Solving Optimization Problems via Maximum Satisfiability: Encodings and Re-Encodings

Jeremias Berg

ACP Doctoral Research Award @ CP2020

September 11, 2020

HELSINKI INSTITUTE FOR INFORMATION TECHNOLOGY

Who am I?

www.jeremiasberg.com

- MSc. in mathematics.
- Graduated with distinction from the University of Helsinki in June 2018.
  - Supervisor: **Prof. Matti Järvisalo**.
  - PhD Committee:
    - **Prof. Peter Stuckey**,
    - **Prof. Lakhdar Sais** and
    - **Prof. Inês Lynce** as the opponent.

Thanks to Matti Järvisalo and Ruben Martins for providing material for these slides

My Doctoral Research

Developing Maximum Satisfiability (MaxSAT) for solving real-world optimization problems.

More specifically:

- Understanding and developing solver independent preprocessing.
- Applying MaxSAT toward NP-hard data analysis problems.

My Doctoral Research

Developing Maximum Satisfiability (MaxSAT) for solving real-world optimization problems.

More specifically:

- Understanding and developing solver independent preprocessing.
- Applying MaxSAT toward NP-hard data analysis problems.

Outline

The talk is divided into two parts:

1. Maximum Satisfiability in real world optimization.
2. Main results of doctoral research.

Maximum Satisfiability

MaxSAT

Exact Boolean optimization paradigm:

- Based on propositional logic
  - i.e. weighted Boolean combinations of binary variables.
- Builds on the success of SAT.
- Expanding range of real-world applications.
- Effective and often orthogonal alternative to e.g.
  - mixed integer programming
  - finite-domain constraint optimization
  - and others

Maximum Satisfiability

MaxSAT

Exact Boolean optimization paradigm:

- Based on propositional logic
  - i.e. weighted Boolean combinations of binary variables.
- Builds on the success of SAT.
- Expanding range of real-world applications.
- Effective and often orthogonal alternative to e.g.
  - mixed integer programming
  - finite-domain constraint optimization
  - and others

Maximum Satisfiability

MaxSAT

Exact Boolean optimization paradigm:

- Based on propositional logic
  - i.e. weighted Boolean combinations of binary variables.
- Builds on the success of SAT.
- Expanding range of real-world applications.
- Effective and often orthogonal alternative to e.g.
  - mixed integer programming
  - finite-domain constraint optimization
  - and others

Maximum Satisfiability

MaxSAT

Exact Boolean optimization paradigm:

- Based on propositional logic
  - i.e. weighted Boolean combinations of binary variables.
- Builds on the success of SAT.
- Expanding range of real-world applications.
- Effective and often orthogonal alternative to e.g.
  - mixed integer programming
  - finite-domain constraint optimization
  - and others

Applications of MaxSAT

Drastically increasing number of successful applications

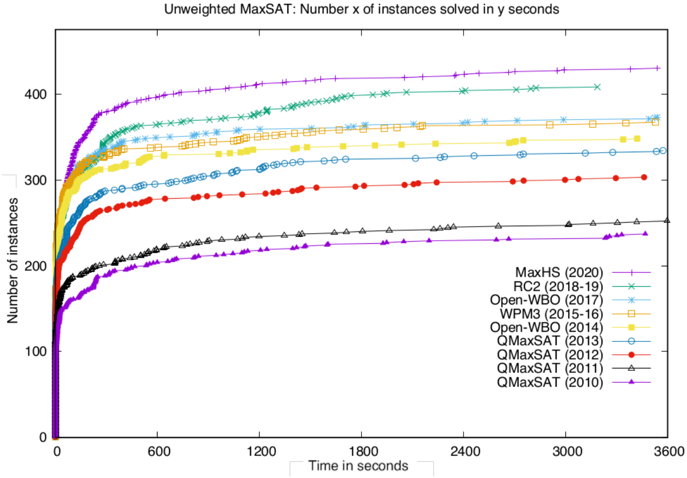
Applications of MaxSAT

Drastically increasing number of successful applications

Central to the increasing success:
Advances in solver technology

Evolution of MaxSAT solvers

Plot provided by Ruben Martins



83% more instances solved in 10 years

MaxSAT solving

Focus of the thesis (and these applications):
SAT-based MaxSAT algorithms

MaxSAT solving

Focus of the thesis (and these applications):
SAT-based MaxSAT algorithms

In more detail:

- Algorithms that make iterative use of SAT solvers.
- Three central ones:
  - *model-improving*,
  - *core-guided* and
  - *implicit hitting set* based.
- Central results of thesis applicable to all.

Results of Thesis Research

Thesis

Six core publications aiming to improve MaxSAT-based solution methods to combinatorial optimization problems.

- Part 1: MaxSAT solving techniques, specifically preprocessing
  IJCAI 2015, ICTAI 2015, CP 2016, ECAI 2016
- Part 2: MaxSAT based approaches to solving data analysis problems
  AISTATS2014, AIJ 2015

Thesis

Six core publications aiming to improve MaxSAT-based solution methods to combinatorial optimization problems.

- Part 1: MaxSAT solving techniques, specifically preprocessing
  IJCAI 2015, ICTAI 2015, CP 2016, ECAI 2016
- Part 2: MaxSAT based approaches to solving data analysis problems
  AISTATS2014, AIJ 2015

Thesis

Six core publications aiming to improve MaxSAT-based solution methods to combinatorial optimization problems.

- Part 1: MaxSAT solving techniques, specifically preprocessing
  IJCAI 2015, ICTAI 2015, CP 2016, ECAI 2016
- Part 2: MaxSAT based approaches to solving data analysis problems
  AISTATS2014, AIJ 2015

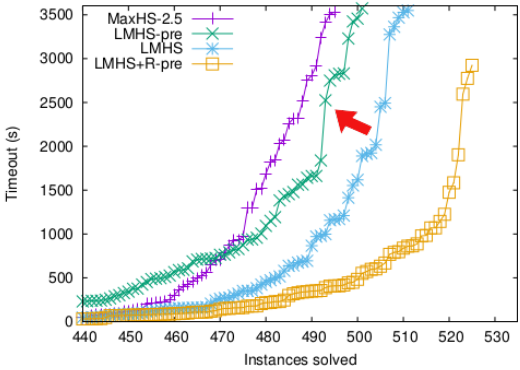
Part 1: MaxSAT Preprocessing

Main results

1. Integrating preprocessing and solving requires care.
   - Straightforward way degrades performance.
   - Our more careful way improves performance.
2. Liftings of SAT preprocessing rules have limited effect on MaxSAT algorithms.
   - Inability of affecting small unsatisfiable subsets of constraints is a fundamental limitation.

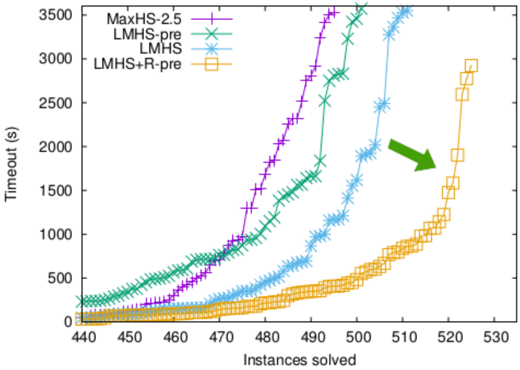
Main results

1. Integrating preprocessing and solving requires care.
   - Straightforward way degrades performance.
   - Our more careful way improves performance.
2. Liftings of SAT preprocessing rules have limited effect on MaxSAT algorithms.
   - Inabil … onstraints is a funda…



Main results

1. Integrating preprocessing and solving requires care.
   - Straightforward way degrades performance.
   - Our more careful way improves performance.
2. Liftings of SAT preprocessing rules have limited effect on MaxSAT algorithms.



Main results

1. Integrating preprocessing and solving requires care.
   - Straightforward way degrades performance.
   - Our more careful way improves performance.
2. Liftings of SAT preprocessing rules have limited effect on MaxSAT algorithms.
   - Inability of affecting small unsatisfiable subsets of constraints is a fundamental limitation.

Main results

1. Integrating preprocessing and solving requires care.
   - Straightforward way degrades performance.
   - Our more careful way improves performance.
2. Liftings of SAT preprocessing rules have limited effect on MaxSAT algorithms.
   - Inability of affecting small unsatisfiable subsets of constraints is a fundamental limitation.

Subsequently

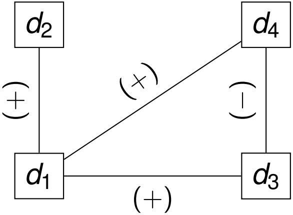
- Increasing number of solvers in MSEs use preprocessing.
- Improved theoretical understanding of MaxSAT preprocessing.

[Berg and Järvisalo(2019)]

Part 2: MaxSAT for Data Analysis

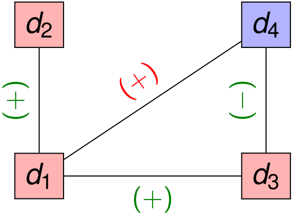
Correlation Clustering

- A set of datapoints to be clustered.
  - Pairwise similarity measure over the points.
- Want a clustering that correlates well with the measure.
  - Similar points $\rightarrow$ same cluster.
  - Dissimilar points $\rightarrow$ different clusters.
  - **Note:** No restriction on number of clusters.



Correlation Clustering

- A set of datapoints to be clustered.
    - Pairwise similarity measure over the points.
- Want a clustering that correlates well with the measure.
    - Similar points $\rightarrow$ same cluster.
    - Dissimilar points $\rightarrow$ different clusters.
    - **Note:** No restriction on number of clusters.

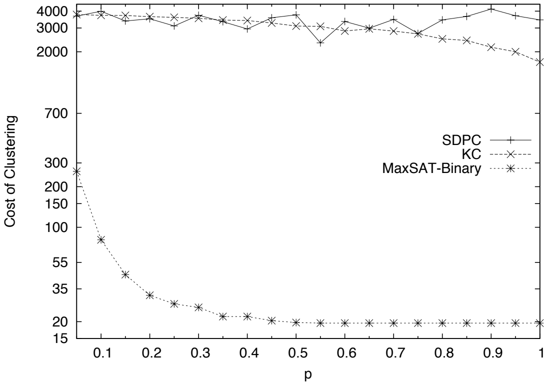


Applications in: biosciences, social network analysis, information retrieval, bioinformatics and microarray data analysis . . . .

Correlation Clustering

Main results

- Solutions obtained with MaxSAT are significantly better than the ones obtained by specialized approximative algorithms.

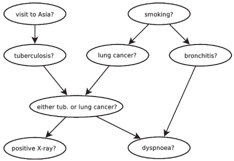


Bounded Treewidth Bayesian Network Structure Learning

BTBNSL

A	S	T	L	B	T&L	X	D
0	1	0	1	0	1	1	1
0	0	1	1	0	1	0	1
1	1	0	1	1	0	0	1
1	0	0	1	1	1	1	1
1	1	1	1	1	1	0	1
1	1	1	1	0	0	1	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

$\rightarrow$



$\rightarrow$

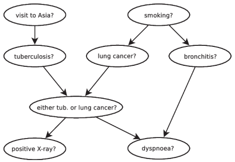
$$P(S \mid T, L, B) = ?$$
$$P(X \mid D, A, B) = ?$$
$$P(A \mid T, L, B) = ?$$
$$\vdots$$

- Starting from observations (data) learn a Bayesian Network that explains the data as well as possible.
  - NP-hard
- Restrict the treewidth of the learned structure.
  - NP-hard

Bounded Treewidth Bayesian Network Structure Learning

A	S	T	L	B	T&L	X	D
0	1	0	1	0	1	1	1
0	0	1	1	0	1	0	1
1	1	0	1	1	0	0	1
1	0	0	1	1	1	1	1
1	1	1	1	1	1	0	1
1	1	1	1	0	0	1	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

$\longrightarrow$



$\longrightarrow$

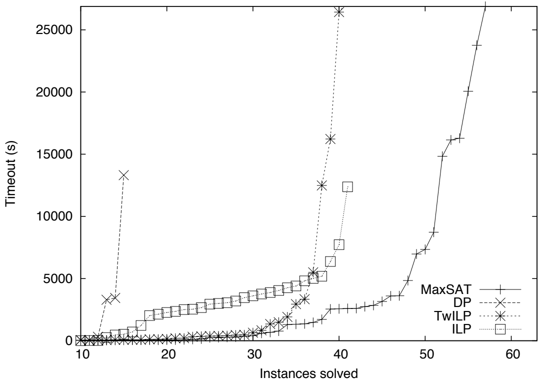
$$P(S \mid T, L, B) = ?$$
$$P(X \mid D, A, B) = ?$$
$$P(A \mid T, L, B) = ?$$
$$\vdots$$

- Starting from observations (data) learn a Bayesian Network that explains the data as well as possible.
  - NP-hard
- Restrict the treewidth of the learned structure.
  - NP-hard

BTBNSL

Main Results

- MaxSAT scales better than other declarative approaches .
- Adding the (highly non-trivial) treewidth constraint to the encoding improves overall performance.



Additional Results of Thesis Research

- Seven related papers developing MaxSAT algorithms and applications,
- Contributions to open-source software

  [Korhonen et al.(2017)Korhonen, Berg, Saikko, and Järvisalo, Saikko et al.(2016)Saikko, Berg, and Järvisalo]
- Openly available benchmark sets [Berg et al.(2018)Berg, Hyttinen, and Järvisalo]
  - Used in MSEs since 2015.

Additional Results of Thesis Research

- Seven related papers developing MaxSAT algorithms and applications,
- Contributions to open-source software
  [Korhonen et al.(2017)Korhonen, Berg, Saikko, and Järvisalo, Saikko et al.(2016)Saikko, Berg, and Järvisalo]
- Openly available benchmark sets [Berg et al.(2018)Berg, Hyttinen, and Järvisalo]
  - Used in MSEs since 2015.

Recent Highlights

- Hybrid algorithms for MaxSAT:
  - Loandra - One of best incomplete solvers in recent MSEs.
    [Berg et al.(2019)Berg, Demirovic, and Stuckey]
  - Best paper award at SAT 2020. [Berg et al.(2020)Berg, Bacchus, and Poole]

Conclusions

Take Home Message:
Modern MaxSAT has developed into an effective and competitive approach to exact optimization.

Conclusions

Take Home Message:
Modern MaxSAT has developed into an effective and competitive approach to exact optimization.

My thesis research played a small part in this.

Conclusions

Take Home Message:
Modern MaxSAT has developed into an effective and competitive approach to exact optimization.
My thesis research played a small part in this.

Future research directions:

- Better hybrid search techniques (e.g. inprocessing).
- Incremental applications of solvers to further improve scalability.
- Improved understanding of the relationship between applications and optimization paradigms.

Conclusions

Take Home Message:
Modern MaxSAT has developed into an effective and competitive approach to exact optimization.
My thesis research played a small part in this.

Future research directions:

- Better hybrid search techniques (e.g. inprocessing).
- Incremental applications of solvers to further improve scalability.
- Improved understanding of the relationship between applications and optimization paradigms.

Conclusions

Take Home Message:
Modern MaxSAT has developed into an effective and competitive approach to exact optimization.
My thesis research played a small part in this.

Future research directions:

- Better hybrid search techniques (e.g. inprocessing).
- Incremental applications of solvers to further improve scalability.
- Improved understanding of the relationship between applications and optimization paradigms.

Conclusions

Take Home Message:
Modern MaxSAT has developed into an effective and competitive approach to exact optimization.
My thesis research played a small part in this.

Future research directions:

- Better hybrid search techniques (e.g. inprocessing).
- Incremental applications of solvers to further improve scalability.
- Improved understanding of the relationship between applications and optimization paradigms.

Thank you!

I would like to thank:

- the Association for Constraint Programming,
- my supervisor Matti Järvisalo,
- the people involved in the nomination:
  - Inês Lynce,
  - Peter Stuckey,
- as well as all my co-authors and collaborators:
  - Paul Saikko,
  - Brandon Malone
- and all of you for watching.

Bibliography I

J. Berg, A. Hyttinen, and M. Järvisalo.
Applications of MaxSAT in Data Analysis.
In *POS@SAT*, volume 59 of *EPiC Series in Computing*, pages 50–64. EasyChair, 2018.

Jeremias Berg and Matti Järvisalo.
Unifying reasoning and core-guided search for maximum satisfiability.
In Francesco Calimeri, Nicola Leone, and Marco Manna, editors, *Logics in Artificial Intelligence - 16th European Conference, JELIA 2019, Rende, Italy, May 7-11, 2019, Proceedings*, volume 11468 of *Lecture Notes in Computer Science*, pages 287–303. Springer, 2019.
doi: 10.1007/978-3-030-19570-0_19.
URL https://doi.org/10.1007/978-3-030-19570-0_19.

Jeremias Berg, Emir Demirovic, and Peter J. Stuckey.
Core-boosted linear search for incomplete maxsat.
In Louis-Martin Rousseau and Kostas Stergiou, editors, *Integration of Constraint Programming, Artificial Intelligence, and Operations Research - 16th International Conference, CPAIOR 2019, Thessaloniki, Greece, June 4-7, 2019, Proceedings*, volume 11494 of *Lecture Notes in Computer Science*, pages 39–56. Springer, 2019.
doi: 10.1007/978-3-030-19212-9_3.
URL https://doi.org/10.1007/978-3-030-19212-9_3.

Jeremias Berg, Fahiem Bacchus, and Alex Poole.
Abstract cores in implicit hitting set maxsat solving.
In Luca Pulina and Martina Seidl, editors, *Theory and Applications of Satisfiability Testing - SAT 2020 - 23rd International Conference, Alghero, Italy, July 3-10, 2020, Proceedings*, volume 12178 of *Lecture Notes in Computer Science*, pages 277–294. Springer, 2020.
doi: 10.1007/978-3-030-51825-7_20.
URL https://doi.org/10.1007/978-3-030-51825-7_20.

Bibliography II

Tuukka Korhonen, Jeremias Berg, Paul Saikko, and Matti Järvisalo.
MaxPre: An extended MaxSAT preprocessor.
In Serge Gaspers and Toby Walsh, editors, *Theory and Applications of Satisfiability Testing - SAT 2017 - 20th International Conference, Melbourne, VIC, Australia, August 28 - September 1, 2017, Proceedings*, volume 10491 of *Lecture Notes in Computer Science*, pages 449–456. Springer, 2017.
doi: 10.1007/978-3-319-66263-3_28.
URL https://doi.org/10.1007/978-3-319-66263-3_28.

Paul Saikko, Jeremias Berg, and Matti Järvisalo.
LMHS: A SAT-IP hybrid MaxSAT solver.
In Nadia Creignou and Daniel Le Berre, editors, *Theory and Applications of Satisfiability Testing - SAT 2016 - 19th International Conference, Bordeaux, France, July 5-8, 2016, Proceedings*, volume 9710 of *Lecture Notes in Computer Science*, pages 539–546. Springer, 2016.