INFERENCE SCHEMES FOR M BEST SOLUTIONS FOR SOFT CSPS

Emma Rollón, Natalia Flerova and Rina Dechter

erollon@lsi.upc.edu, flerova@ics.uci.edu, dechter@ics.uci.edu
Universitat Politècnica de Catalunya
University of California, Irvine

Summary

- Optimization problems:
  - Finding the best solution
  - Finding the m-best solutions

- Applications of the m-best solutions:
  - Set of diverse solutions desired (e.g., haplotaping)
  - Constraints are hard to formalized (e.g., portfolio mgmt)
  - Sensitivity analysis (e.g., biological sequence alignment)

Summary

- Previous works on the m-best tasks:
  - Compute the m-best solutions by successively computing the best solution, each time using a slightly different reformulation of the original problem.
    - Lawler, 1972; Nilsson, 1998; Yanover and Weiss, 2004;
  - Compute the m best solutions in a single pass of algorithm, using message passing/propagation
    - Seroussi and Golmard, 1994 ; Elliot, 2007;

Summary

- Our contribution:
  - We provide a formalization of the m-best task within the unifying framework of c-semiring, making many known inference schemes immediately applicable.
  - In particular, we focus on Graphical Models and extend:
    - Bucket Elimination (exact algorithm)
    - Mini-Bucket Elimination (bounding algorithm)
  - We show how to tighten the bound on the best solution using a bound on the m-best solutions.

1-best vs. m-best optimization

1-best optimization	m-best optimization

Variables : $X = \{X_1, X_2, X_3, \ldots, X_n\}$

Finite domain values: $D = \{D_1, D_2, \ldots, D_n\}$

Objective function: $F : \bigotimes_{i=1}^{n} D_i \rightarrow A$

A is a totally ordered set (<) of valuations

$F(t)$ such that

$\forall t' \notin t : F(t) \leq F(t')$

$\{F(t_1), \ldots, F(t_m)\}$ such that

$\forall t' \notin t_i, t_j \quad 1 \leq i \leq \mathrm{j} \leq \mathrm{m} \quad F(t_i) \leq F(t_j) \leq F(t')$

Graphical Model

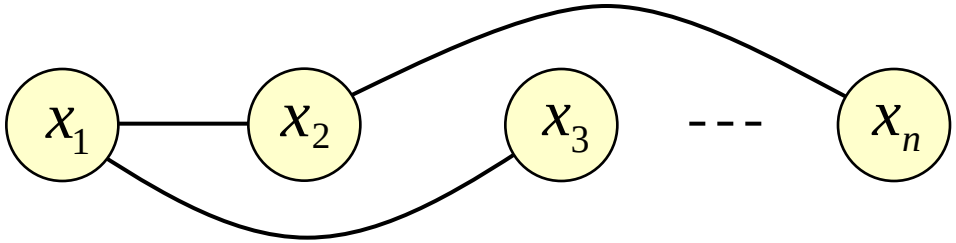
- $X = \{x_1, \ldots, x_n\}$: a set of variables
- $D = \{D_1, \ldots, D_n\}$: a set of domain values
- $\{f_1, \ldots, f_e\}$: a set of local functions

$$f_j : D_Y \rightarrow A$$

$$Y \subseteq X \quad \text{scope of} f_j$$

- $\otimes$: combination operator over functions

- Interaction graph:



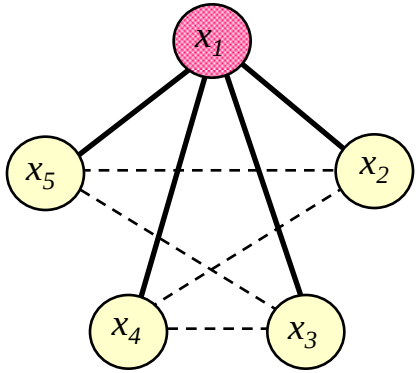
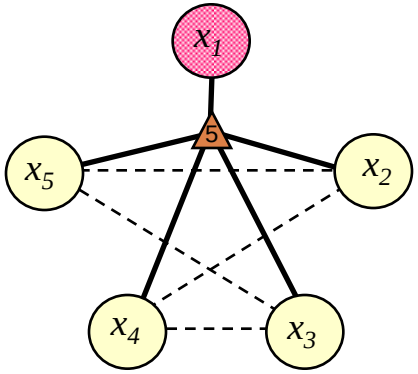
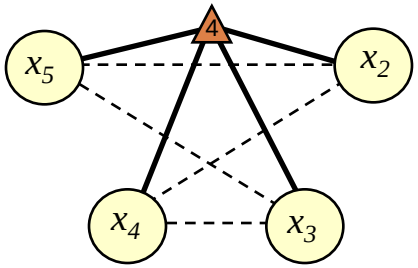
Graphical Model

- Global view (objective function): $\mathrm{F}(X) = \bigotimes_{k=1}^{e} f_k$
- Reasoning task: $\mathrm{F}(X) \Downarrow_X$
- Particular instantiations:

A	$\otimes$	$\Downarrow$	Task
$N \cup \{\infty\}$	$+$	$\min$	WCSP
$[0...1]$	$\times$	$\max$	MPE
$[0...1]$	$\times$	$+$	P(e)

Bucket Elimination

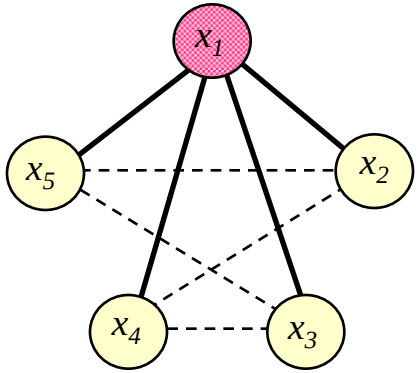
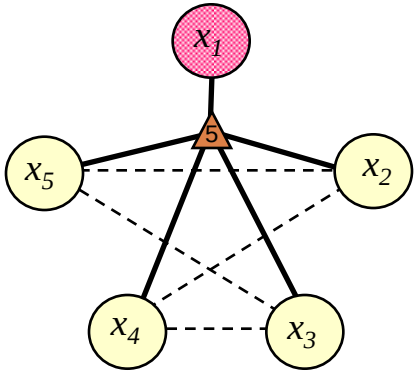
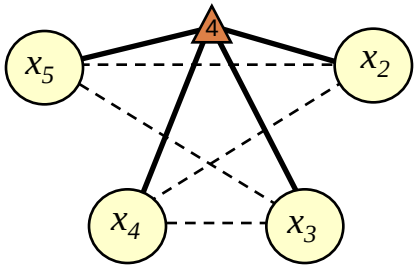
BE

Select a var	Combination $\otimes$	Marginalization $\Downarrow$	Output
			$\lambda() = \left(\bigotimes_{k=1}^{e} f_k \right) \Downarrow_X$

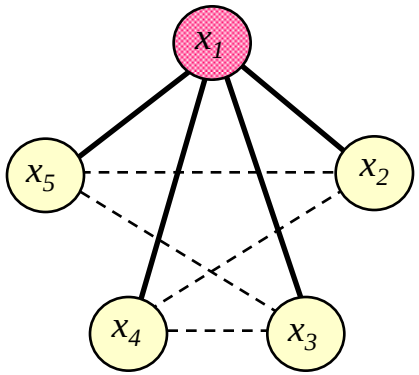
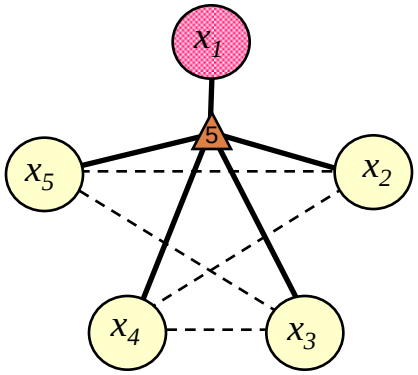
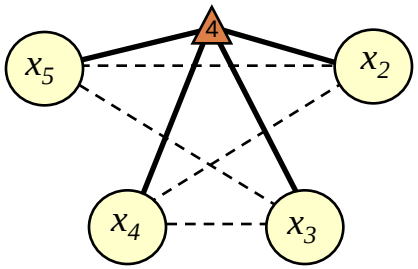
Complete and correct: whenever the task can be defined over a semiring

[Shafer et. Al, Srivanas et al, Kohlas et al.]

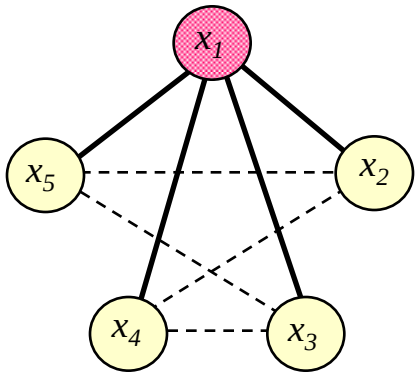
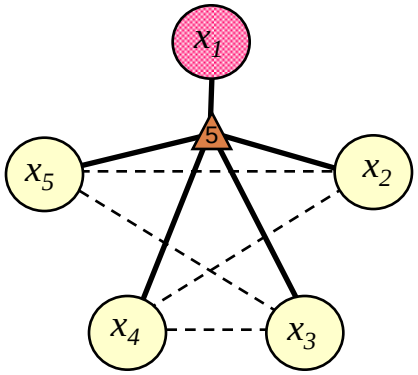
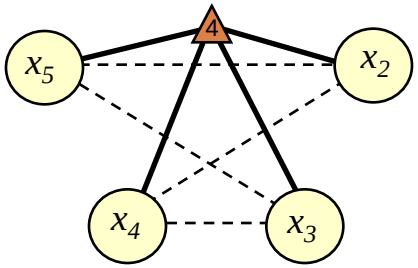
Bucket Elimination

	Select a var	Combination $\otimes$	Marginalization $\Downarrow$	Output
BE				$\lambda() = \left(\bigotimes_{k=1}^{e} f_k \right) \Downarrow_X$
elim-opt		$\otimes \equiv +$	$\Downarrow \equiv \min$	$\lambda() = c_1$ = WCSP

Bucket Elimination

	Select a var	Combination $\otimes$	Marginalization $\Downarrow$	Output
BE				$\lambda() = \left(\bigotimes_{k=1}^{e} f_k \right) \Downarrow_X$
elim-opt		$\otimes \equiv +$	$\Downarrow \equiv \min$	$\lambda() = c_1$ = WCSP
elim-m-opt		$\otimes \equiv ??$	$\Downarrow \equiv ??$	$\lambda() = \{c_1,\ldots,c_m\}$ =m-best WCSP

Bucket Elimination

	Select a var	Combination $\otimes$	Marginalization $\Downarrow$	Output
BE				$\lambda() = \left(\bigotimes_{k=1}^{e} f_k \right) \Downarrow_X$
elim-opt		$\otimes \equiv +$	$\Downarrow \equiv \min$ Each tuple is the best cost extension to x1	$\lambda() = c_1$ = WCSP
elim-m-opt		$\otimes \equiv ??$	$\Downarrow \equiv ??$ Each tuple has to be the m-best cost extensions to x1	$\lambda() = \{c_1,\ldots,c_m\}$ =m-best WCSP

From 1-best to m-best optimization

	WCSP	m-best WCSP (m = 2)
Functions	$f : l(Y) \rightarrow [0...1]$ $f(t) = c_1$	$\bar{f} : l(Y) \rightarrow [0...1]^m$ $\bar{f}(t) = \{c_1, c_2\}$
$\otimes$	$+$ $f(t) = 6; g(t) = 5$	$\overline{+}$ $\bar{f}(t) = \{3,5\}; \bar{g}(t) = \{1,6\}$
$\Downarrow$	$\min$ $f(x \leftarrow a) = 6;\ f(x \leftarrow b) = 5$	$\overline{\min}$ $\bar{f}(x \leftarrow a) = \{3,5\}; \bar{f}(x \leftarrow b) = \{1,6\}$

From 1-best to m-best optimization

	WCSP	m-best WCSP (m = 2)
Functions	$f : l(Y) \rightarrow [0...1]$ $f(t) = c_1$	$\bar{f} : l(Y) \rightarrow [0...1]^m$ $\bar{f}(t) = \{c_1, c_2\}$
$\otimes$	$+$ $f(t) = 6; g(t) = 5$ $f(t) + g(t) = 6 + 5 = 11$	$\overline{+}$ $\bar{f}(t) = \{3,5\}; \bar{g}(t) = \{1,6\}$ $\bar{f}(t) \overline{+} \bar{g}(t) = \{4, \cancel{9}, 6, \cancel{11}\} = \{4,6\}$
$\Downarrow$	$\min$ $f(x \leftarrow a) = 6;\ f(x \leftarrow b) = 5$ $\min_x f = \min\{5,6\} = 5$	$\overline{\min}$ $\bar{f}(x \leftarrow a) = \{3,5\}; \bar{f}(x \leftarrow b) = \{1,6\}$ $\overline{\min_x} \bar{f} = \{3, \cancel{5}, 1, \cancel{6}\} = \{1,3\}$

How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$

$1 + 2$

How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$

$\boxed{1+2}$ 1st best

How to combine two ordered sets

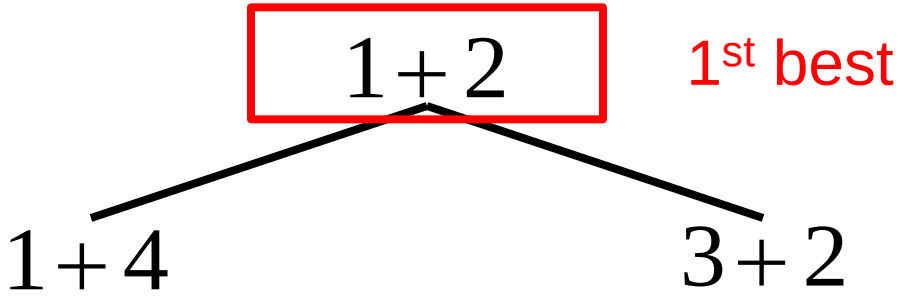
$S = \{1,3,6\}$ $T = \{2,4,5\}$

$\boxed{1+2}$ 1st best

$1+4$

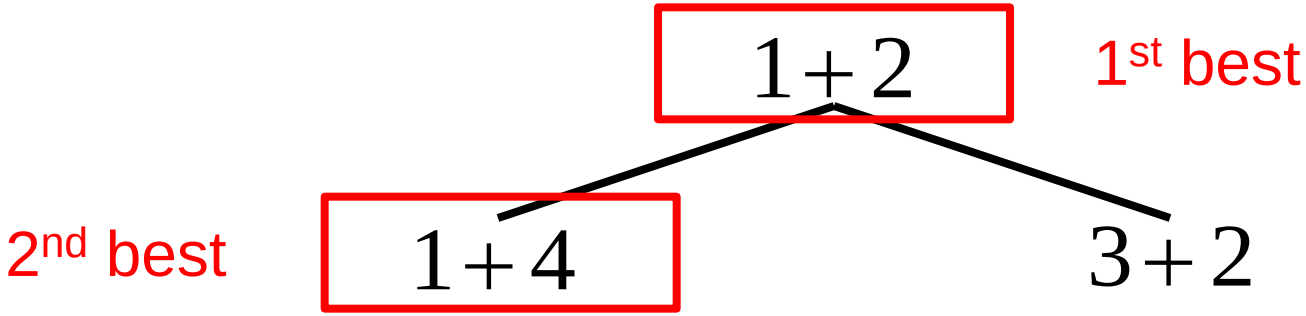
How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$



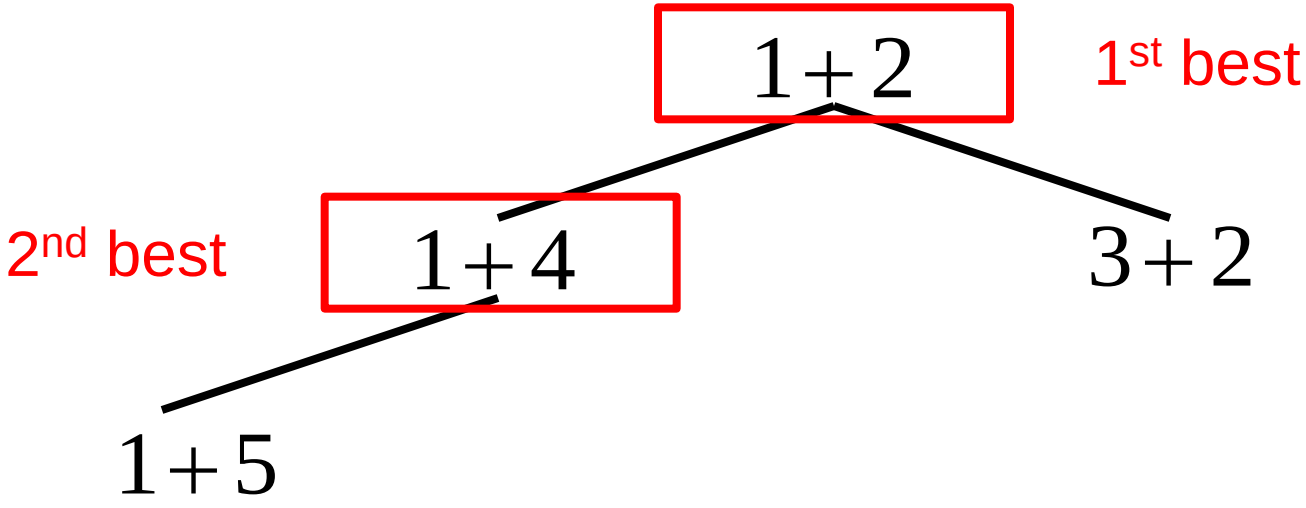
How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$



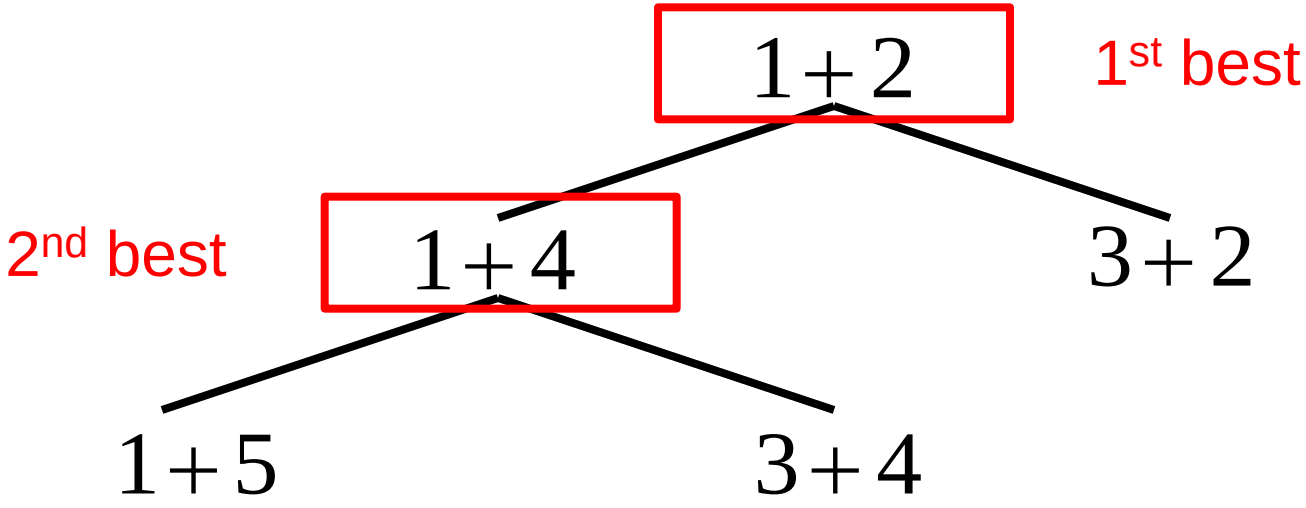
How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$



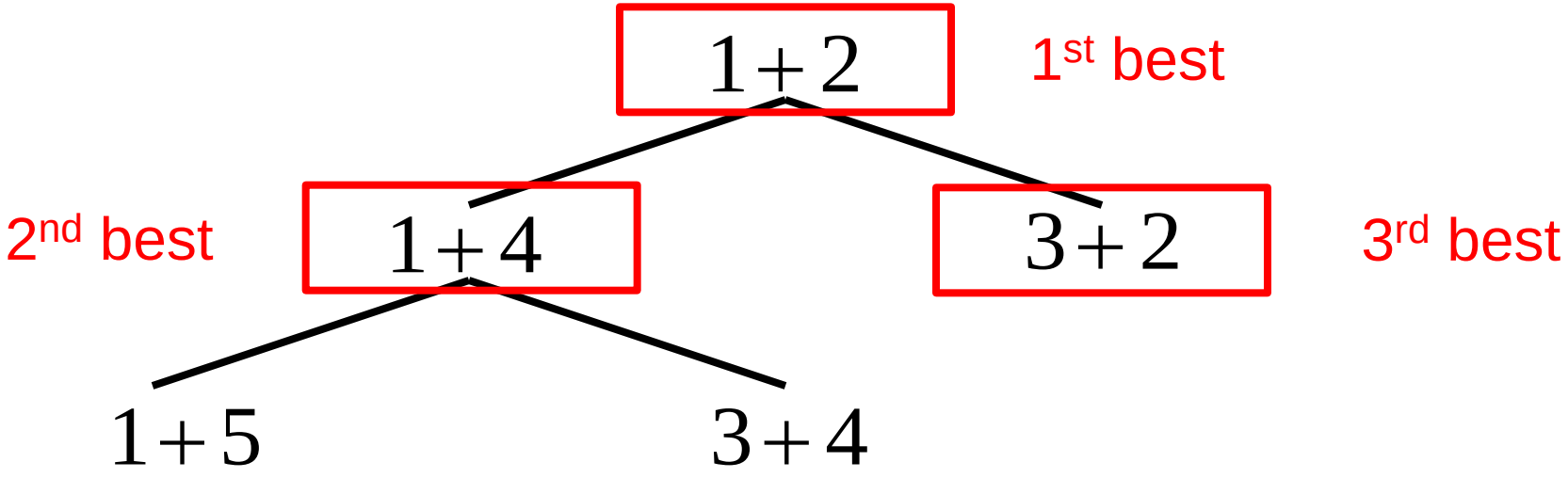
How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$



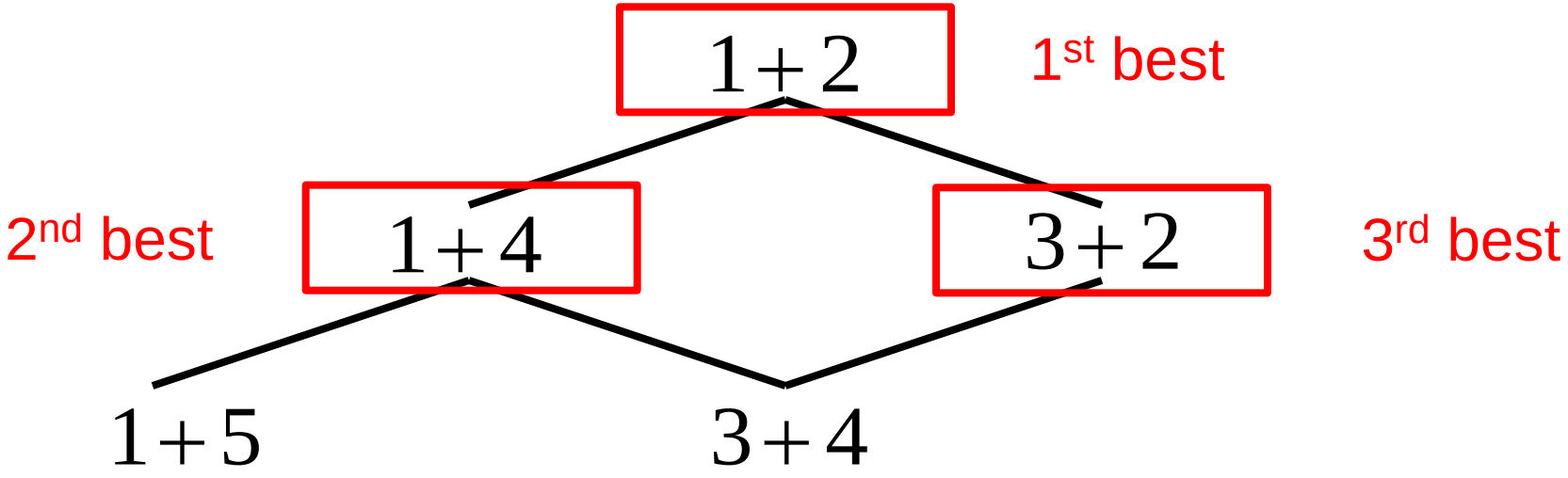
How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$



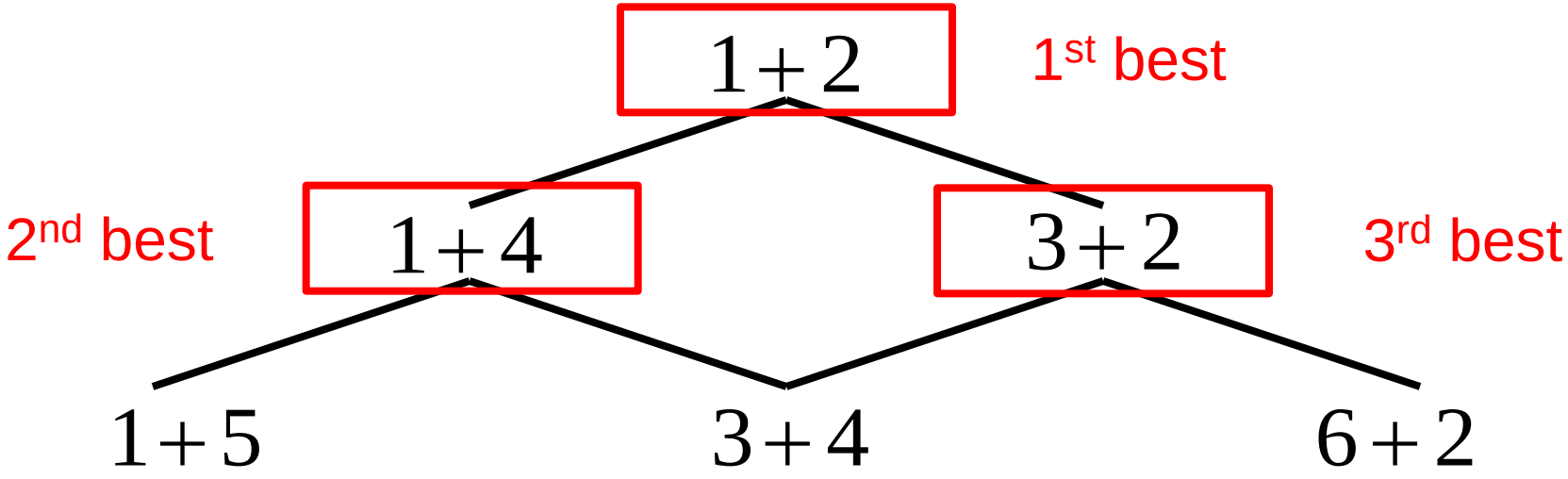
How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$



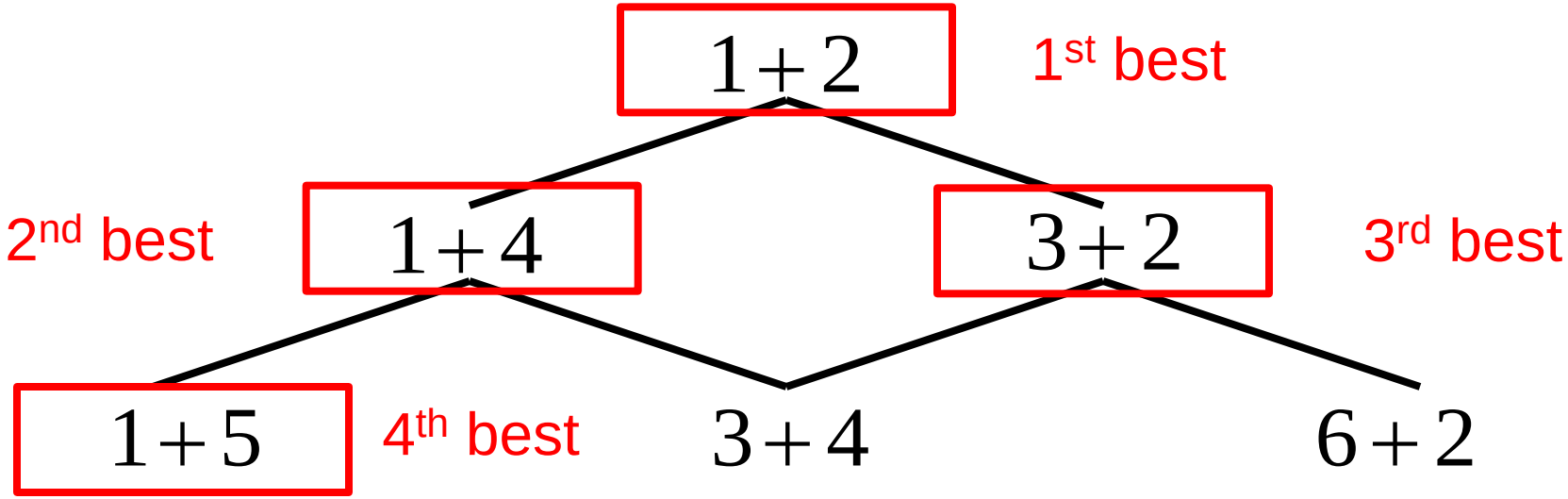
How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$



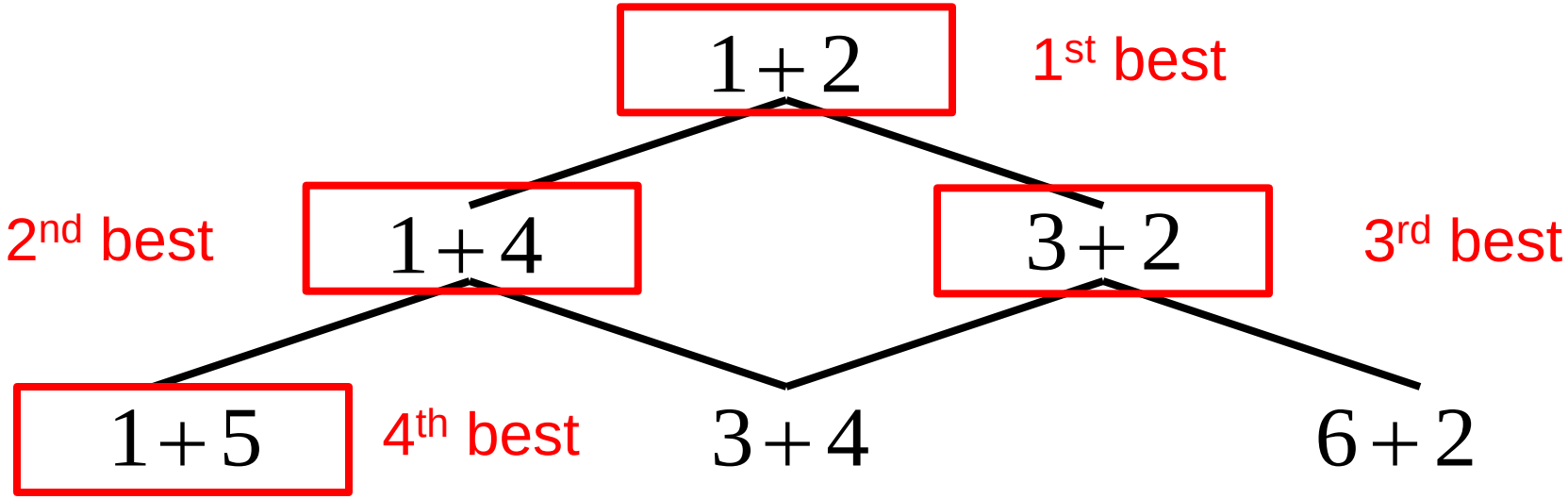
How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$



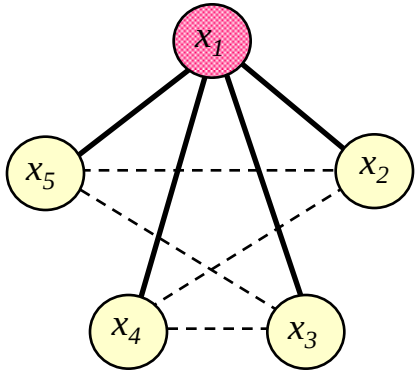
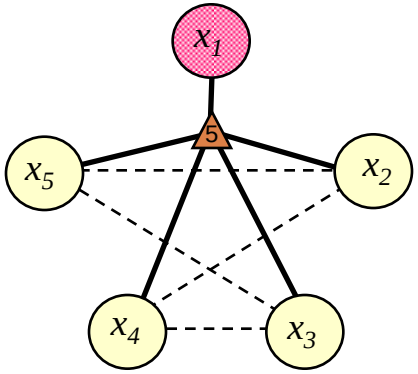
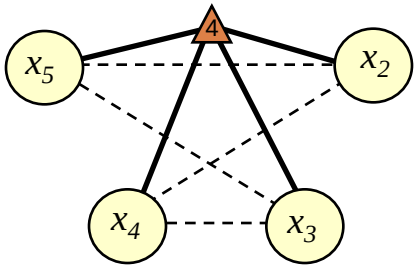
How to combine two ordered sets

$S = \{1,3,6\}$ $T = \{2,4,5\}$

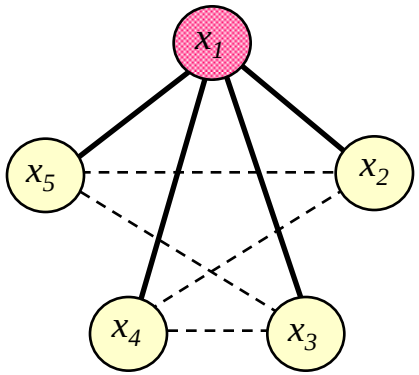
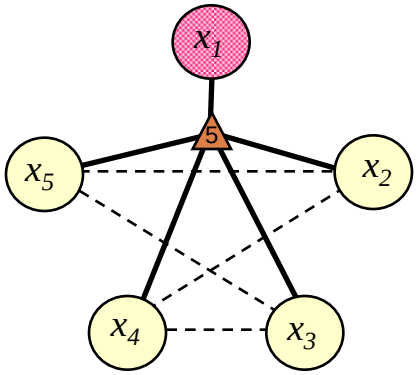
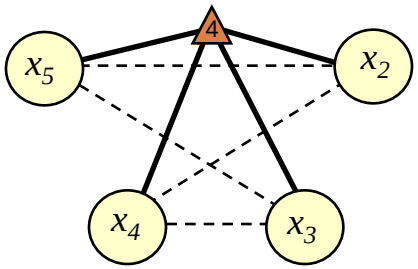


$O(m^2) \longrightarrow O(m * \log(m+1))$

Bucket Elimination

	Select a var	Combination $\otimes$	Marginalization $\Downarrow$	Output
BE				$\lambda() = \left(\bigotimes_{k=1}^{e} f_k \right) \Downarrow_X$
elim-opt		$\otimes \equiv +$	$\Downarrow \equiv \min$	$\lambda() = c_1$ = WCSP
elim-m-opt		$\otimes \equiv \overline{+}$	$\Downarrow \equiv \overline{\min}$	$\lambda() = \{c_1,..,c_m\}$ =m-best WCSP

Bucket Elimination

	Select a var	Combination $\otimes$	Marginalization $\Downarrow$	Output
BE				$\lambda() = \left(\bigotimes_{k=1}^{e} f_k \right) \Downarrow_X$
elim-opt		$\otimes \equiv +$	$\Downarrow \equiv \min$	$\lambda() = c_1$ = WCSP
elim-m-opt		$\otimes \equiv \overline{+}$	$\Downarrow \equiv \overline{\min}$	$\lambda() = \{c_1,..,c_m\}$ = m-best WCSP

Correct and complete: the m-best problem can be formulated as a **commutative semiring** using the new operators

Semirings

- A **commutative semiring** is a triplet $(A,\otimes,\oplus)$, where operators satisfy three axioms:
  - A1. The operation $\oplus$ is associative, commutative and idempotent, and there
  - is an additive identity element called 0 such that $a \oplus 0 = a$ for all $a \in A$.
  - A2. The operation $\otimes$ is also associative and commutative, and there is a
  - multiplicative identity element called 1 such that $a \otimes 1 = a$ for all $a \in A$
  - A3. $\otimes$ distributes over $\oplus$, i.e., $(a \otimes b) \oplus (a \otimes c) = a \otimes (b \oplus c)$
- **Example:** *MPE task* is defined over semiring $K = (R,\times,\max)$, *a CSP* is defined over semiring $K = (\{0, 1\}, \wedge, \vee)$, and *a Weighted CSP* is defined over semiring $K = (N \cup \{\infty\},+,\min)$.
- **It was showed that the correctness of inference algorithms over a reasoning task P is ensured whenever P is defined over a semiring.** [Shafer et. Al, Srivanas et al, Kohlas et al.]

m-space and semirings

Consider an optimization problem: $F : \bigotimes_{i=1}^{n} D_i \rightarrow A$

Let ***S*** be a subset of a set of valuation ***A***.

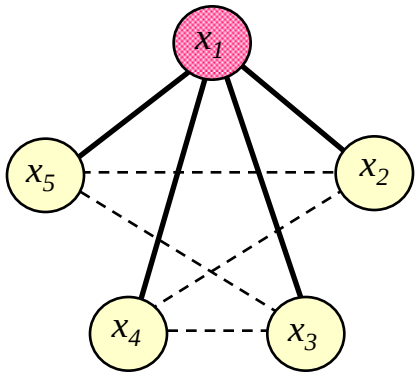
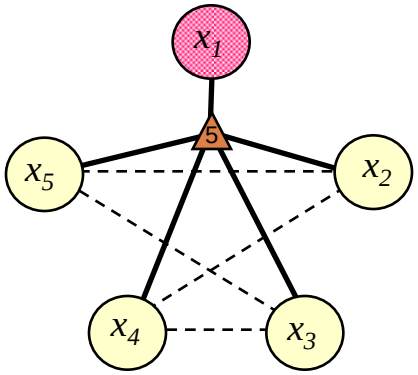
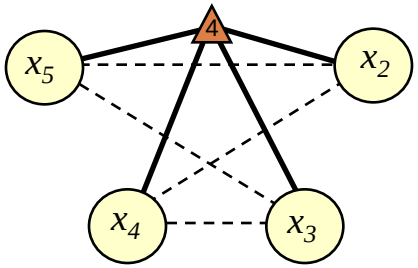
We define the set of ordered m-best elements of ***S*** $S^m = \{s_1, \ldots, s_j\}$ such that $s_1 \leq s_2 \leq \ldots \leq s_j$ where *j*=*m* if |*S*|≥*m* and *j*=|*S*| otherwise, and $\forall s \notin S^m, s_j \leq s$

m-space of *A*: denoted $\overline{A}$, is the set of subsets of ordered m-best elements of ***A***.

Operators $\overline{+}$ and $\overline{\min}$ described above are defined over m-space

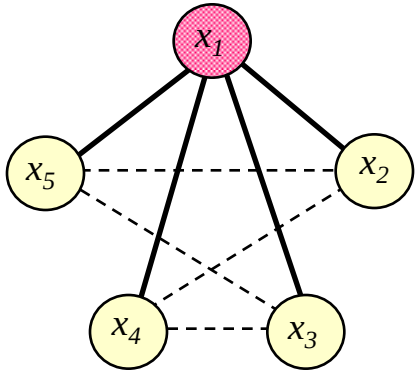
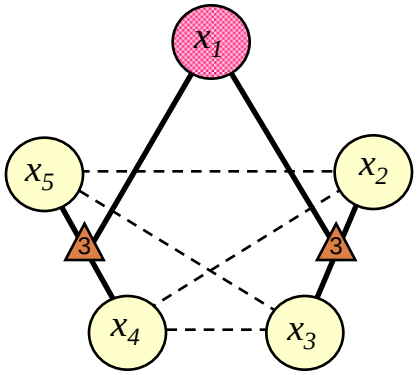
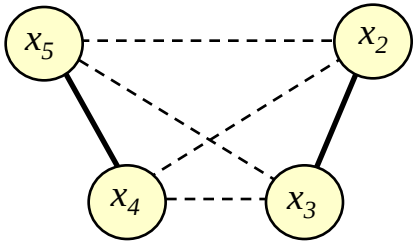
Theorem1: the triplet $(\overline{A}, \overline{+}, \overline{\min})$ is a semiring, defining the m-best WCSP task.

Bucket Elimination

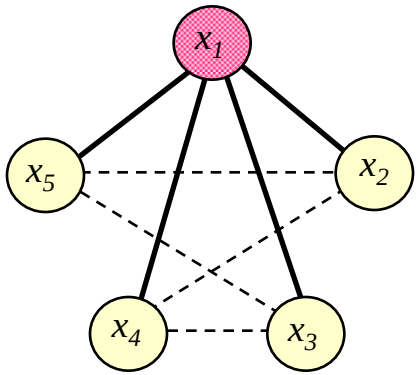
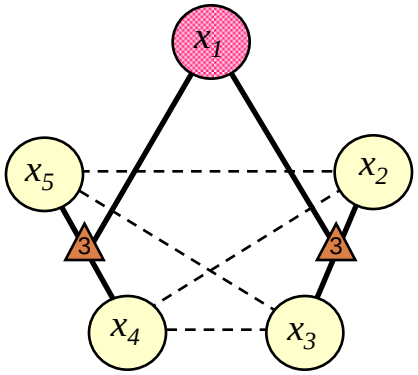
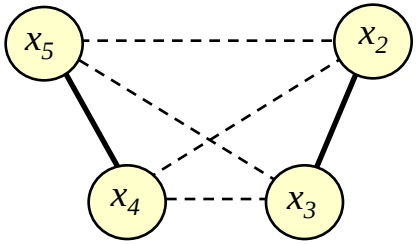
	Select a var	Combination $\otimes$	Marginalization $\Downarrow$	Output
BE				$\lambda() = \left(\bigotimes_{k=1}^{e} f_k \right) \Downarrow_X$
elim-opt		$\otimes \equiv +$	$\Downarrow \equiv \min$	$\lambda() = c_1$ = WCSP
	Time: $O(\exp(w^*+1) \cdot n)$		Space: $O(\exp(w^*) \cdot n)$	
elim-m-opt		$\otimes \equiv \overline{+}$	$\Downarrow \equiv \overline{\min}$	$\lambda() = \{c_1,..,c_m\}$ =m-best WCSP
	Time: $O(\exp(w^*+1) \cdot n \cdot m \cdot \log(m+1))$		Space: $O(\exp(w^*+1) \cdot n \cdot m)$	

Mini-Bucket Elimination

MBE(z = 3)

Select a var	Combination $\otimes$	Marginalization $\Downarrow$	Output
			$\lambda() \geq \left(\bigotimes_{k=1}^{e} f_k \right) \Downarrow_X$

Mini-Bucket Elimination

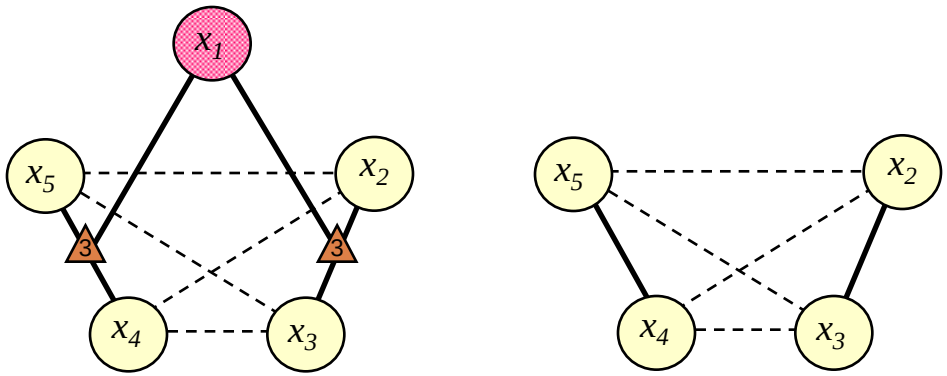
	Select a var	Combination $\otimes$	Marginalization $\Downarrow$	Output
MBE(z = 3)				$\lambda() \geq \left(\bigotimes_{k=1}^{e} f_k \right) \Downarrow_X$
mbe-opt		$\otimes \equiv +$	$\Downarrow \equiv \min$	$\lambda() = a_1 \leq$ Cost(WCSP)
	Time: $O(\exp(z+1) \cdot n)$		Space: $O(\exp(z) \cdot n)$	
mbe-m-opt		$\otimes \equiv \overline{+}$	$\Downarrow \equiv \overline{\min}$	$\lambda() = \{a_1,..,a_m\}$? m-best WCSP
	Time: $O(\exp(z+1) \cdot n \cdot m \cdot \log(m+1))$		Space: $O(\exp(z+1) \cdot n \cdot m)$	

mbe-m-opt Output

m-best WCSP = $\{c_1, c_2, c_3, \ldots, c_k\}$

$c_1 \leq c_2 \leq c_3 \leq \ldots \leq c_k$

mbe-m-opt

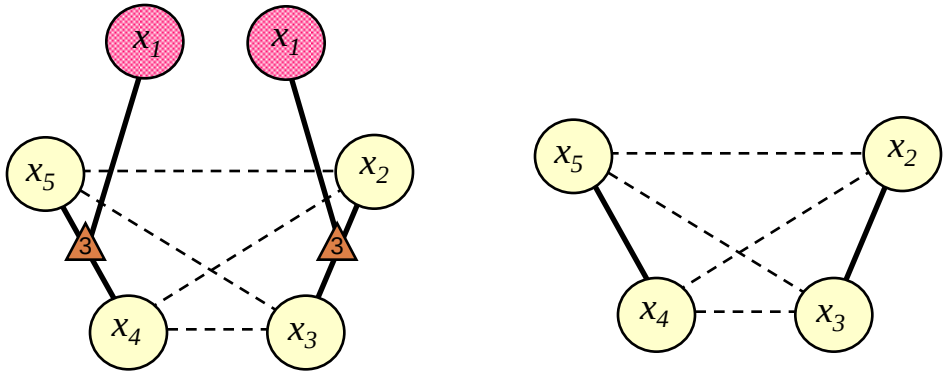


mbe-m-opt Output

m-best WCSP = $\{c_1, c_2, c_3, \ldots, c_k\}$

$c_1 \leq c_2 \leq c_3 \leq \ldots \leq c_k$

mbe-m-opt

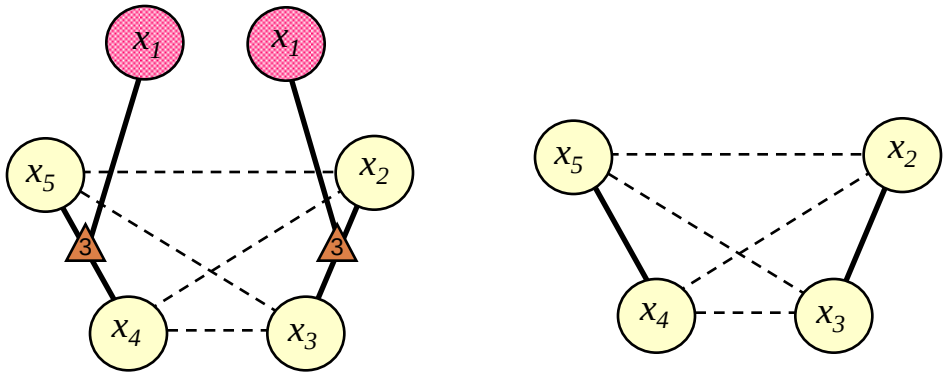


mbe-m-opt Output

m-best WCSP = $\{c_1, c_2, c_3, \ldots, c_k\}$

$c_1 \leq c_2 \leq c_3 \leq \ldots \leq c_k$

mbe-m-opt



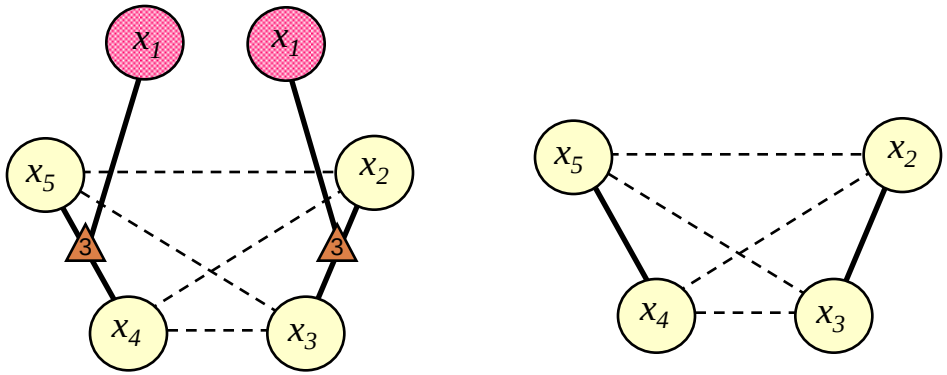
$\lambda() = \{ c_1, c_2, \ldots, c_3, \ldots, \ldots, c_k\}$

mbe-m-opt Output

m-best WCSP = $\{c_1, c_2, c_3, \ldots, c_k\}$

$c_1 \leq c_2 \leq c_3 \leq \ldots \leq c_k$

mbe-m-opt



$\lambda() = \{a_1, a_2, a_3, c_1, a_4, a_5, c_2, a_6, a_7, a_8, \ldots, a_j, c_3, \ldots, a_{j+l}, \ldots, c_k\}$

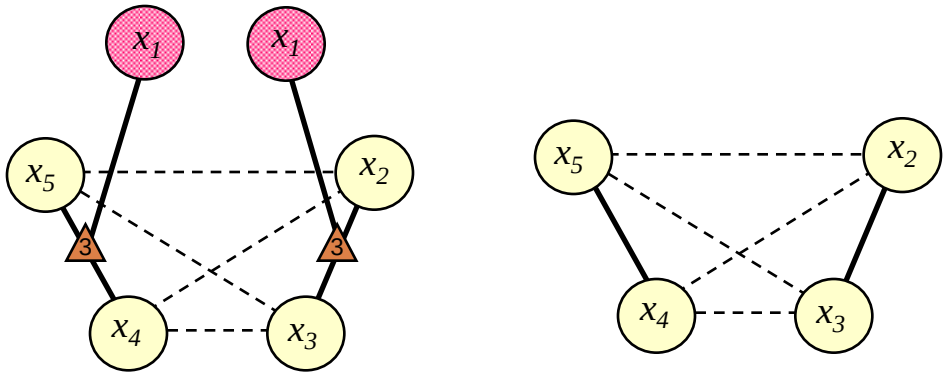
$a_1 \leq a_2 \leq a_3 \leq c_1 \leq a_4 \leq \ldots \leq c_k$

mbe-m-opt Output

m-best WCSP = $\{c_1, c_2, c_3, \ldots, c_k\}$

$c_1 \leq c_2 \leq c_3 \leq \ldots \leq c_k$

mbe-m-opt



$\lambda() = \{a_1, a_2, a_3, c_1, a_4, a_5, c_2\}$

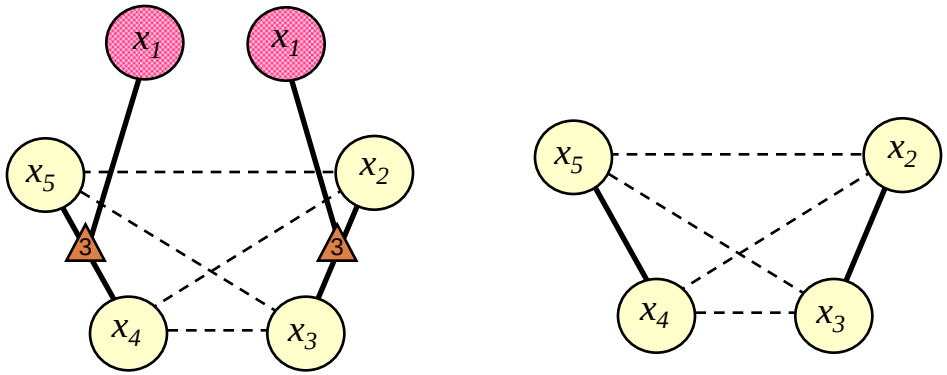
$a_1 \geq a_2 \geq a_3 \geq c_1 \geq \ldots \geq c_k$

mbe-m-opt Output

m-best WCSP = $\{c_1, c_2, c_3, \ldots, c_k\}$

$c_1 \leq c_2 \leq c_3 \leq \ldots \leq c_k$

mbe-m-opt



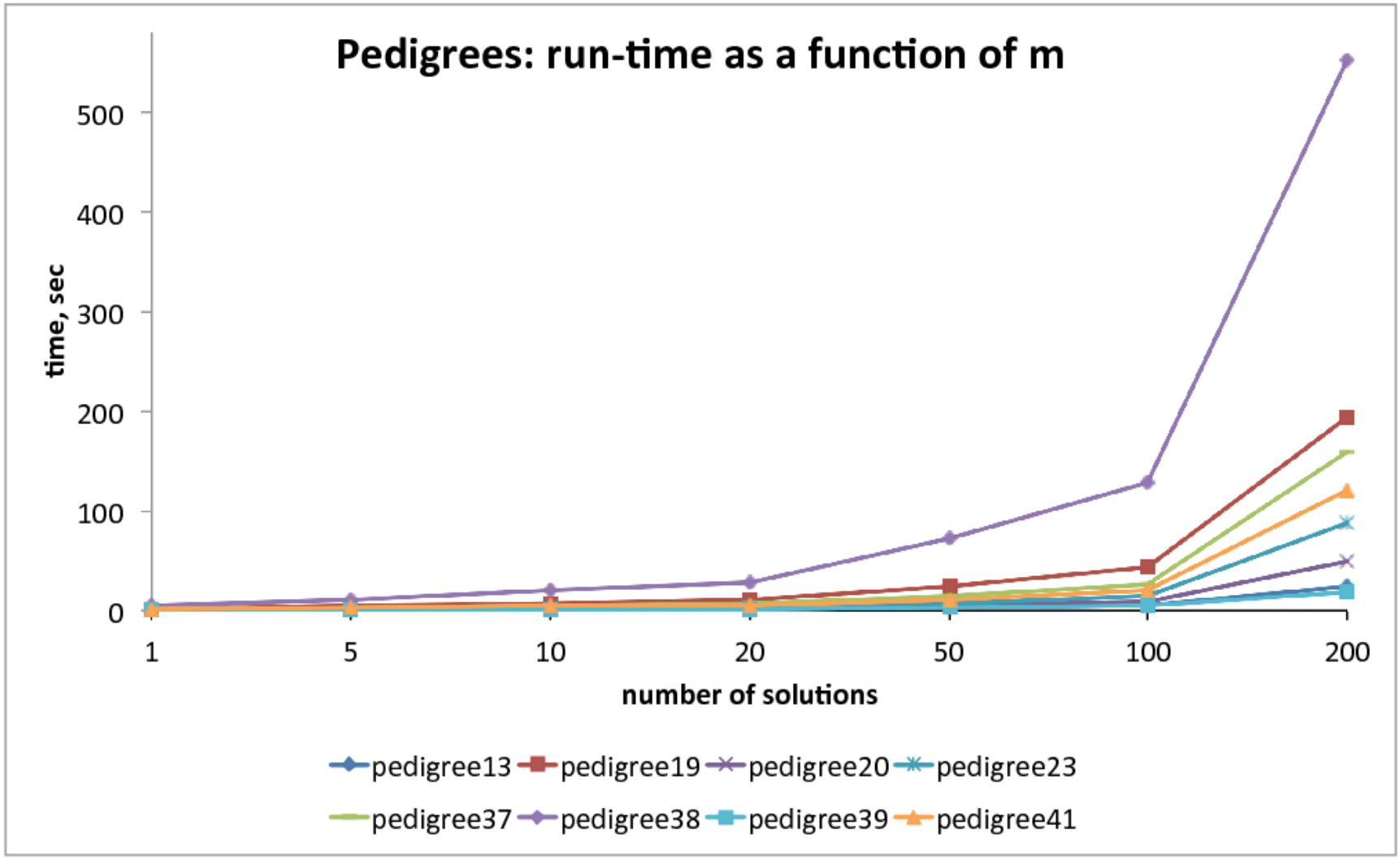
$\lambda() = \{a_1, a_2, a_3\}$

$a_1 \geq a_2 \geq a_3 \geq c_1 \geq \ldots \geq c_k$

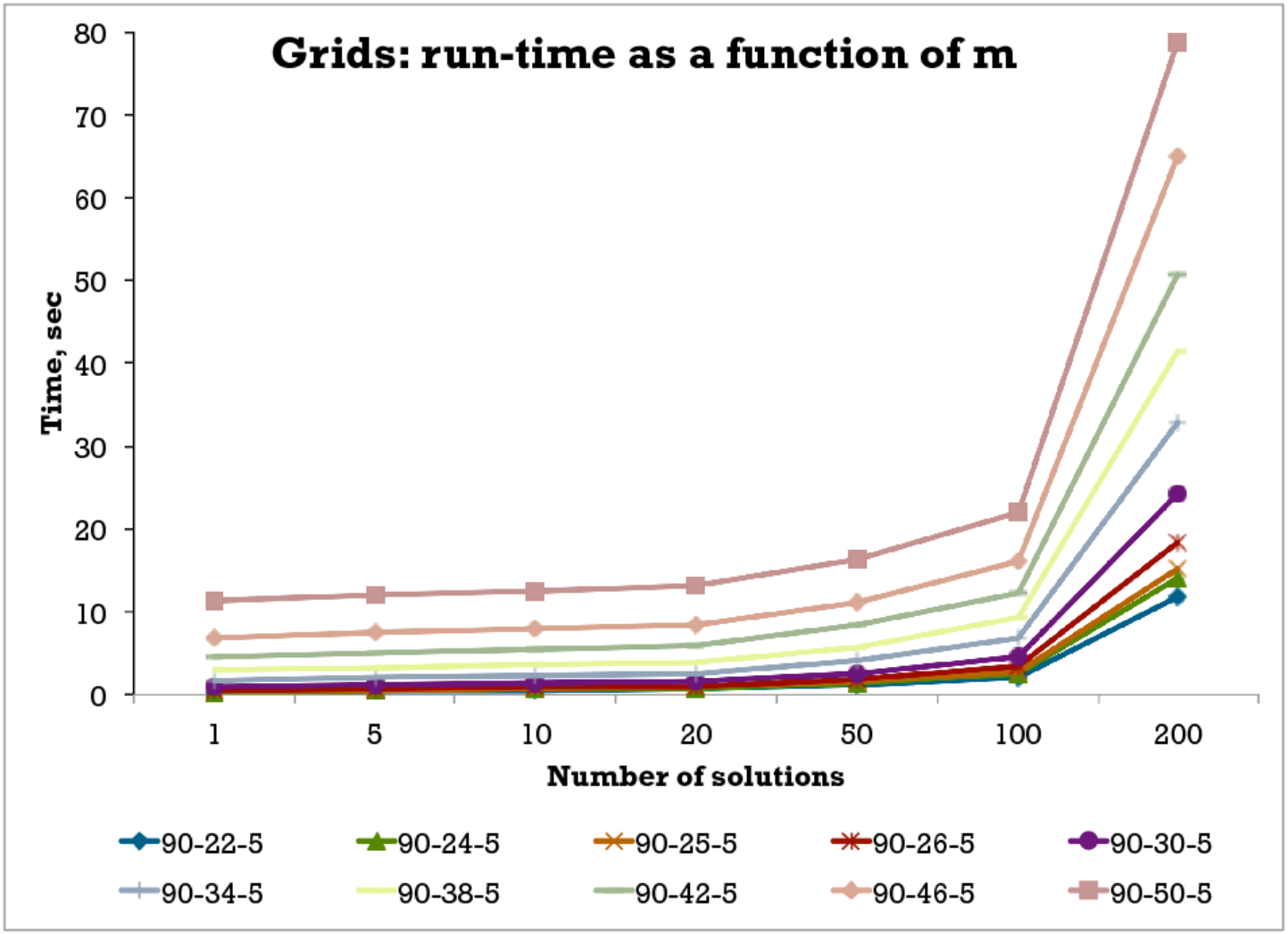
Empirical Evaluation

- Benchmarks (UAI 2008 competition):
  - Linkage analysis networks (pedigree)
  - Grid networks
  - WCSPs
- Algorithm: mbe-m-opt($z = 10$)
- Evaluate:
  - The runtime as a function of the number of solutions m
  - The improvement of the 1-bound as a function of m

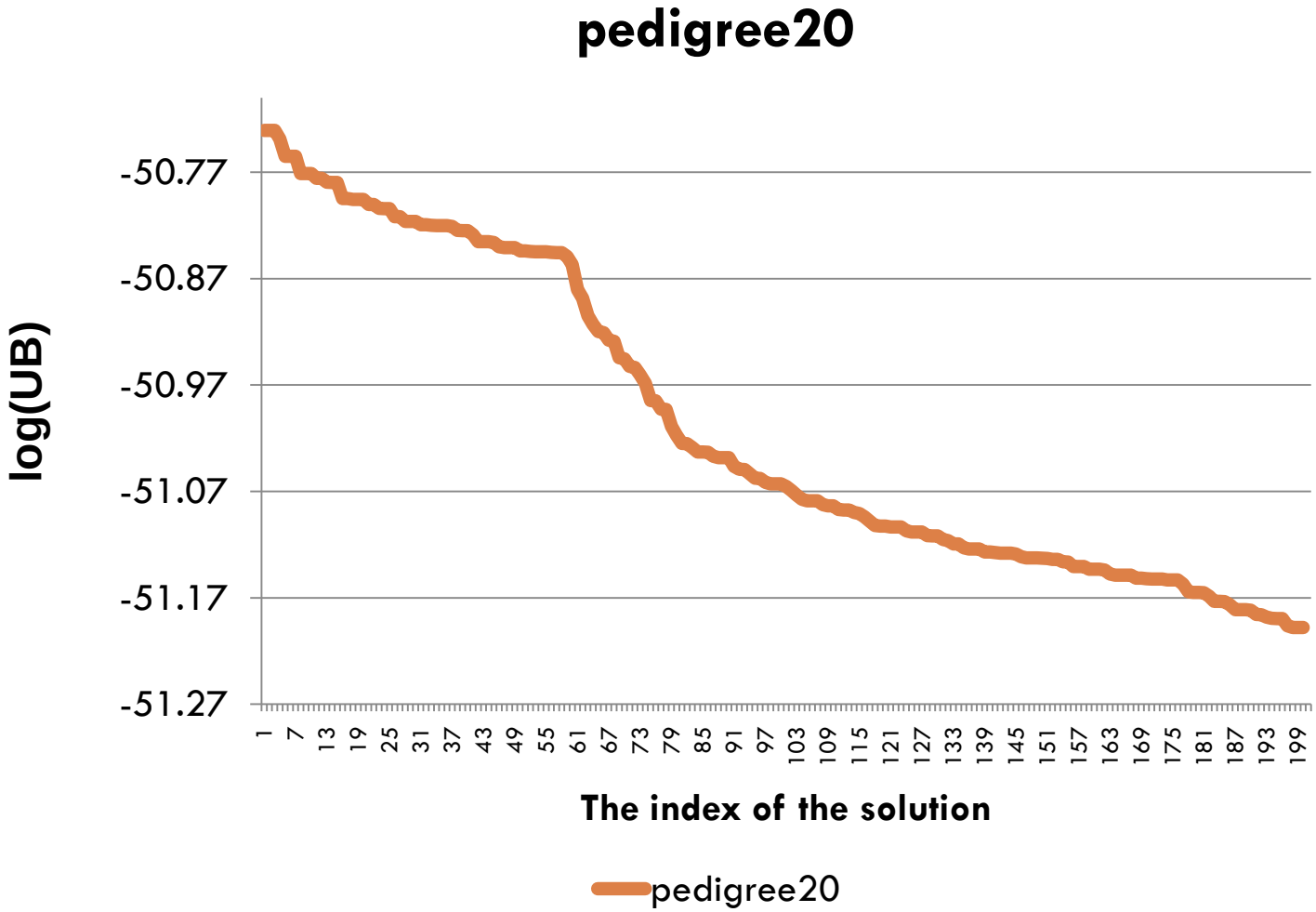
Runtime as a function of m



Runtime as a function of m

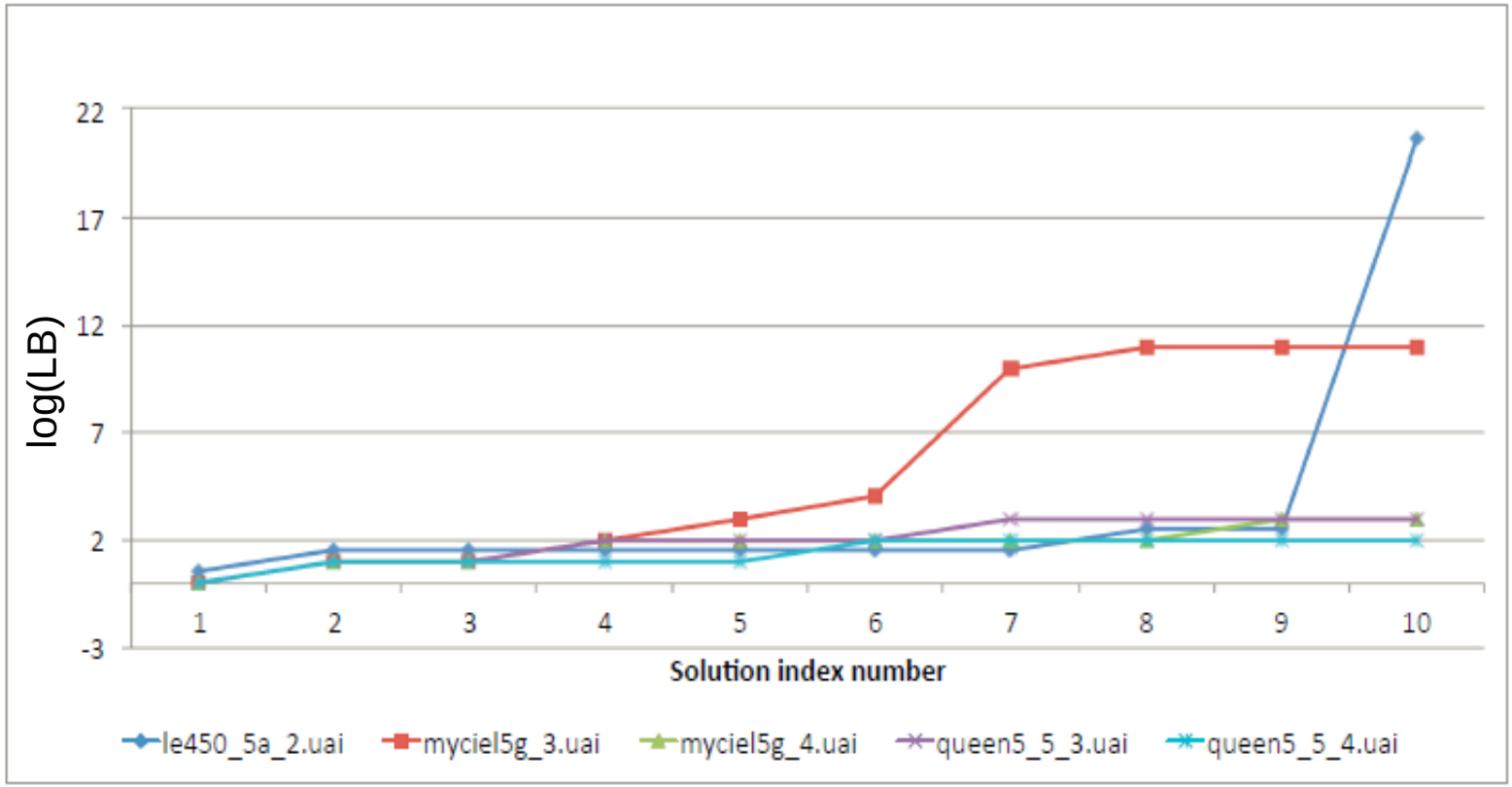


Bound's improvement as a function of *m*



Bound's improvement as a function of *m*

WCSP instances



Conlusions

- We presented two new bucket elimination algorithms for solving the m-best task by extending the combination and marginalization operators.
- The same extension yields a general formalization of m-best task over a semiring and can be used for:
  - solving other optimization problems
  - applying other known inference algorithms to m-best task.
- Future work:
  - Improve the empirical evaluation.
  - Investigate an extension of the loopy-belief propagation for the m-best task.
  - Investigate the use of bounds on m-best solutions as possible heuristics

INFERENCE SCHEMES FOR M BEST SOLUTIONS FOR SOFT CSPS

Emma Rollón, Natalia Flerova and Rina Dechter

erollon@lsi.upc.edu, flerova@ics.uci.edu, dechter@ics.uci.edu
Universitat Politècnica de Catalunya
University of California, Irvine