Solving Max SAT and #SAT on structured CNF formulas

Sigve Hortemo Sæther, Jan Arne Telle, Martin Vatshelle

University of Bergen

July 14, 2014

Outline

- Equivalence of assignments

Outline

- Equivalence of assignments
- CNF formulas with few equivalence classes

Outline

- Equivalence of assignments
- CNF formulas with few equivalence classes
- Algorithm for MAX SAT

Outline

- Equivalence of assignments
- CNF formulas with few equivalence classes
- Algorithm for MAX SAT
- Polynomial time solvable cases

Equivalence of assignments to CNF formulas

- When are two truth assignments to a CNF formula equivalent?

Equivalence of assignments to CNF formulas

- When are two truth assignments to a CNF formula equivalent?
- When they satisfy the same set of clauses.

Equivalence of assignments to CNF formulas

- When are two truth assignments to a CNF formula equivalent?
- When they satisfy the same set of clauses.
- The number of equivalence classes is called the **ps-value**.

Equivalence of assignments to CNF formulas

- When are two truth assignments to a CNF formula equivalent?
- When they satisfy the same set of clauses.
- The number of equivalence classes is called the **ps-value**.

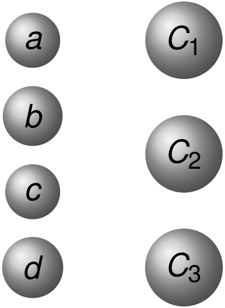
- We need two standard definitions before we give a sufficient condition for low ps-value.

Incidence graph

$$(a \vee b \vee \neg c) \wedge (\neg a \vee c \vee \neg d) \wedge (\neg b \vee d)$$

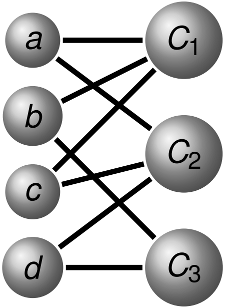
Incidence graph

$$(a \vee b \vee \neg c) \wedge (\neg a \vee c \vee \neg d) \wedge (\neg b \vee d)$$



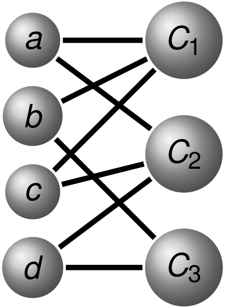
Incidence graph

$$(a \vee b \vee \neg c) \wedge (\neg a \vee c \vee \neg d) \wedge (\neg b \vee d)$$



Incidence graph

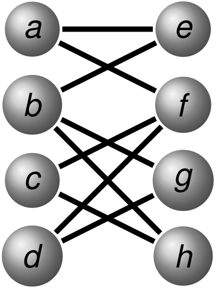
$$(a \vee b \vee \neg c) \wedge (\neg a \vee c \vee \neg d) \wedge (\neg b \vee d)$$



- For a CNF formula F we denote the incedence graph by $I(F)$.

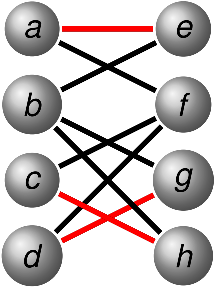
Induced matching

Given a bipartite graph M is an induced matching if:



Induced matching

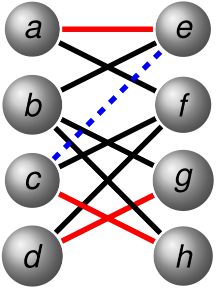
Given a bipartite graph *M* is an induced matching if:



1. *M* is a matching.

Induced matching

Given a bipartite graph M is an induced matching if:



1. M is a matching.
2. No other edge is adjacent to 2 edges in M.

CNF formulas with low ps-value

Lemma

Let F be a CNF formula and k be the maximum size of an induced matching in $I(F)$.
The ps-value of F is at most $|\texttt{cla}(F)|^k + 1$

CNF formulas with low ps-value

Lemma

Let F be a CNF formula and k be the maximum size of an induced matching in $I(F)$.
The ps-value of F is at most $|\texttt{cla}(F)|^k + 1$

Sketch of proof

1. *Let U be the set of unsatisfied clauses, and L the variables appearing in some clause of U, then there is a unique assignment for L.*
2. *There exist $U' \subseteq U$ of size at most k such that L is uniquely defined by U'.*
3. *There is at most $|\texttt{cla}(F)|^k + 1$ choices for U'*

CNF formulas with low ps-value

Any CNF formula F where $I(F)$ has a maximum induced matching of size 1 will be on the form:

CNF formulas with low ps-value

Any CNF formula F where $I(F)$ has a maximum induced matching of size 1 will be on the form:

$$(a \vee \neg b)$$

CNF formulas with low ps-value

Any CNF formula F where $I(F)$ has a maximum induced matching of size 1 will be on the form:

$$(a \vee \neg b) \wedge (\neg a \vee b \vee c)$$

CNF formulas with low ps-value

Any CNF formula F where $I(F)$ has a maximum induced matching of size 1 will be on the form:

$$(a \vee \neg b) \wedge (\neg a \vee b \vee c) \wedge (a \vee b \vee c \vee d)$$

CNF formulas with low ps-value

Any CNF formula F where $I(F)$ has a maximum induced matching of size 1 will be on the form:

$$(a \vee \neg b) \wedge (\neg a \vee b \vee c) \wedge (a \vee b \vee c \vee d) \ \ldots$$

DP algorithm strategy

1. Make an ordering of clauses and variables.

DP algorithm strategy

1. Make an ordering of clauses and variables.

   Let the i first elements of the ordering define a cut S_i.
2. Store one solution for each equivalence class of the cut S_i.

DP algorithm strategy

1. Make an ordering of clauses and variables.

   Let the i first elements of the ordering define a cut S_i.
2. Store one solution for each equivalence class of the cut S_i.
3. Use solutions for S_i to find solutions of S_{i+1}.

DP algorithm strategy

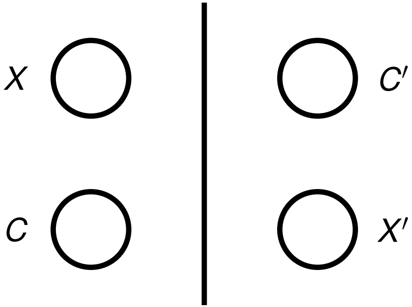
1. Make an ordering of clauses and variables.

   Let the i first elements of the ordering define a cut S_i.
2. Store one solution for each equivalence class of the cut S_i.
3. Use solutions for S_i to find solutions of S_{i+1}.

Note: the linear ordering can be replaced by a tree-like decomposition.

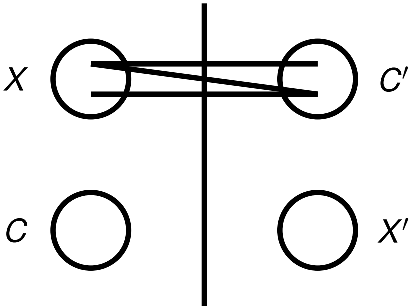
Equivalence over a cut

Let F be a CNF formula, $X \subseteq var(F)$ and $C \subseteq cla(F)$ defines a cut.



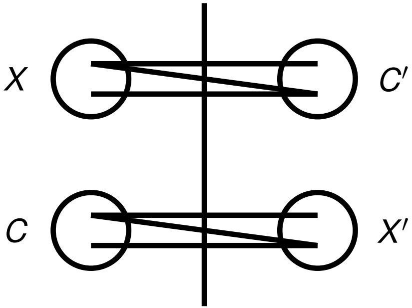
Equivalence over a cut

Let F be a CNF formula, $X \subseteq var(F)$ and $C \subseteq cla(F)$ defines a cut.



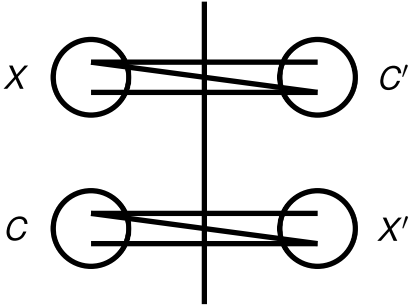
Equivalence over a cut

Let F be a CNF formula, $X \subseteq var(F)$ and $C \subseteq cla(F)$ defines a cut.



Equivalence over a cut

Let F be a CNF formula, $X \subseteq var(F)$ and $C \subseteq cla(F)$ defines a cut.



We define **ps-value of a cut** as the ps-value of the formula where each variable is removed from all clauses on the same side as the variable.

PS-width

1. The **ps-width** of a decomposition (ordering) is the max ps-value over all cuts defined by the decomposition.
2. The **ps-width** of a CNF formula is the min ps-width over all decompositions.

Runningtime for MAX SAT

Theorem

Given a formula F over n variables and m clauses and of total size s, and a decomposition of F of $\mathtt{ps}$-width k, we solve #SAT, and weighted MAXSAT in time $\mathcal{O}(k^3 s(m+n))$.

If the decomposition is a linear order the runningtime can be improved by a factor k.

Formulas of linear PS-width

We say a formula has an interval order if:

- Each variable and clause can be assigned an interval of the real line.
- Such that a variable x is in a clause c if and only if the interval of x intersects the interval of c.

Incidence graphs of such formulas are called interval bigraphs.

Formulas of linear PS-width

We say a formula has an interval order if:

- Each variable and clause can be assigned an interval of the real line.
- Such that a variable x is in a clause c if and only if the interval of x intersects the interval of c.

Incidence graphs of such formulas are called interval bigraphs.

Many other classes of bipartite intersection graphs correspond to formulas which have polynomial ps-width.

Relation to graph decomposition

- A tree-decomposition of $I(F)$ of tree-width $O(log(n))$ can be turned into a decomposition of polynomial ps-width.
- A clique-decomposition of $I(F)$ of constant clique-width can be turned into a decomposition of polynomial ps-width.

Relation to graph decomposition

- A tree-decomposition of $I(F)$ of tree-width $O(log(n))$ can be turned into a decomposition of polynomial ps-width.
- A clique-decomposition of $I(F)$ of constant clique-width can be turned into a decomposition of polynomial ps-width.
- An active community study a wide range of width parameters such as:
  tree-width, branch-width, rank-width, boolean-width, clique-width, cut-width, MM-width, MIM-width ...
  Bounding any of these parameters would prove polynomial ps-width.

Future research

1. Can we approximate ps-width?
2. Can we recognize graphs of MIM-width 1?
3. Does real world SAT instances have low ps-width?

THANK YOU