability Models on Rankings, *Journal of Mathematical Psychology*, vol. 35, no. 3, pp. 294-318.
- P. Diaconis (1988), Group Representations in Probability and Statistics, *Institute of Mathematical Statistics.*
- M. A. Fligner and J. S. Verducci (1986), Distance based Ranking Models, *Journal of Royal Statistical Society, Series B*, vol. 48, no. 3, pp. 359-369.
- R. L. Plackett (1975), The Analysis of Permutations, *Applied Statistics*, vol. 24, no. 10, pp. 193-202.
- D. R. Luce (1959), Individual Choice Behaviour, *Wiley.*
- R. A. Bradley AND M. E. Terry (1952), Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons, *Biometrika*, vol. 39, no. 3, pp. 324-345.
- L. L. Thurstone (1927), A law of comparative judgment, *Psychological Review*, vol 34, no. 4, pp. 273-286.

Motivation

Probability Models on Rankings

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

Mallows

Distance-based

$$P(\sigma) = \frac{1}{\psi(\theta)} e^{-\sum_{j=1}^{n-1} \theta_j S_j(\sigma, \sigma_0)}$$

Generalized Mallows

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

Plackett-Luce

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

Bradley-Terry

Order statistics

Apply same idea in constrained COPs…

Do probability models for constrained spaces exist?

No,

At least, we do not know them…

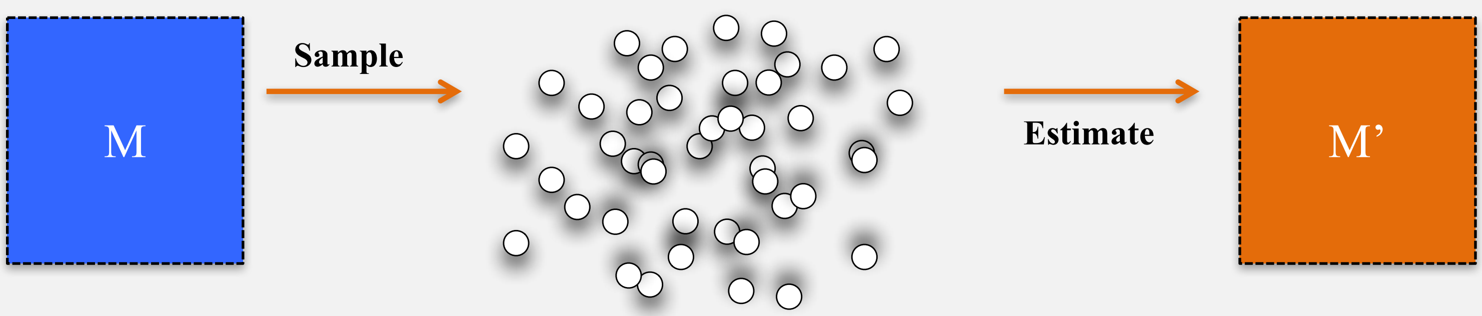
Designing probability models

Requirements

1 $\forall x \in \Omega, \quad 0 \leq P(x) \leq 1$

2 $\sum_{i=1}^{|\Omega|} P(x_i) = 1$

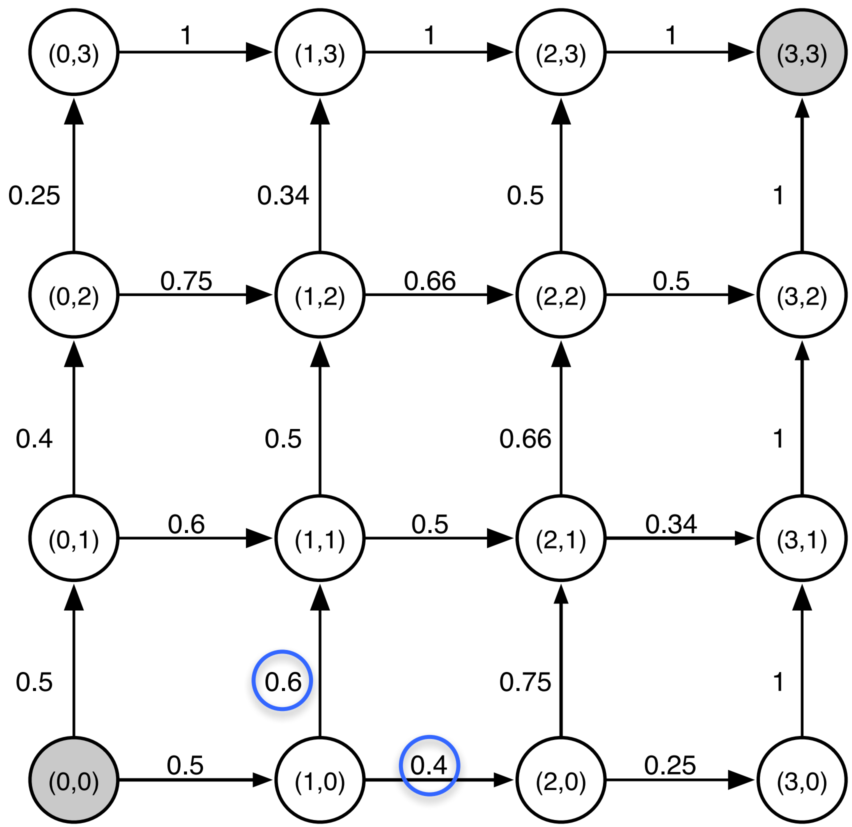
3 Develop efficient **learning** and **sampling** methods



$$M = M'$$

A Square Lattice

The Probability Model



Square Lattice for n=6 GPP instance

Solutions are modelled as paths on a square lattice of $(n/2+1)^2$ vertices

Vertices: the number of ones and zeros at that stage

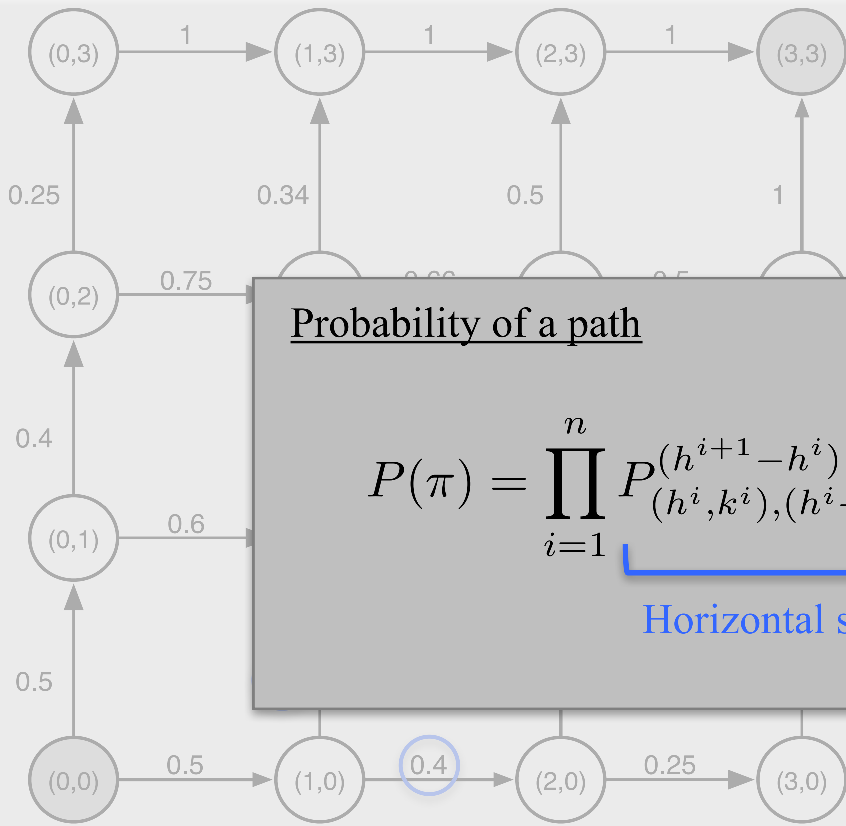
Edges: the probability of moving from one vertex to another.

$$\mathbf{x} = (1, 0, 0, 1, 0, 1)$$

$$\pi = ((0,1), (1,1), (2,1), (2,2), (3,2), (3,3))$$

A Square Lattice

The Probability Model



Square Lattice for n=4 GPP instance

Solutions are modelled as paths on a square lattice of $(n/2+1)^2$ vertices

…nber of ones …hat stage

…bability of …e vertex to …er

$$\mathbf{x} = (1, 0, 0, 1, 0, 1)$$

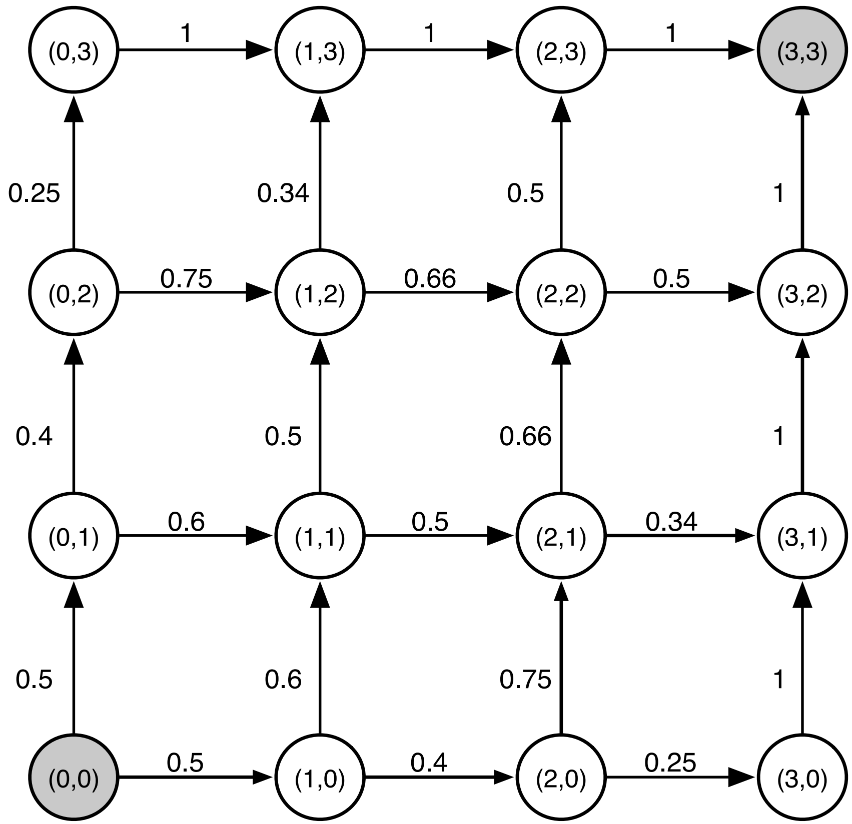
<u>Probability of a path</u>

$$P(\pi) = \prod_{i=1}^{n} \underbrace{P^{(h^{i+1}-h^{i})}_{(h^i,k^i),(h^i+1,k^i)}}_{\text{Horizontal step}} \underbrace{P^{(k^{i+1}-k^{i})}_{(h^i,k^i),(h^i,k^i+1)}}_{\text{Vertical step}}$$

$$\pi = ((0,1),(1,1),(2,1),(2,2),(3,2),(3,3))$$

A Square Lattice

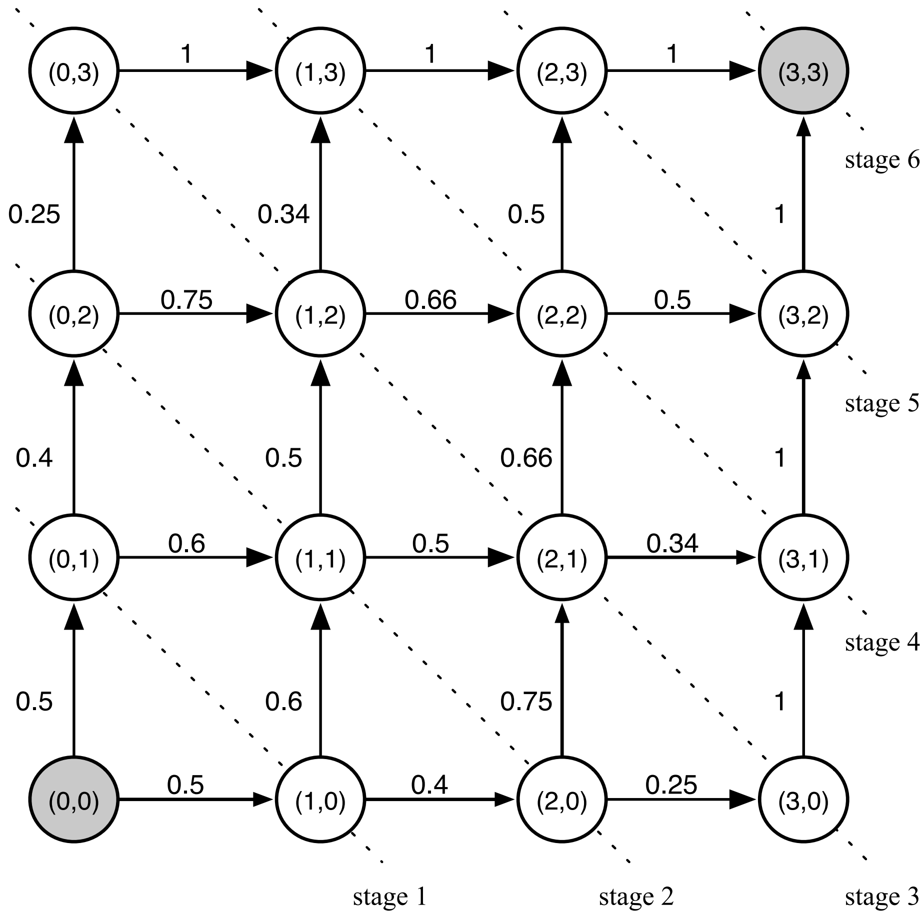
The Probability Model - Sampling



Square Lattice for n=6 GPP instance

A Square Lattice

The Probability Model - Sampling



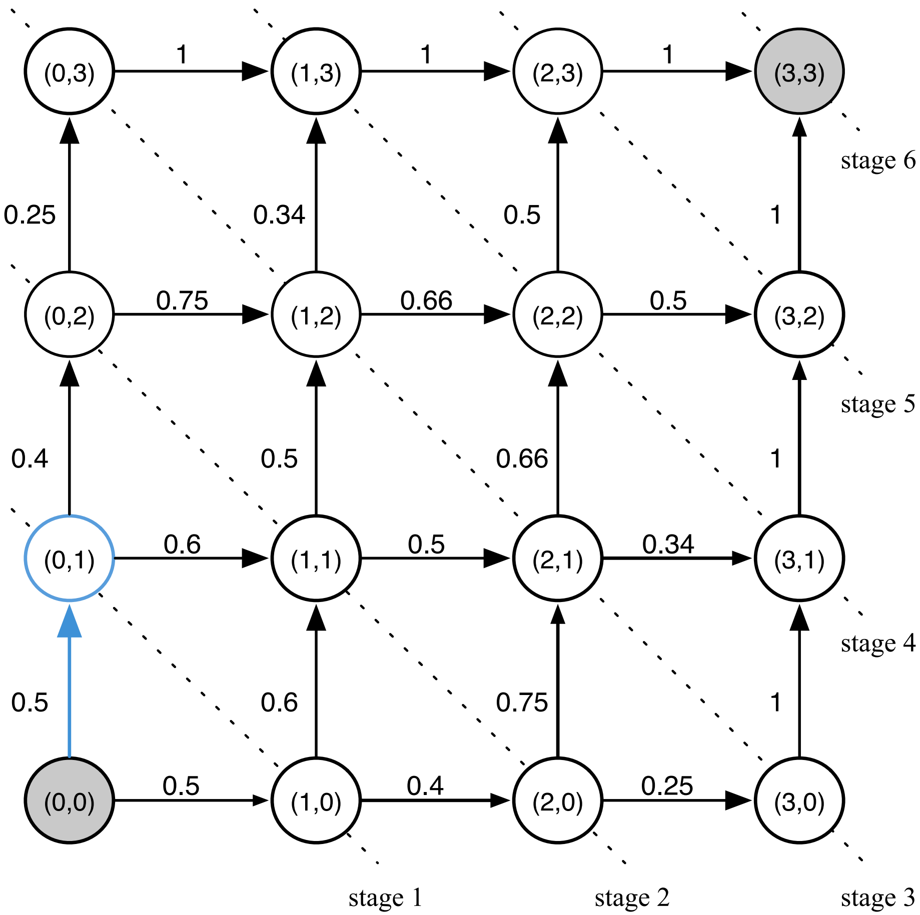
Square Lattice for n=6 GPP instance

Solutions are sampled at *n* stages

At each stage a decision has to be taken: up/right

A Square Lattice

The Probability Model - Sampling



Square Lattice for n=6 GPP instance

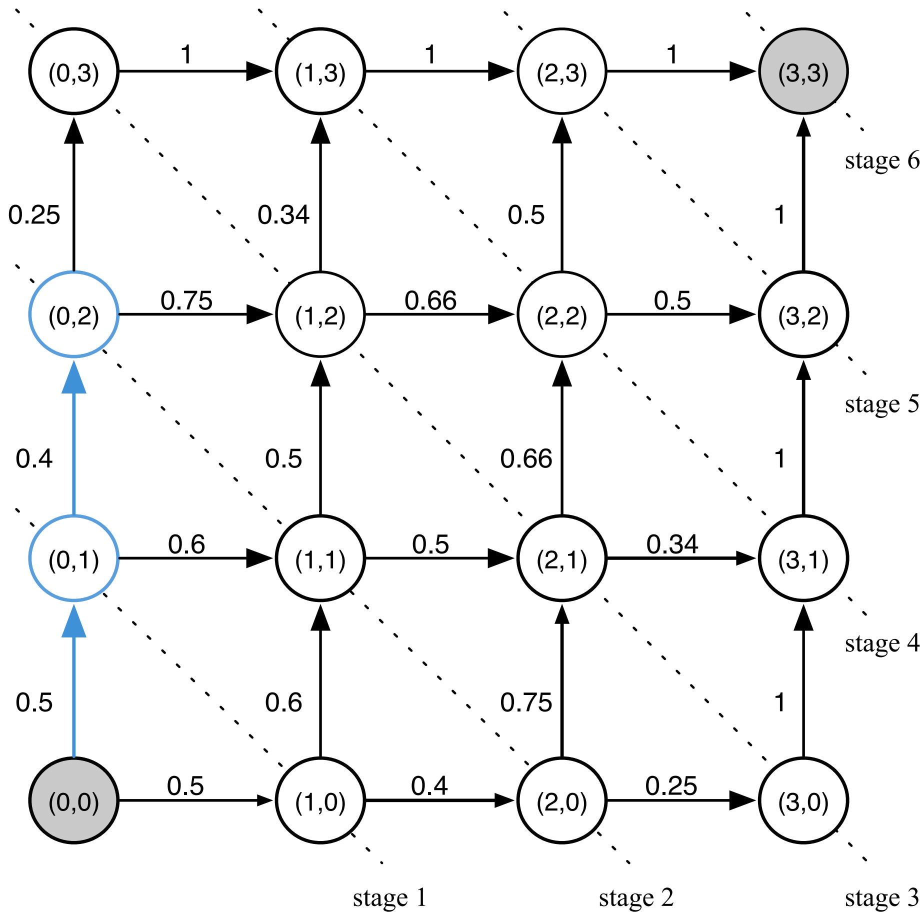
Solutions are sampled at *n* stages

At each stage a decision has to be taken: up/right

$$\mathbf{x} = (1$$

A Square Lattice

The Probability Model - Sampling



Square Lattice for n=6 GPP instance

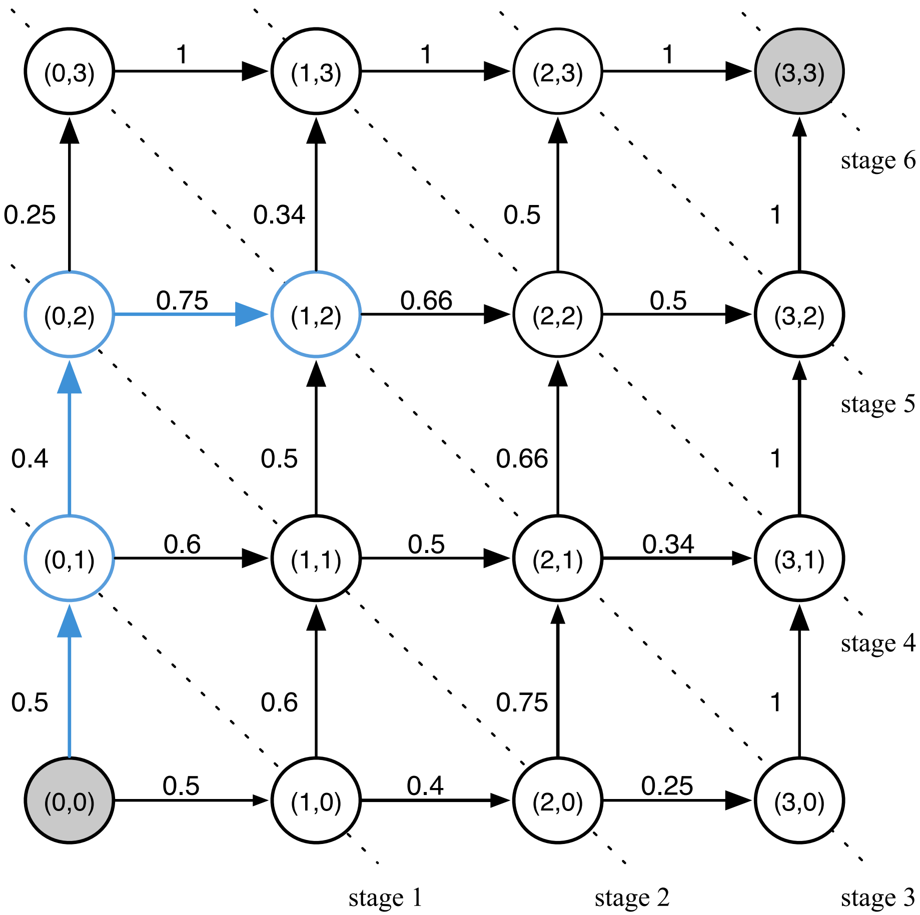
Solutions are sampled at *n* stages

At each stage a decision has to be taken: up/right

$$\mathbf{x} = (1, 1$$

A Square Lattice

The Probability Model - Sampling



Square Lattice for n=6 GPP instance

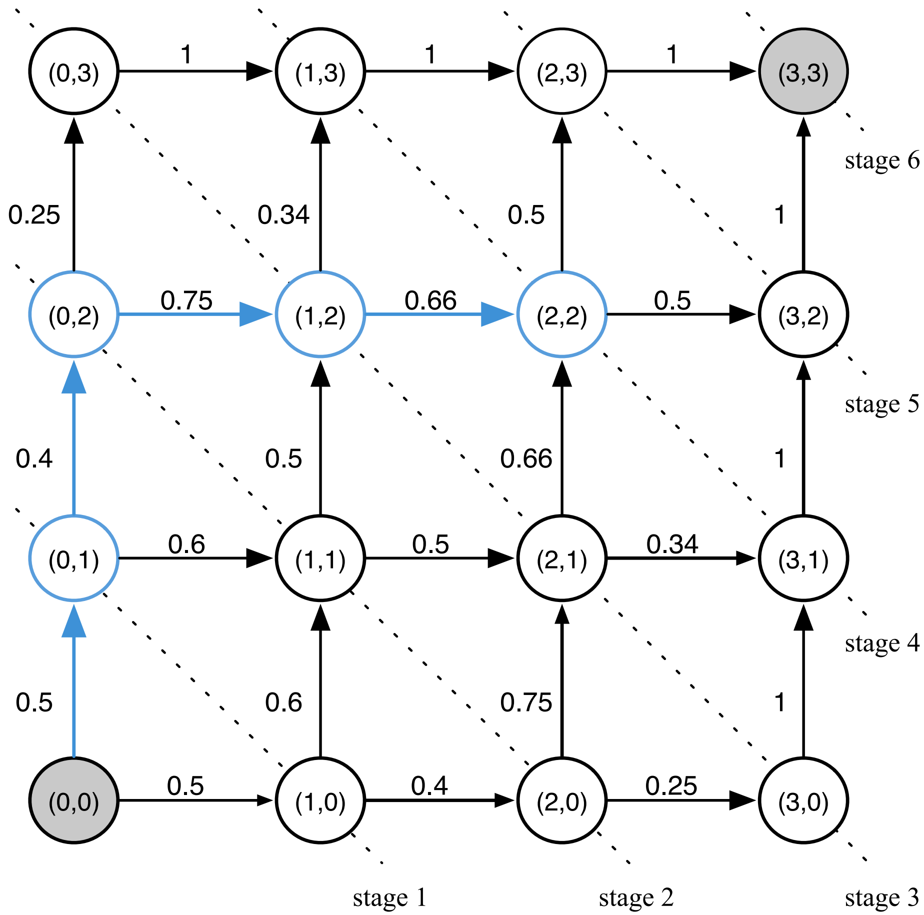
Solutions are sampled at n stages

At each stage a decision has to be taken: up/right

$$\mathbf{x} = (1, 1, 0$$

A Square Lattice

The Probability Model - Sampling



Square Lattice for n=6 GPP instance

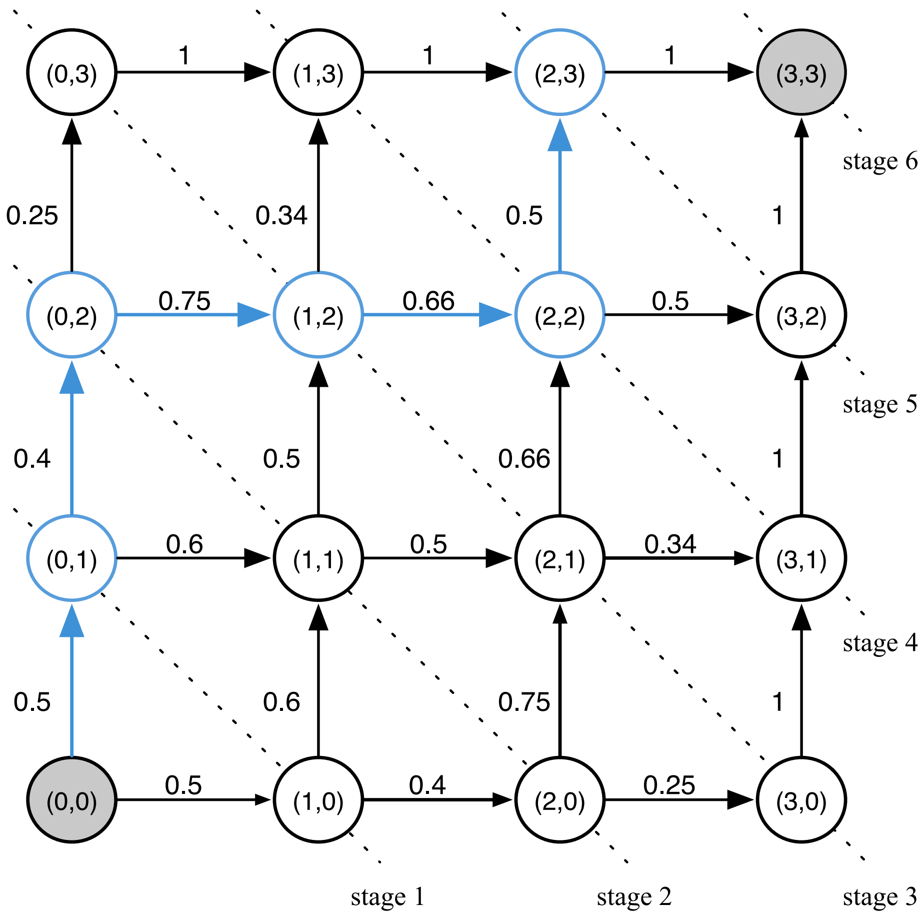
Solutions are sampled at n stages

At each stage a decision has to be taken: up/right

$$\mathbf{x} = (1, 1, 0, 0$$

A Square Lattice

The Probability Model - Sampling



Square Lattice for n=6 GPP instance

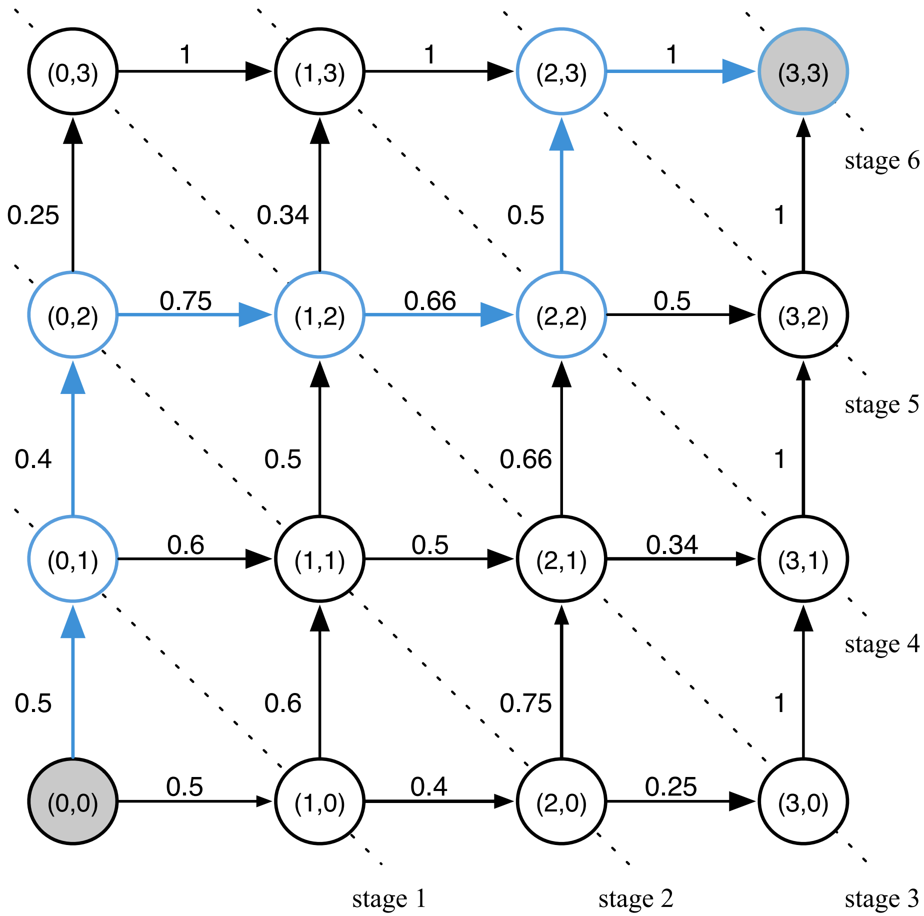
Solutions are sampled at *n* stages

At each stage a decision has to be taken: up/right

$$\mathbf{x} = (1, 1, 0, 0, 1$$

A Square Lattice

The Probability Model - Sampling



Square Lattice for n=6 GPP instance

Solutions are sampled at *n* stages

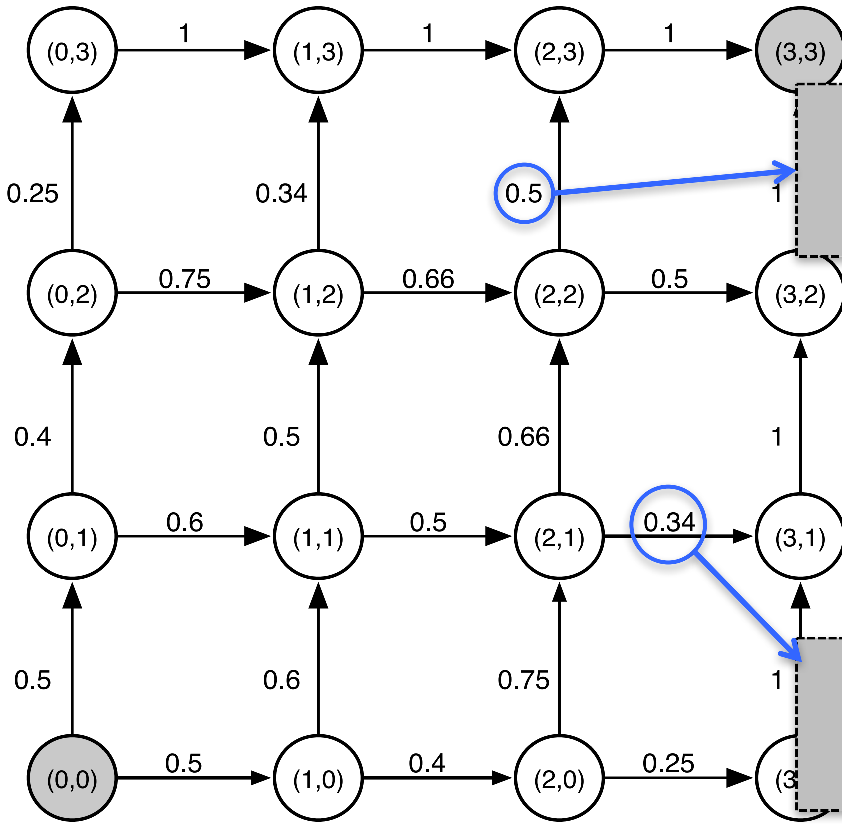
At each stage a decision has to be taken: up/right

$$\mathbf{x} = (1, 1, 0, 0, 1, 0)$$

As labels are interchangeable, we sample first a 0

A Square Lattice

The Probability Model – Estimating parameters



$$P_{(h,k),(h,k+1)} = \frac{\#\ \text{solutions moved from } (h,k) \text{ to } (h,k+1)}{\#\ \text{solutions reached } (h,k)}$$

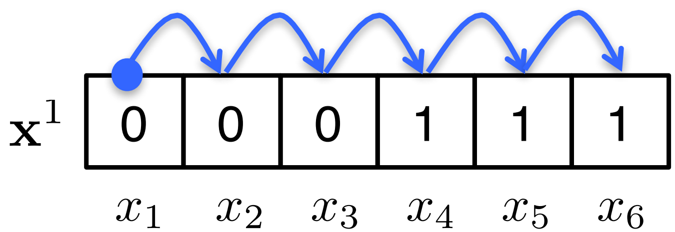
$$P_{(h,k),(h+1,k)} = \frac{\#\ \text{solutions moved from } (h,k) \text{ to } (h+1,k)}{\#\ \text{solutions reached } (h,k)}$$

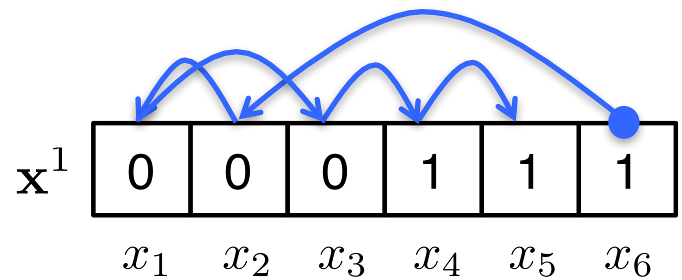
Square Lattice for n=6 GPP instance

A Square Lattice

The Probability Model – The order of variables

In which order should we visit the variables?



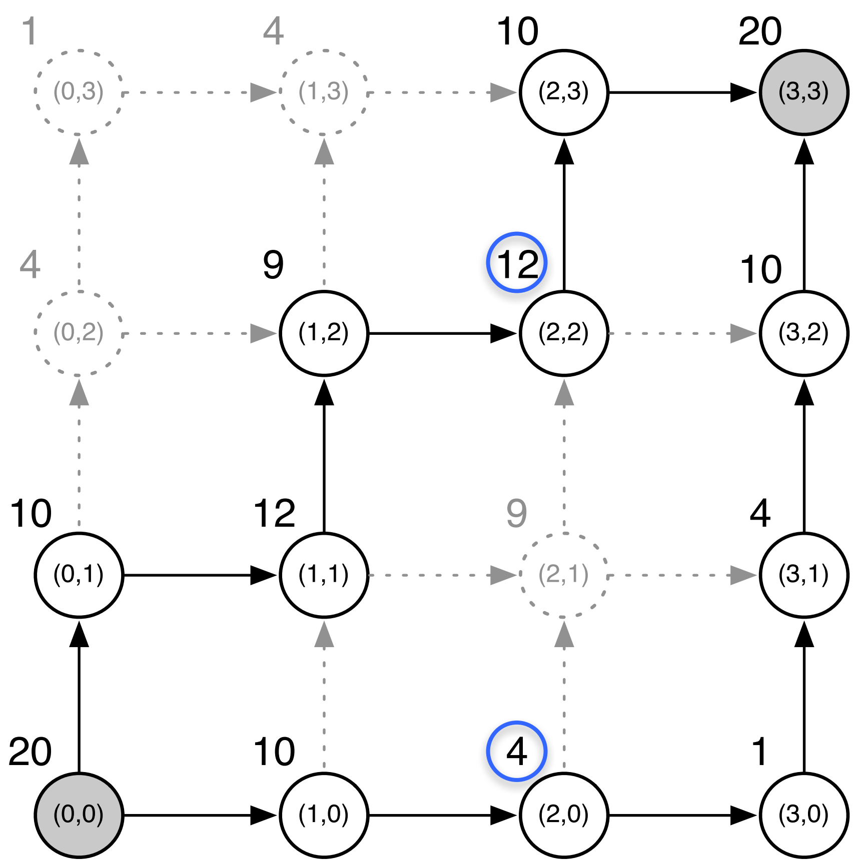


$$\sigma = 123456 \qquad \sigma' = 621345$$

Is there any difference?

A Square Lattice

The Probability Model – The order of variables



Square Lattice for n=6 GPP instance

Vertex close to the diagonal are more frequently visited

Taking a decision on these points is largely uncertain:

At (2,2) -- the solution has 2 zeros and 2 ones.

At (2,0), the preceeding positions are filled with zeros.

A Square Lattice

The Probability Model – The order of variables

- Map the best solution to the external border

$$\mathbf{x} = (1,0,0,1,1,0)$$

  - First visit 0000…, and then ...1111.

$$\sigma = (\{2,3,6\},\{1,4,5\})$$

  - And the order within the subsets?

From each set, choose the item that minimizes the cut-size, in alternated rounds.

Experimental Study

Experimental Setting

UMDA *, TREE*[1] , Lattice
(*) adapted sampling

Algorithms

Pop-size: 10n
Sel-size: 5n
Off-size: 10n
Max evals.: $100n^2$
10 repetitions

Parameters

22 Instances (Johnson et al.)

G-type and U-type

n=124, 250, 500, 1000

Benchmarks

Edges with no observations are assigned uniform probability

Constructive at every 40 iterations

Sample
$x_0 = 0$

Lattice Settings

[1]M. Pelikan, S. Tsutsui, and R. Kalapala, *Dependency Trees, Permutations and Quadratic Assignment Problem,* Medal Report No. 2007003 Tech. Rep. 2007.

Experimental Study

Results - Performance

Instance	Best Fitness	ARPD		
		Lattice	UMDA	Tree
G124.02	13	0,32	0,61	0,19
G124.16	449	0,02	0,05	0,01
G250.01	31	0,33	0,49	0,20
G250.02	118	0,07	0,14	0,06
G250.04	360	0,04	0,10	0,03
G250.08	830	0,01	0,05	0,01
G500.005	61	0,30	0,40	0,08
G500.01	234	0,09	0,21	0,07
G500.02	642	0,03	0,11	0,03
G500.04	1754	0,02	0,06	0,02
G1000.0025	131	2,96	3,20	0,74
G1000.005	496	1,22	1,28	0,88
G1000.01	1420	0,56	0,66	0,62
G1000.02	3450	0,35	0,40	0,39
U500.05	23	1,17	1,89	0,57
U500.10	61	1,05	1,12	0,57
U500.20	185	0,56	0,87	0,44
U500.40	412	0,41	0,38	0,28
U1000.05	77	1,62	12,83	2,39
U1000.10	170	1,67	11,67	3,73
U1000.20	352	1,67	10,58	4,94
U1000.40	862	1,53	3,24	2,29

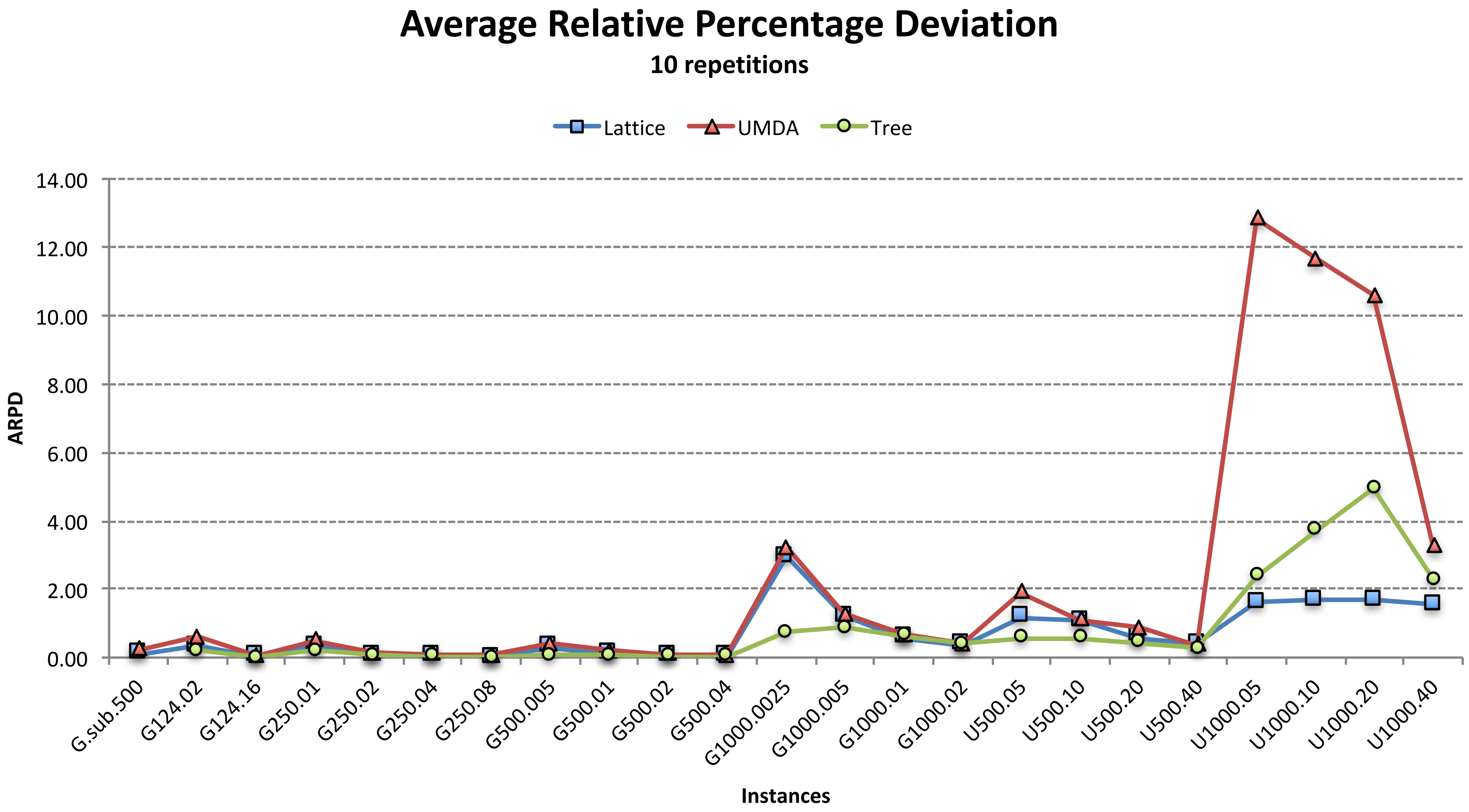
Experimental Study

Results - Performance

Instance	Best Fitness	ARPD		
		Lattice	UMDA	Tree
G124.02	13	0,32	0,61	**0,19**
G124.16	449	0,02	0,05	**0,01**
G250.01	31	0,33	0,49	**0,20**
G250.02	118	0,07	0,14	**0,06**
G250.04	360	0,04	0,10	**0,03**
G250.08	830	**0,01**	0,05	**0,01**
G500.005	61	0,30	0,40	**0,08**
G500.01	234	0,09	0,21	**0,07**
G500.02	642	**0,03**	0,11	**0,03**
G500.04	1754	**0,02**	0,06	**0,02**
G1000.0025	131	2,96	3,20	**0,74**
G1000.005	496	1,22	1,28	**0,88**
G1000.01	1420	**0,56**	0,66	0,62
G1000.02	3450	**0,35**	0,40	0,39
U500.05	23	1,17	1,89	**0,57**
U500.10	61	1,05	1,12	**0,57**
U500.20	185	0,56	0,87	**0,44**
U500.40	412	0,41	0,38	**0,28**
U1000.05	77	**1,62**	12,83	2,39
U1000.10	170	**1,67**	11,67	3,73
U1000.20	352	**1,67**	10,58	4,94
U1000.40	862	**1,53**	3,24	2,29

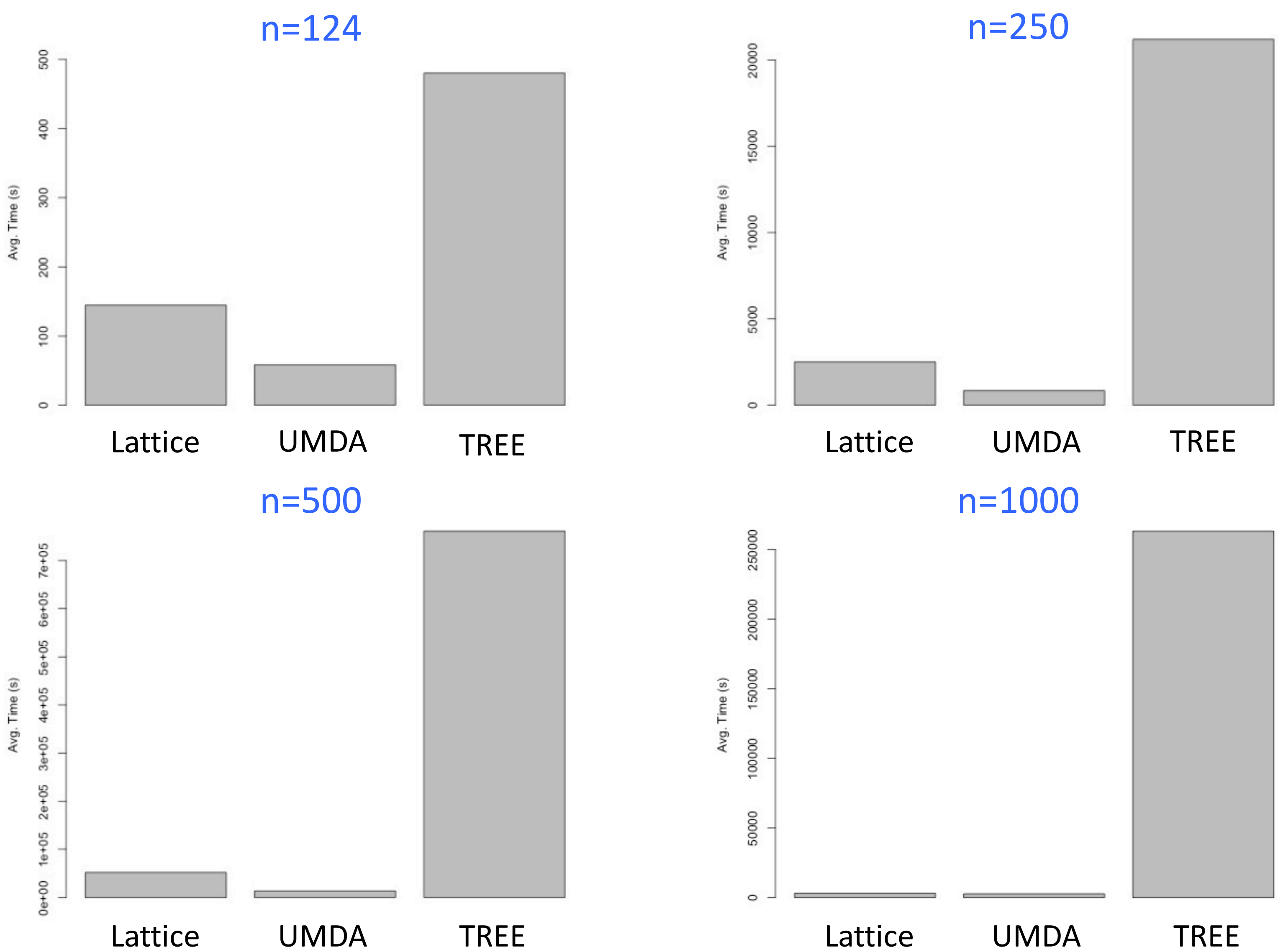
Experimental Study

Results - Performance



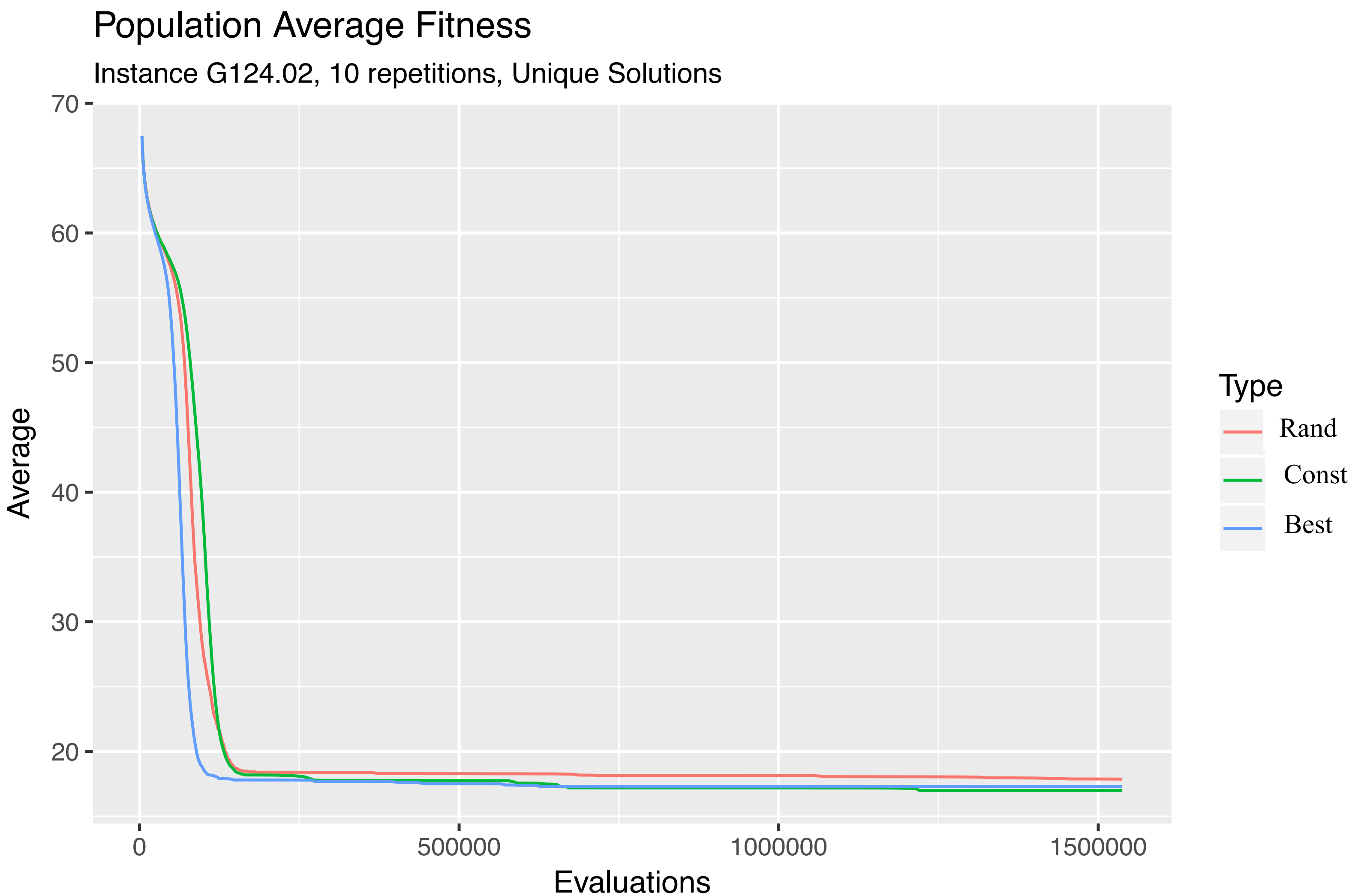
Experimental Study

Results – Time Consumption



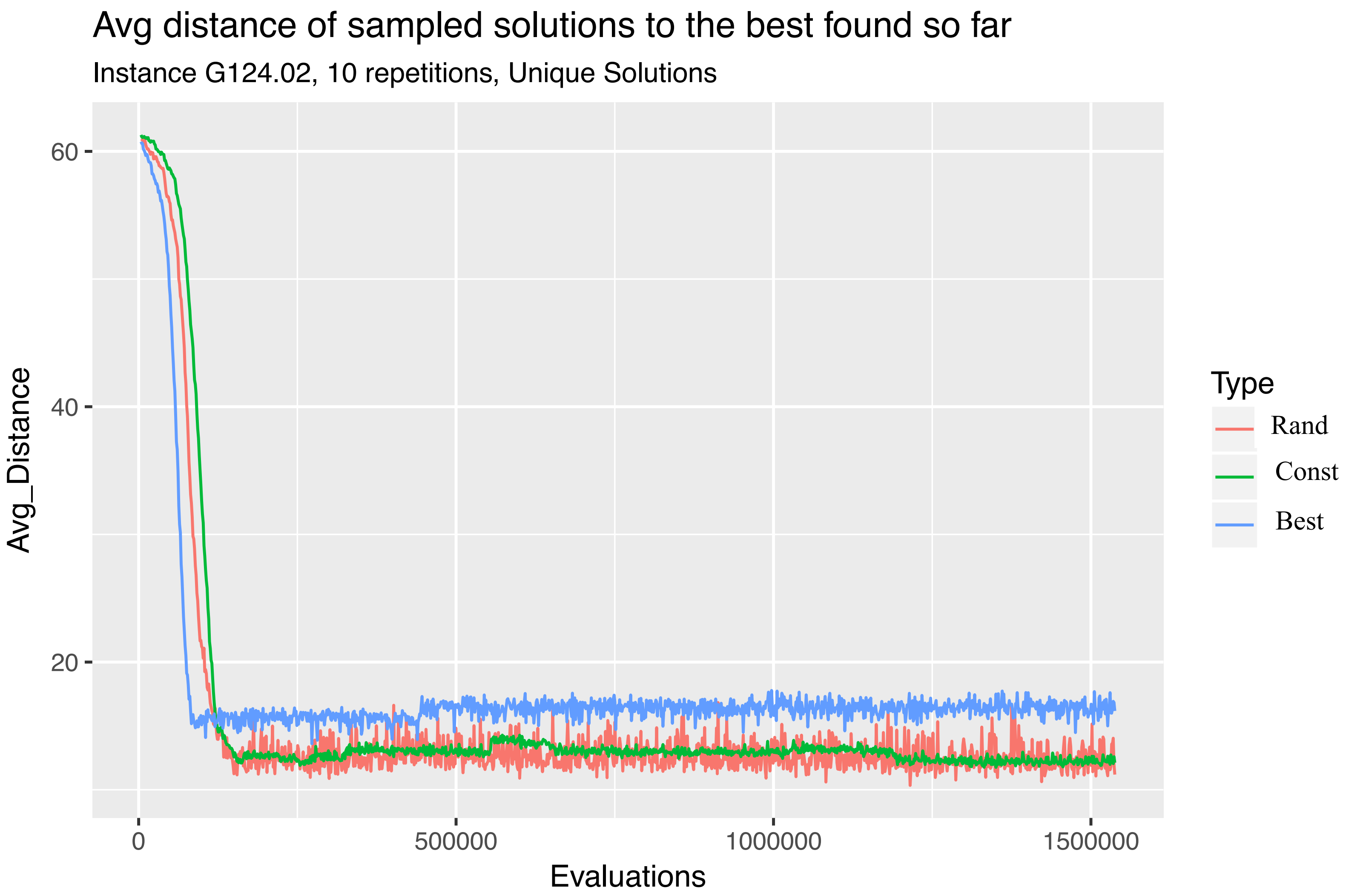
Experimental Study

The influence of the order



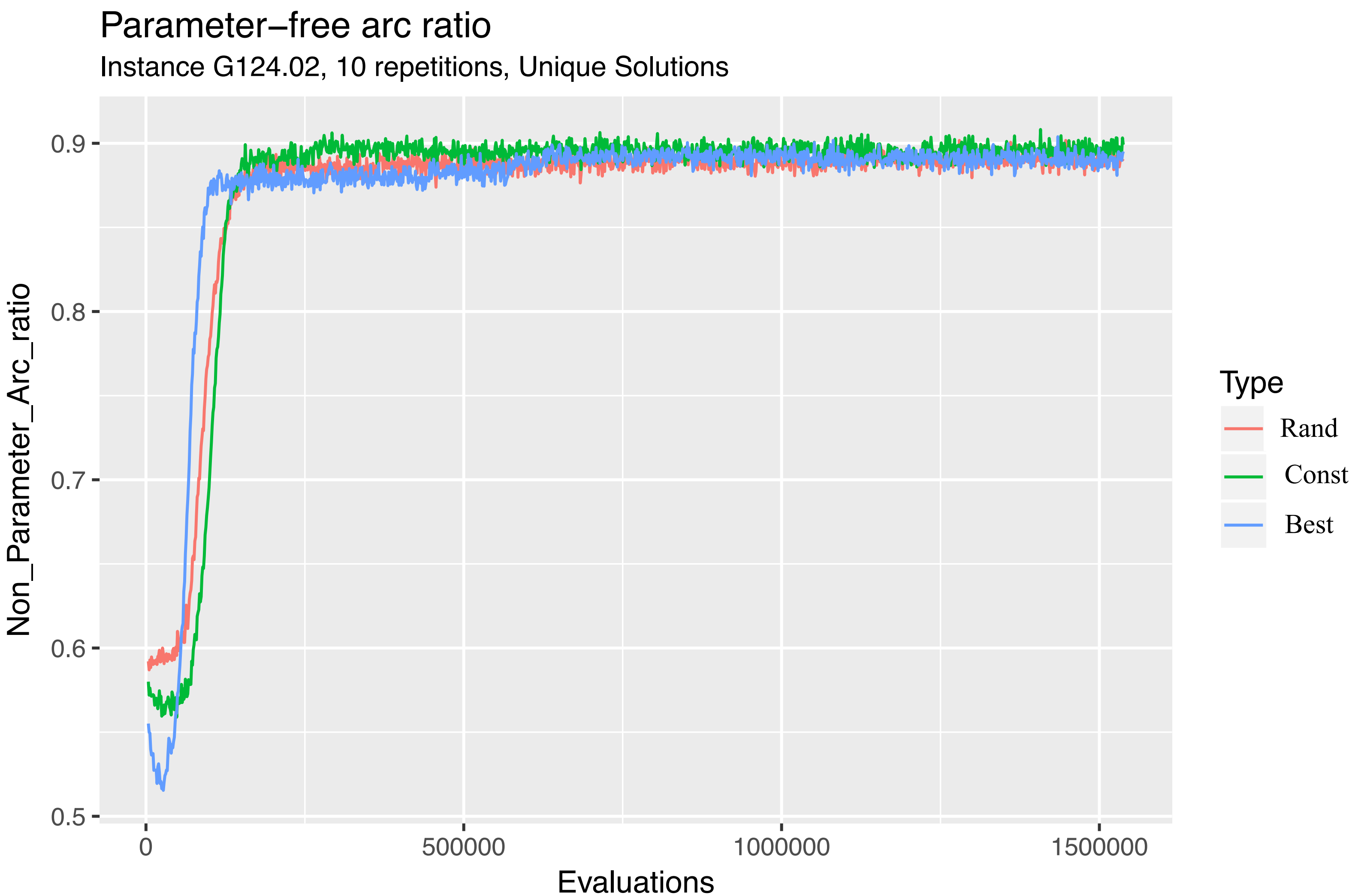
Experimental Study

The influence of the order



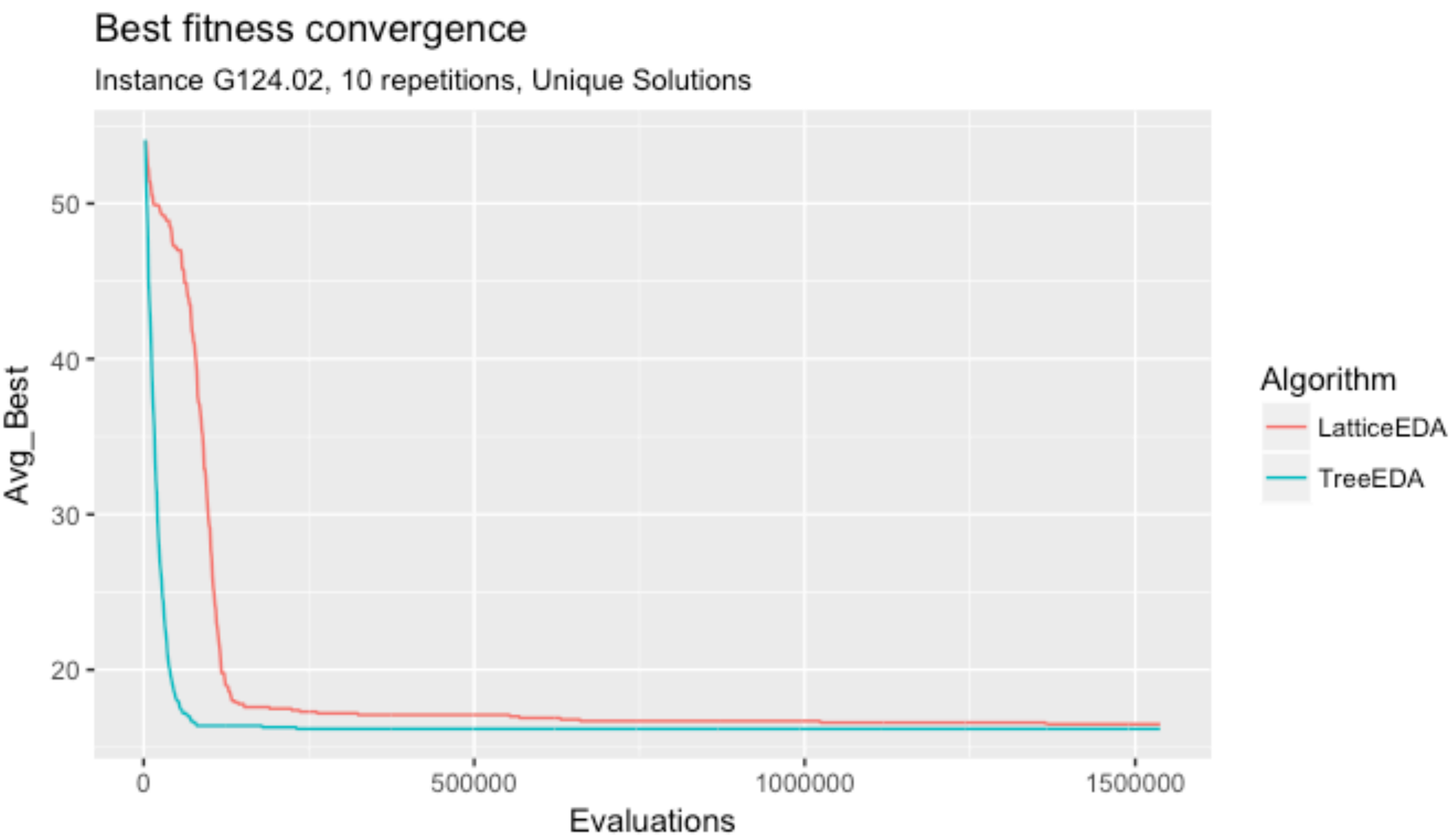
Experimental Study

The influence of the order



Experimental Study

Convergence Comparison



Experimental Study

Results - Performance

Instance	Best Fitness	ARPD Lattice	Lattice 2	Lattice 4	UMDA	Tree
G124.02	13	0,32	0,32	0,45	0,61	**0,19**
G124.16	449	0,02	0,02	0,02	0,05	**0,01**
G250.01	31	0,33	0,44	0,49	0,49	**0,20**
G250.02	118	0,07	0,09	0,05	0,14	**0,06**
G250.04	360	0,04	0,04	0,04	0,10	**0,03**
G250.08	830	**0,01**	0,02	0,02	0,05	**0,01**
G500.005	61	0,30	0,46	0,46	0,40	**0,08**
G500.01	234	0,09	0,12	0,13	0,21	**0,07**
G500.02	642	**0,03**	0,04	0,06	0,11	**0,03**
G500.04	1754	**0,02**	0,03	0,03	0,06	**0,02**
G1000.0025	131	2,96	--	--	3,20	**0,74**
G1000.005	496	1,22	--	--	1,28	**0,88**
G1000.01	1420	**0,56**	--	--	0,66	0,62
G1000.02	3450	**0,35**	--	--	0,40	0,39
U500.05	23	1,17	1,30	1,04	1,89	**0,57**
U500.10	61	1,05	0,81	1,20	1,12	**0,57**
U500.20	185	0,56	0,62	0,45	0,87	**0,44**
U500.40	412	0,41	0,58	0,83	0,38	**0,28**
U1000.05	77	**1,62**	--	--	12,83	2,39
U1000.10	170	**1,67**	--	--	11,67	3,73
U1000.20	352	**1,67**	--	--	10,58	4,94
U1000.40	862	**1,53**	--	--	3,24	2,29

Experimental Study

Results - Performance

Instance	Best Fitness	ARPD				
		Lattice	Lattice 2	Lattice 4	UMDA	Tree
G124.02	13	0,32	0,32	~~0,45~~	0,61	**0,19**
G124.16	449	0,02	0,02	0,02	0,05	**0,01**
G250.01	31	0,33	~~0,44~~	~~0,49~~	0,49	**0,20**
G250.02	118	0,07	~~0,09~~	**0,05**	0,14	0,06
G250.04	360	0,04	0,04	0,04	0,10	**0,03**
G250.08	830	**0,01**	~~0,02~~	~~0,02~~	0,05	**0,01**
G500.005	61	0,30	~~0,46~~	~~0,46~~	0,40	**0,08**
G500.01	234	0,09	~~0,12~~	~~0,13~~	0,21	**0,07**
G500.02	642	**0,03**	~~0,04~~	~~0,06~~	0,11	**0,03**
G500.04	1754	**0,02**	~~0,03~~	~~0,03~~	0,06	**0,02**
G1000.0025	131	2,96	--	--	3,20	**0,74**
G1000.005	496	1,22	--	--	1,28	**0,88**
G1000.01	1420	**0,56**	--	--	0,66	0,62
G1000.02	3450	**0,35**	--	--	0,40	0,39
U500.05	23	1,17	~~1,30~~	1,04	1,89	**0,57**
U500.10	61	1,05	~~0,81~~	~~1,20~~	1,12	**0,57**
U500.20	185	0,56	~~0,62~~	0,45	0,87	**0,44**
U500.40	412	0,41	~~0,58~~	~~0,83~~	0,38	**0,28**
U1000.05	77	**1,62**	--	--	12,83	2,39
U1000.10	170	**1,67**	--	--	11,67	3,73
U1000.20	352	**1,67**	--	--	10,58	4,94
U1000.40	862	**1,53**	--	--	3,24	2,29

Conclusions

The experiments support the validity of our research line:

Designing probability models exclusively on the set of feasible solutions

Competitive for small instances, and better in large instances

Low time complexity

Future Work

Many aspects to be faced in this work

Develop the idea of uncertainty of the paths in the lattice

Analyze the effect of other orderings

Use information of the population as Tree (mutual informations)

Ordering of the variables

Larger benchmarks

Understand the dynamics of the Lattice for different problem sizes

Experimentation

Square Lattice for $K \geq 3$

The probability model

Future Research Lines

Other models

Distance-based exponential probability models

A distance-metric

Spread parameter

Central solution

$$P(x) = \frac{e^{-\theta d(x,\bar{x})}}{\sum_{i=1}^{k} e^{-\theta d(x_i,\bar{x})}}$$

Normalization function

The size of the search space. GPP:

$$k = \binom{n}{n/2}$$

Develop efficient **learning** and **sampling** methods

Challenges

$$L(\theta,\bar{x}\mid x) = \prod_{i=1} \frac{e}{\sum_{j=1} e^{-\theta d(x_j,\bar{x})}} \Rightarrow \log L(\theta,\bar{x}\mid x) = -\sum_{i=1}\left(\theta d(x_i,\bar{x}) + \log\left(\sum_{j=1} e^{-\theta d}\right.\right.$$

Lo siguiente es derivar la función e igualar a 0:

$$\log L(\theta,\bar{x}\mid x) = -\sum_{i=1}^{N}\theta d(x_i,\bar{x}) - \sum_{i=1}^{N}\log\underbrace{\left(\sum_{j=1}^{K} e^{-\theta d(x_j,\bar{x})}\right)}_{f(\theta)} = -\sum_{i=1}^{N}\theta d(x_i,\bar{x}) - N\log\left(\sum_{j=1}^{K} e^{-\theta d(x_j,\bar{x})}\right.$$

$$\frac{d\log f(\theta)}{d\theta} = \frac{f'(\theta)}{f(\theta)}$$

derivar respecto a θ

$$= -\sum_{i=1}^{N} d(x_i,\bar{x}) + N \cdot \frac{\sum_{j=1}^{K} d(x_j,\bar{x})\cdot e^{-\theta d(x_j,\bar{x})}}{\sum_{j=1}^{K} e^{-\theta d(x_j,\bar{x})}} = -\sum_{i=1}^{N} d(x_i,\bar{x}) + N\cdot\sum_{j=1}^{K}\left(\frac{d(x_j,\bar{x})\cdot e^{-\theta d(x_j,\bar{x})}}{e^{-\theta d(x_j,\bar{x})}}\right)$$

$$= -\sum_{i=1}^{N} d(x_i,\bar{x}) + N\sum_{j=1}^{K} d(x_j,\bar{x}) = 0$$

← El parámetro Theta ha desaparecido... :(

$$= -N\bar{d} + N\sum_{j=1}^{K} d(x_j,\bar{x}) = N\left(-\bar{d} + \sum_{j=1}^{K} d(x_j,\bar{x})\right) = 0$$

X ← ESTA MAL

$$= -N\bar{d} + N\,\frac{\sum_{j=1}^{K} d(x_j,\bar{x})\cdot e^{-\theta d(x_j,\bar{x})}}{\sum_{j=1}^{K} e^{-\theta d(x_j,\bar{x})}} = 0$$

← ¿Podemos aproximar la Theta mediante un método numérico?

Otra posible vía reformulando la constante de normalización:

$$P(x) = \frac{e^{-\theta d(x,\bar{x})}}{\sum_{j=0,\ j\%2==0}^{n}\binom{n/2}{j/2}^2 e^{-\theta j}}$$

y por lo tanto el log-likelihood queda como:

$$-N\bar{d} + N\left(\frac{\sum_{j=0}^{n} j\binom{n/2}{j/2}^2 e^{-\theta j}}{\sum_{j=0}^{n}\binom{n/2}{j/2}^2 e^{-\theta j}}\right) = 0$$

Dealing with constraints in estimation of distribution algorithms: a different approach

Josu Ceberio, Alexander Mendiburu, Jose A. Lozano

Thank you for your attention!!!

Intelligent Systems Group

eman ta zabal zazu

Universidad del País Vasco Euskal Herriko Unibertsitatea

FACULTY OF COMPUTER SCIENCE
UNIVERSITY OF THE BASQUE COUNTRY