The Nemhauser-Trotter Reduction and Lifted Message Passing for the Weighted CSP

Hong Xu T. K. Satish Kumar Sven Koenig
hongx@usc.edu, tkskwork@gmail.com, skoenig@usc.edu

June 8, 2017

University of Southern California
the 14th International Conference on Integration of Artificial Intelligence and Operations Research Techniques in Constraint Programming (CPAIOR 2017)
Padova, Italy

Agenda

- The Weighted Constraint Satisfaction Problem (WCSP)
- The Constraint Composite Graph (CCG)
- Computational Techniques Facilitated by the CCG
  - The Nemhauser-Trotter (NT) Reduction
  - Min-Sum Message Passing (MSMP)
- Conclusion

Executive Summary

Using the Constraint Composite Graph (CCG) of a WCSP,

- The Nemhauser-Trotter (NT) Reduction, a polynomial-time procedure, can solve about 1/8 of the benchmark instances without search.
- The Min-Sum Message Passing (MSMP) algorithm, widely used in the probabilistic reasoning community, produces significantly better solutions on the CCG than on the WCSP's original form. This further bridges the probabilistic reasoning and CP communities.

Agenda

- The Weighted Constraint Satisfaction Problem (WCSP)
- The Constraint Composite Graph (CCG)
- Computational Techniques Facilitated by the CCG
    - The Nemhauser-Trotter (NT) Reduction
    - Min-Sum Message Passing (MSMP)
- Conclusion

The Weighted Constraint Satisfaction Problem: Motivation

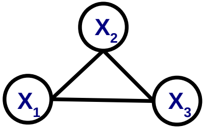
Many real-world problems can be solved using the WCSP:

- RNA motif localization (Zytnicki et al. 2008)
- Communication through noisy channels using Error Correcting Codes in Information Theory (Yedidia et al. 2003)
- Medical and mechanical diagnostics (Milho et al. 2000; Muscettola et al. 1998)
- Energy minimization in Computer Vision (Kolmogorov 2005)
- ...

Weighted Constraint Satisfaction Problem (WCSP)

- N variables $\underline{x} = \{X_1, X_2, \ldots, X_N\}$.
- Each variable X_i has a discrete-valued domain D_i.
- M weighted constraints $\{E_{s_1}, E_{s_2}, \ldots, E_{s_M}\}$.
- Each constraint E_s specifies the weight for each combination of assignments of values to a subset s of the variables.
- Find an optimal assignment of values to these variables so as to minimize the total weight: $E(\underline{x}) = \sum_{i=1}^{M} E_{s_i}(\underline{x}_{s_i})$.
- Known to be NP-hard.

WCSP Example on Boolean Variables



X_1	
0	0.7
1	0.2

X_2	
0	0.3
1	0.8

X_3	
0	0.1
1	1.0

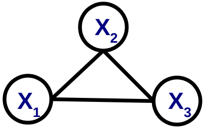
X_1 \ X_2	0	1
0	0.5	0.6
1	0.7	0.3

X_2 \ X_3	0	1
0	0.6	1.3
1	1.0	1.1

X_1 \ X_3	0	1
0	0.4	0.9
1	0.7	0.8

$$E(X_1, X_2, X_3) = E_1(X_1) + E_2(X_2) + E_3(X_3) + E_{12}(X_1, X_2) + E_{13}(X_1, X_3) + E_{23}(X_2, X_3)$$

WCSP Example: Evaluate the Assignment $X_1 = 0, X_2 = 0, X_3 = 1$



X_1	
0	0.7
1	0.2

X_2	
0	0.3
1	0.8

X_3	
0	0.1
1	1.0

X_1 \ X_2	0	1
0	0.5	0.6
1	0.7	0.3

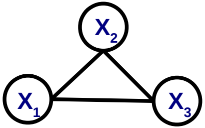
X_2 \ X_3	0	1
0	0.6	1.3
1	1.0	1.1

X_1 \ X_3	0	1
0	0.4	0.9
1	0.7	0.8

$$E(X_1 = 0, X_2 = 0, X_3 = 1) = 0.7 + 0.3 + 1.0 + 0.5 + 1.3 + 0.9 = 4.7$$

(This is not an optimal solution.)

WCSP Example: Evaluate the Assignment $X_1 = 1, X_2 = 0, X_3 = 0$



X_1	
0	0.7
1	0.2

X_2	
0	0.3
1	0.8

X_3	
0	0.1
1	1.0

X_1 \ X_2	0	1
0	0.5	0.6
1	0.7	0.3

X_2 \ X_3	0	1
0	0.6	1.3
1	1.0	1.1

X_1 \ X_3	0	1
0	0.4	0.9
1	0.7	0.8

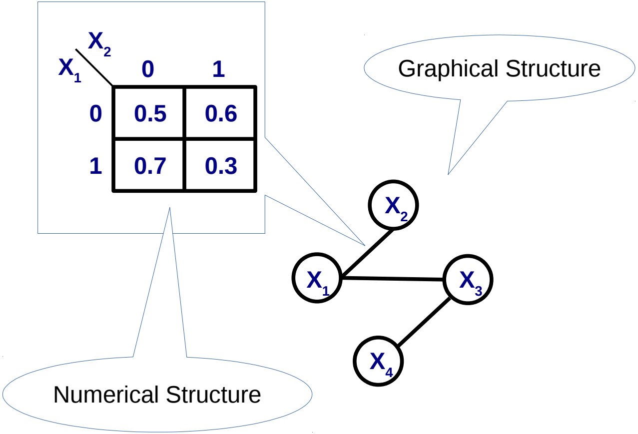
$$E(X_1 = 1, X_2 = 0, X_3 = 0) = 0.2 + 0.3 + 0.1 + 0.7 + 0.6 + 0.7 = 2.6$$

This is an optimal solution. Using brute force, it requires exponential time to find.

Agenda

- The Weighted Constraint Satisfaction Problem (WCSP)
- The Constraint Composite Graph (CCG)
- Computational Techniques Facilitated by the CCG
  - The Nemhauser-Trotter (NT) Reduction
  - Min-Sum Message Passing (MSMP)
- Conclusion

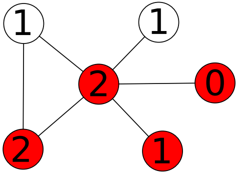
Two Forms of Structure in WCSP



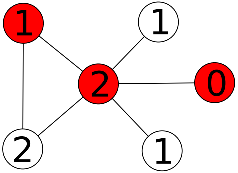
- Graphical: Which variables are in which constraints?
- Numerical: How does each constraint relate the variables in it?

How can we exploit both forms of structure computationally?

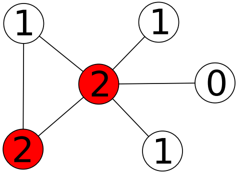
Minimum Weighted Vertex Cover (MWVC)



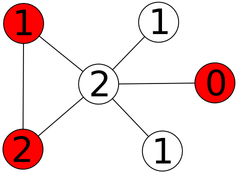
(a) ✗



(b) ✓



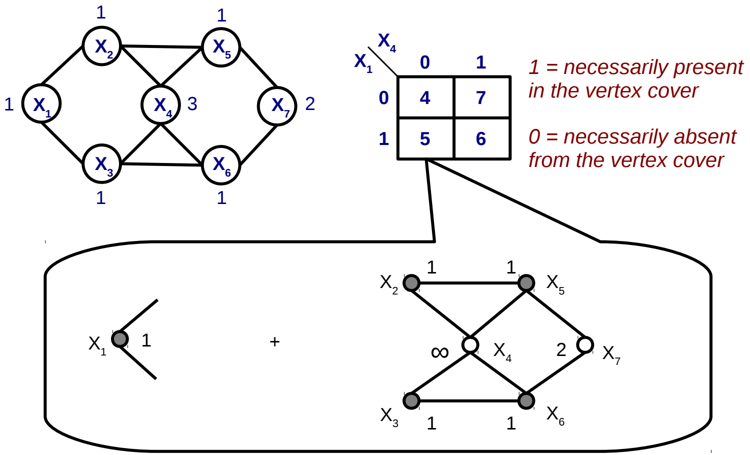
(c) ✗



(d) ✗

Each vertex is associated with a non-negative weight. Sum of the weights on the vertices in the VC is minimized.

Projection of Minimum Weighted Vertex Cover onto an Independent Set



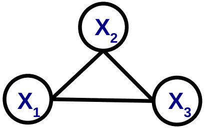
(Kumar 2008, Fig. 2)

Projection of MWVC onto an Independent Set

Assuming Boolean variables in WCSPs

- Observation: The projection of MWVC onto an independent set looks similar to a weighted constraint.
- Question 1: Can we build the lifted graphical representation for any given weighted constraint? This is answered by (Kumar 2008).
- Question 2: What is the benefit of doing so?

Lifted Representations: Example



X_1	
0	0.7
1	0.2

X_2	
0	0.3
1	0.8

X_3	
0	0.1
1	1.0

X_1 \ X_2	0	1
0	0.5	0.6
1	0.7	0.3

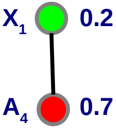
X_2 \ X_3	0	1
0	0.6	1.3
1	1.0	1.1

X_1 \ X_3	0	1
0	0.4	0.9
1	0.7	0.8

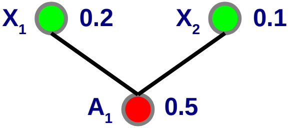
$$E(X_1, X_2, X_3) = E_1(X_1) + E_2(X_2) + E_3(X_3) + E_{12}(X_1, X_2) + E_{13}(X_1, X_3) + E_{23}(X_2, X_3)$$

Lifted Representations: Example

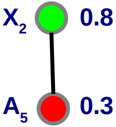
X_1	
0	0.7
1	0.2



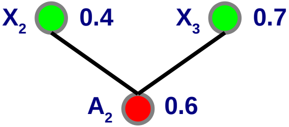
X_1 \ X_2	0	1
0	0.5	0.6
1	0.7	0.3



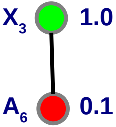
X_2	
0	0.3
1	0.8



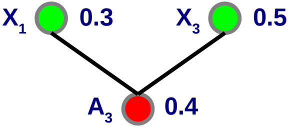
X_2 \ X_3	0	1
0	0.6	1.3
1	1.0	1.1



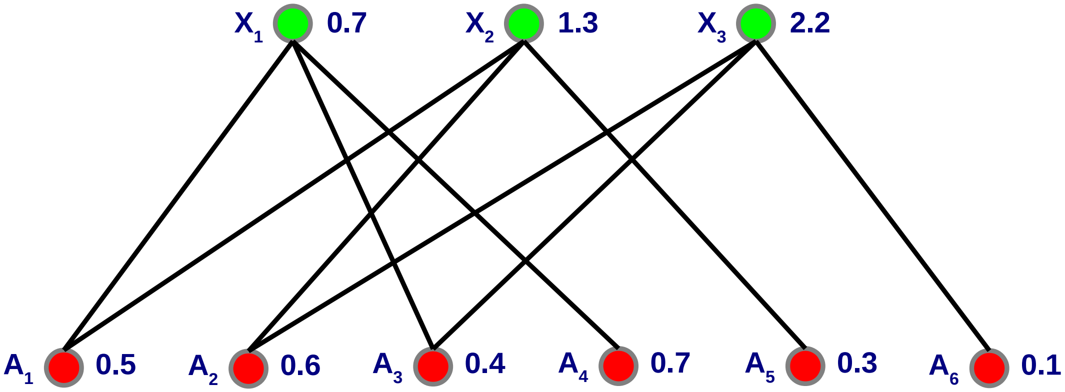
X_3	
0	0.1
1	1.0



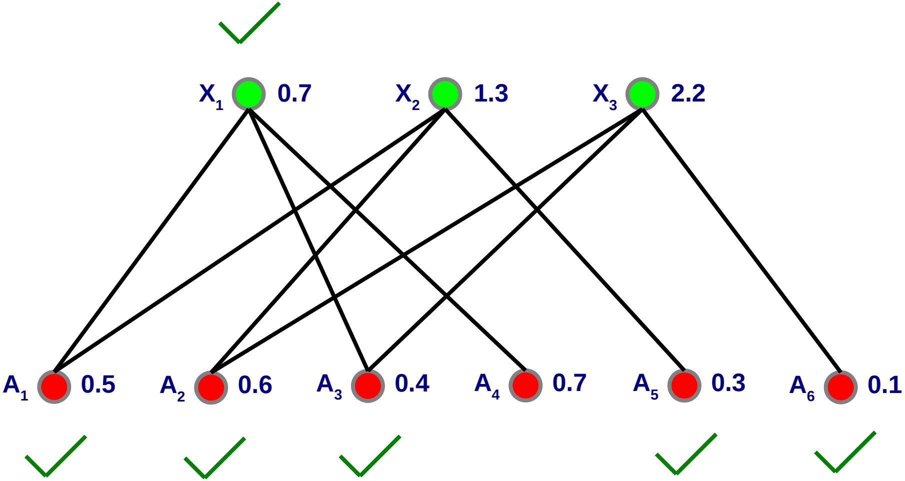
X_1 \ X_3	0	1
0	0.4	0.9
1	0.7	0.8



Constraint Composite Graph (CCG)



MWVC on the Constraint Composite Graph (CCG)



An MWVC of the CCG encodes an optimal solution of the original WCSP!

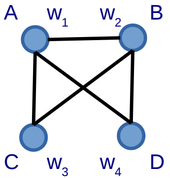
Agenda

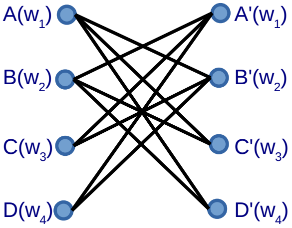
- The Weighted Constraint Satisfaction Problem (WCSP)
- The Constraint Composite Graph (CCG)
- Computational Techniques Facilitated by the CCG
  - The Nemhauser-Trotter (NT) Reduction
  - Min-Sum Message Passing (MSMP)
- Conclusion

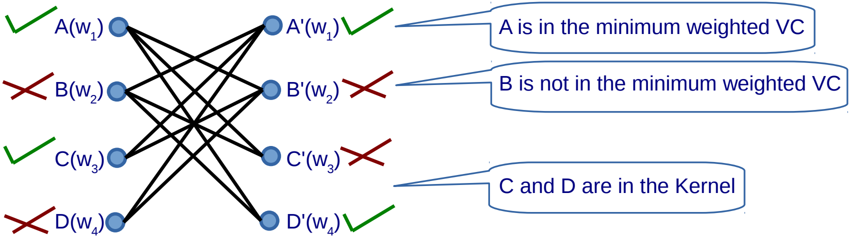
Agenda

- The Weighted Constraint Satisfaction Problem (WCSP)
- The Constraint Composite Graph (CCG)
- Computational Techniques Facilitated by the CCG
    - The Nemhauser-Trotter (NT) Reduction
    - Min-Sum Message Passing (MSMP)
- Conclusion

The Nemhauser-Trotter (NT) Reduction



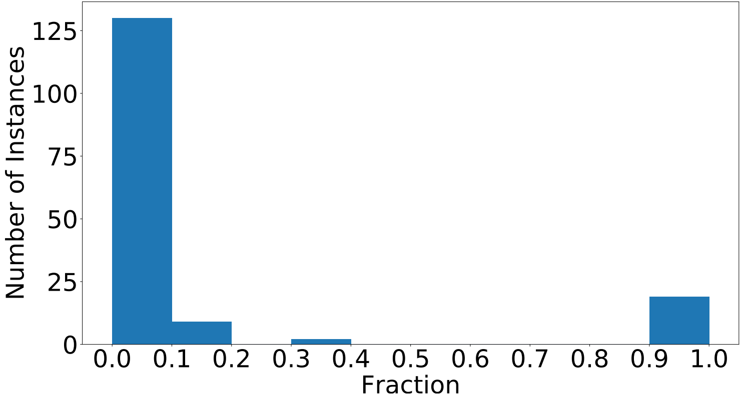




Experimental Evaluation: Instances

- The UAI 2014 Inference Competition: PR and MMAP benchmark instances (Up to 10 thousands variables and constraints)
  - Converted to WCSP instances by taking negative logarithms normalization.
- WCSP Instances from (Hurley et al. 2016) (Up to less than 1 million variables and millions of constraints)
  - The Probabilistic Inference Challenge 2011
  - The Computer Vision and Pattern Recognition OpenGM2 benchmark
  - The Weighted Partial MaxSAT Evaluation 2013
  - The MaxCSP 2008 Competition
  - The MiniZinc Challenge 2012 & 2013
  - The CFLib (a library of cost function networks)
- Only instances in which variables have only binary domains are used.

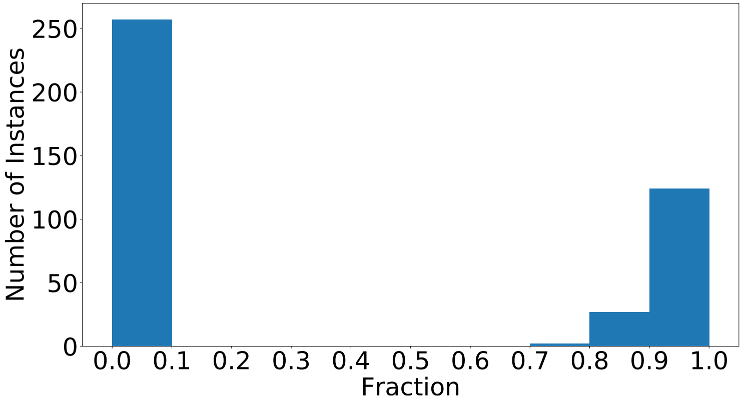
Experimental Evaluation: Results



Benchmark instances from UAI 2014 Inference Competition:

19 out of 160 benchmark instances solved by the NT reduction

Experimental Evaluation: Results



Benchmark instances from (Hurley et al. 2016):

53 out of 410 benchmark instances solved by the NT reduction

Agenda

- The Weighted Constraint Satisfaction Problem (WCSP)
- The Constraint Composite Graph (CCG)
- Computational Techniques Facilitated by the CCG
  - The Nemhauser-Trotter (NT) Reduction
  - Min-Sum Message Passing (MSMP)
- Conclusion

Min-Sum Message Passing (MSMP) Algorithms

- Min-Sum Message Passing Algorithms
  - are variants of belief propagation
  - are widely used
  - have information passed locally between variables and constraints
- Original MSMP Algorithm
  - Perform MSMP on WCSPs directly
  - Messages are passed between variables and constraints
- Lifted MSMP Algorithm
  - Perform MSMP on the MWVC problem instance of the CCG
  - Messages are passed between adjacent vertices

Operations on Tables: Min

$\min_{X_1}$ {

X_1 \ X_2	0	1
0	1	2
1	4	3

} =

X_1	
0	1
1	3

Operations on Tables: Sum

X_1 \ X_2	0	1
0	1	2
1	4	3

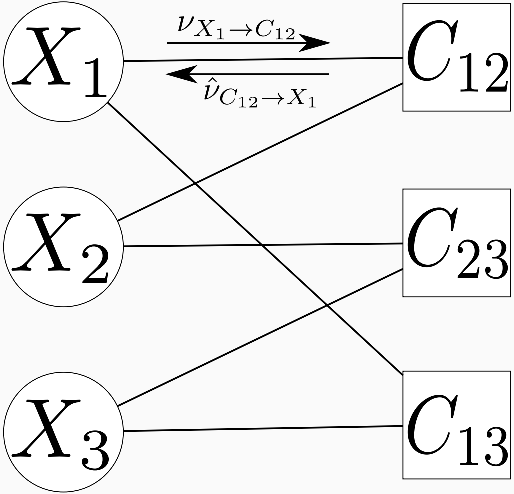
$+$

X_1	
0	5
1	6

$=$

X_1 \ X_2	0	1
0	$1 + 5 = 6$	$2 + 5 = 7$
1	$4 + 6 = 10$	$3 + 6 = 9$

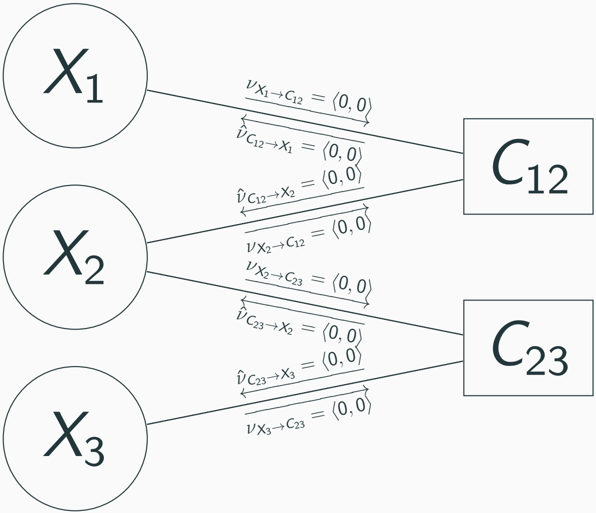
Original MSMP Algorithm: Message Passing for the WCSP



(Xu et al. 2017, Fig. 1)

- A message is a table over the single variable, which is the sender or the receiver.
- A vertex of k neighbors
  1. applies sum on the messages from its $k-1$ neighbors and internal constraint table, and
  2. applies min on the summation result and sends the resulting table to its k^{th} neighbor.

Original MSMP Algorithm: Example



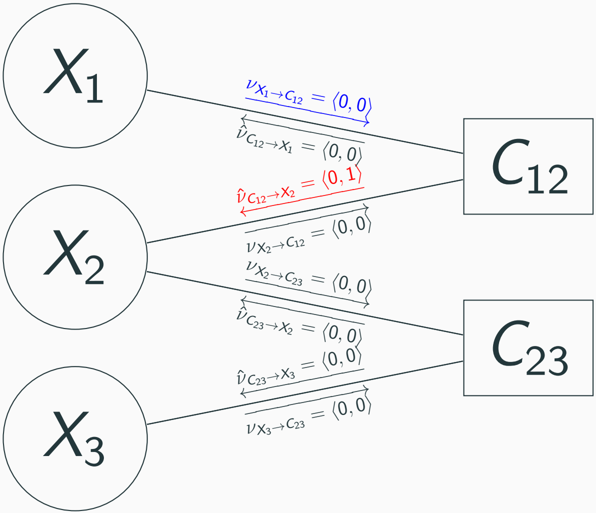
X_1 \ X_2	0	1
0	2	3
1	1	2

(a) C_{12}

X_2 \ X_3	0	1
0	1	4
1	2	2

(b) C_{23}

Original MSMP Algorithm: Example



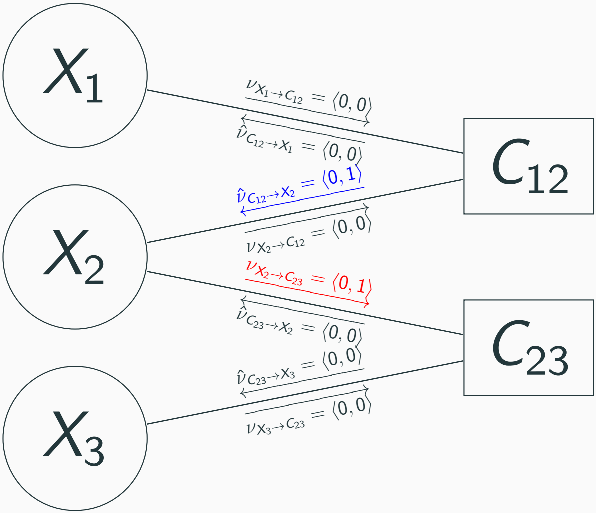
X_1 \ X_2	0	1
0	2	3
1	1	2

(a) C_{12}

X_2 \ X_3	0	1
0	1	4
1	2	2

(b) C_{23}

Original MSMP Algorithm: Example



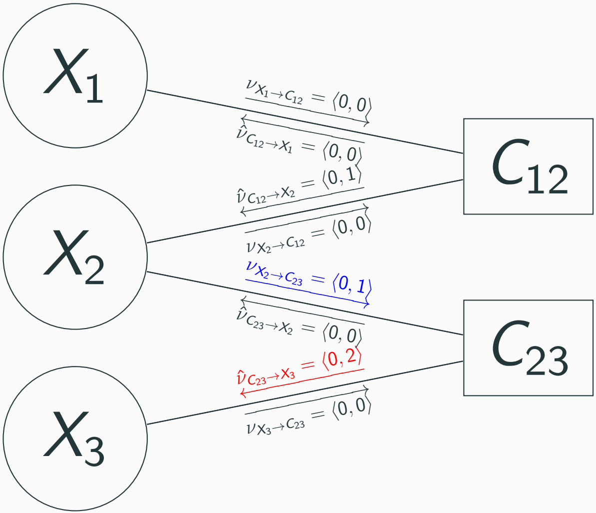
X_1 \ X_2	0	1
0	2	3
1	1	2

(a) C_{12}

X_2 \ X_3	0	1
0	1	4
1	2	2

(b) C_{23}

Original MSMP Algorithm: Example



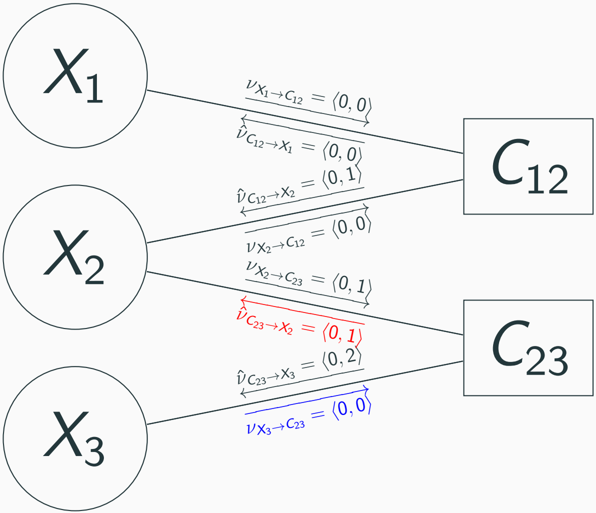
X_1 \ X_2	0	1
0	2	3
1	1	2

(a) C_{12}

X_2 \ X_3	0	1
0	1	4
1	2	2

(b) C_{23}

Original MSMP Algorithm: Example



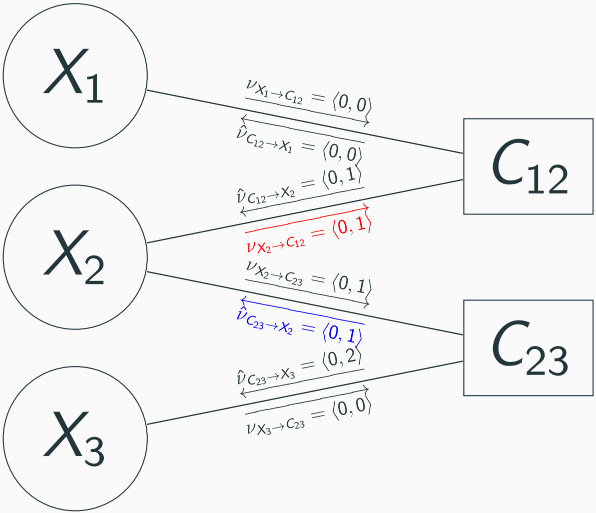
X_1 \ X_2	0	1
0	2	3
1	1	2

(a) C_{12}

X_2 \ X_3	0	1
0	1	4
1	2	2

(b) C_{23}

Original MSMP Algorithm: Example



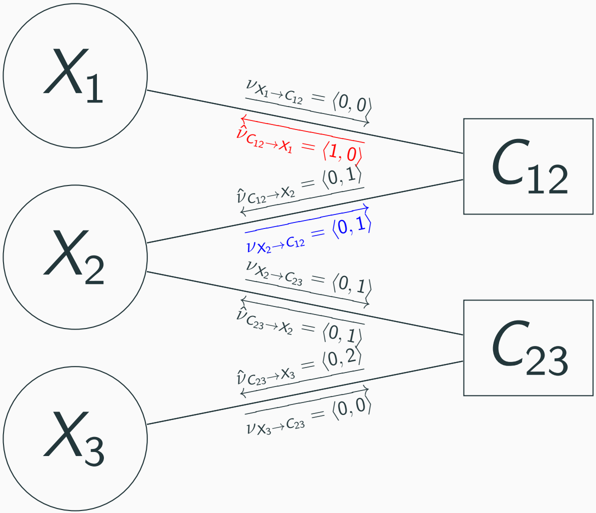
X_1 \ X_2	0	1
0	2	3
1	1	2

(a) C_{12}

X_2 \ X_3	0	1
0	1	4
1	2	2

(b) C_{23}

Original MSMP Algorithm: Example



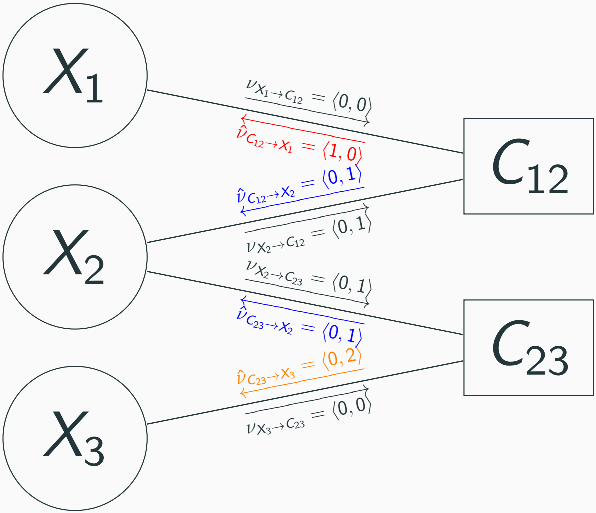
X_1 \ X_2	0	1
0	2	3
1	1	2

(a) C_{12}

X_2 \ X_3	0	1
0	1	4
1	2	2

(b) C_{23}

Original MSMP Algorithm: Example



- $X_1 = 1$ minimizes $\hat{\nu}_{C_{12}\to X_1}(X_1)$
- $X_2 = 0$ minimizes $\hat{\nu}_{C_{12}\to X_2}(X_2) + \hat{\nu}_{C_{23}\to X_2}(X_2)$
- $X_3 = 0$ minimizes $\hat{\nu}_{C_{23}\to X_3}(X_3)$
- Optimal solution: $X_1 = 1, X_2 = 0, X_3 = 0$

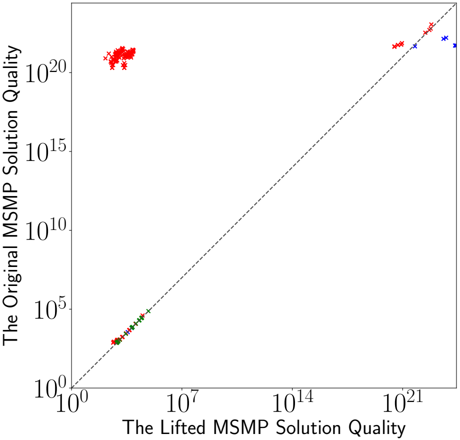
Lifted MSMP Algorithm: Finding an MWVC on the CCG

- Treat MWVC problems on the CCG as WCSPs and apply the MSMP algorithm on it.
- Messages are simplified passed between adjacent vertices.

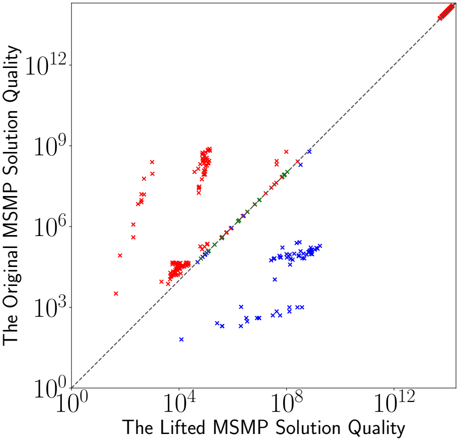
Experimental Evaluation: Setup

- Use the same benchmark instances as before.
- Solutions are reported if the MSMP algorithms do not terminate in 5 min.
- Optimal solutions are computed using toulbar2 (Hurley et al. 2016) or integer linear programming.
- Experiments were performed on a GNU/Linux workstation with an Intel Xeon processor E3-1240 v3 (8MB Cache, 3.4GHz) and 16GB RAM.

Experimental Evaluation: Results — Solution Quality

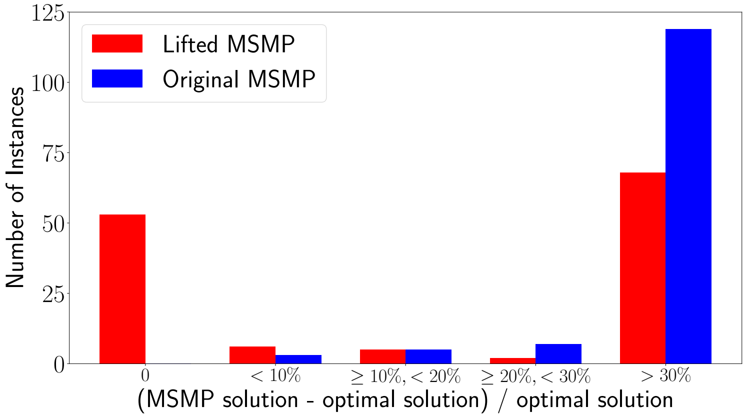


(a) Benchmark instances from the UAI 2014 Inference Competition: 126/9/18 above/below/close to the diagonal dashed line



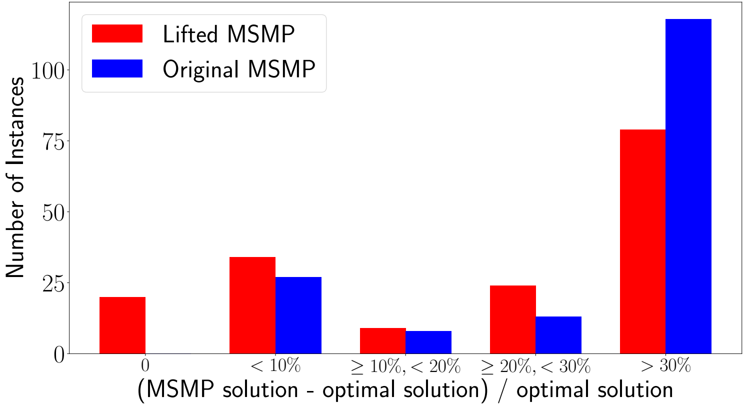
(b) Benchmark instances from (Hurley et al. 2016): 222/68/19 above/below/close to the diagonal dashed line

Experimental Evaluation: Results — Solution Quality



UAI 2014 Inference Competition: Compare qualities of solution with the optimal solutions.

Experimental Evaluation: Results — Solution Quality



Benmark instances from (Hurley et al. 2016): Compare qualities of solution with the optimal solutions.

Experimental Evaluation: Results — Convergence

Benchmark Instance Set	Neither	Both	Original	Lifted
UAI 2014 Inference Competition	25	4	124	0
(Hurley et al. 2016)	258	7	44	0

(Xu et al. 2017, Tab. 1)

- **Neither**: Neither of the MSMP algorithms terminates in 5 min.
- **Both**: Both of the MSMP algorithms terminate in 5 min.
- **Original**: Only the original MSMP algorithm terminates in 5 min.
- **Lifted**: Only the lifted MSMP algorithm terminates in 5 min.

Agenda

- The Weighted Constraint Satisfaction Problem (WCSP)
- The Constraint Composite Graph (CCG)
- Computational Techniques Facilitated by the CCG
  - The Nemhauser-Trotter (NT) Reduction
  - Min-Sum Message Passing (MSMP)
- Conclusion

Conclusion

- NT reduction on the CCG is effective for many benchmark instances.
  - The NT reduction could determine the optimal values of all variables for about 1/8 of the benchmark instances without search.
- We revived the MSMP algorithm for solving the WCSP by applying it on its CCG instead of its original form.
  - The lifted MSMP algorithm produced solutions that are significantly better than the original MSMP algorithm in general.
  - The lifted MSMP algorithm produced solutions that are close to optimal for a large fraction of benchmark instances.
  - However, the lifted MSMP algorithm is less advantageous in terms of convergence.
  - (Future work) Both MSMP algorithms can be easily adjusted to distributed settings.

References I

Barry Hurley, Barry O'Sullivan, David Allouche, George Katsirelos, Thomas Schiex, Matthias Zytnicki, and Simon de Givry. "Multi-language evaluation of exact solvers in graphical model discrete optimization". In: *Constraints* 21.3 (2016), pp. 413–434.

Vladimir Kolmogorov. *Primal-dual Algorithm for Convex Markov Random Fields.* Tech. rep. MSR-TR-2005-117. Microsoft Research, 2005.

T. K. Satish Kumar. "A framework for hybrid tractability results in boolean weighted constraint satisfaction problems". In: *the International Conference on Principles and Practice of Constraint Programming*. Springer, 2008, pp. 282–297.

Isabel Milho, Ana Fred, Jorge Albano, Nuno Baptista, and Paulo Sena. "An Auxiliary System for Medical Diagnosis Based on Bayesian Belief Networks". In: *Portuguese Conference on Pattern Recognition*. 2000.

Nicola Muscettola, P. Pandurang Nayak, Barney Pell, and Brian C. Williams. "Remote Agent: to boldly go where no {AI} system has gone before". In: *Artificial Intelligence* 103.1–2 (1998), pp. 5–47.

References II

Hong Xu, T. K. Satish Kumar, and Sven Koenig. "The Nemhauser-Trotter Reduction and Lifted Message Passing for the Weighted CSP". In: *the 14th International Conference on Integration of Artificial Intelligence and Operations Research Techniques in Constraint Programming (CPAIOR)*. 2017.

Jonathan S Yedidia, William T Freeman, and Yair Weiss. "Understanding belief propagation and its generalizations". In: *Exploring Artificial Intelligence in the New Millennium* 8 (2003), pp. 236–239.

Matthias Zytnicki, Christine Gaspin, and Thomas Schiex. "DARN! A Weighted Constraint Solver for RNA Motif Localization". In: *Constraints* 13.1 (2008), pp. 91–109.