The Complexity of Minimum Equivalent Formula

Dave Buchfuhrer Chris Umans

California Institute of Technology

August 1, 2007

Outline

1. Problem Definition
2. The Reduction
   - Modified Succinct Set Cover
   - The Idea
3. The Proof
   - Ensuring one z
   - Ensuring Positive z
   - Ensuring z under 2^{nd}-level AND
   - Other sub-formula is $\overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_\ell}}$

The Problem

- Given a formula F and a maximum depth d, does there exist an equivalent depth d formula of size at most k?
- In Σ_2^p
- Known to be Σ_2^p-complete for $d = 2$ (Umans '98)
- Open problem mentioned often in literature for $d = 3$
- Difficult due to inherent need to incorporate formula lower bounds into reduction
- We show that this problem is Σ_2^p-complete for $d \geq 3$, as well as unlimited depth

The Problem

- Given a formula *F* and a maximum depth *d*, does there exist an equivalent depth *d* formula of size at most *k*?
- In Σ_2^p
- Known to be Σ_2^p-complete for $d = 2$ (Umans '98)
- Open problem mentioned often in literature for $d = 3$
- Difficult due to inherent need to incorporate formula lower bounds into reduction
- We show that this problem is Σ_2^p-complete for $d \geq 3$, as well as unlimited depth

The Problem

- Given a formula F and a maximum depth d, does there exist an equivalent depth d formula of size at most k?
- In Σ_2^p
- Known to be Σ_2^p-complete for $d = 2$ (Umans '98)
- Open problem mentioned often in literature for $d = 3$
- Difficult due to inherent need to incorporate formula lower bounds into reduction
- We show that this problem is Σ_2^p-complete for $d \geq 3$, as well as unlimited depth

The Problem

- Given a formula F and a maximum depth d, does there exist an equivalent depth d formula of size at most k?
- In Σ_2^p
- Known to be Σ_2^p-complete for $d = 2$ (Umans '98)
- Open problem mentioned often in literature for $d = 3$
- Difficult due to inherent need to incorporate formula lower bounds into reduction
- We show that this problem is Σ_2^p-complete for $d \geq 3$, as well as unlimited depth

The Problem

- Given a formula F and a maximum depth d, does there exist an equivalent depth d formula of size at most k?
- In Σ_2^p
- Known to be Σ_2^p-complete for $d = 2$ (Umans '98)
- Open problem mentioned often in literature for $d = 3$
- Difficult due to inherent need to incorporate formula lower bounds into reduction
- We show that this problem is Σ_2^p-complete for $d \geq 3$, as well as unlimited depth

The Problem

- Given a formula F and a maximum depth d, does there exist an equivalent depth d formula of size at most k?
- In Σ_2^p
- Known to be Σ_2^p-complete for $d = 2$ (Umans '98)
- Open problem mentioned often in literature for $d = 3$
- Difficult due to inherent need to incorporate formula lower bounds into reduction
- We show that this problem is Σ_2^p-complete for $d \geq 3$, as well as unlimited depth

Outline

Modified Succinct Set Cover

Definition

SSC: Given a DNF D, a collection of variables x_i and an integer k, is there a tautology of the form $D \vee \overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_k}}$?

Definition

MSSC: Same as SSC, but we ask not for a tautology, but to accept everything but the all true assignment.

Modified Succinct Set Cover

Definition

SSC: Given a DNF D, a collection of variables x_i and an integer k, is there a tautology of the form $D \vee \overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_k}}$?

Definition

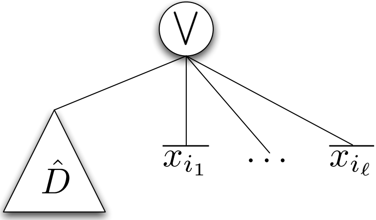
MSSC: Same as SSC, but we ask not for a tautology, but to accept everything but the all true assignment.

Outline

1. Problem Definition
2. The Reduction
   - Modified Succinct Set Cover
   - The Idea
3. The Proof
   - Ensuring one z
   - Ensuring Positive z
   - Ensuring z under 2^{nd}-level AND
   - Other sub-formula is $\overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_\ell}}$

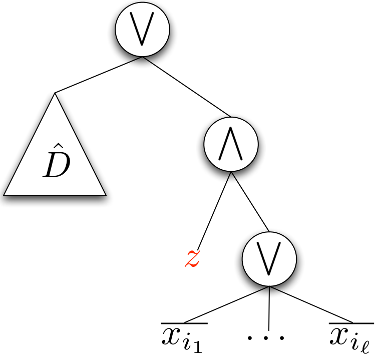
Ideally...

We'd like to ask about a minimum formula for $D \vee \overline{x_1} \vee \cdots \vee \overline{x_n}$ know that it is of the form



Enforcement

In order to enforce this form, we add a control variable z, and ask about $D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$, and the minimum formula will be of the form



Ensuring the z works

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

Three things must occur before the previous form is achieved

- z occurs only once
- z occurs positively
- z is located under a second-level AND gate
- The other sub-formula is $\overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_\ell}}$

Ensuring the z works

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

Three things must occur before the previous form is achieved

- z occurs only once
- z occurs positively
- z is located under a second-level AND gate
- The other sub-formula is $\overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_\ell}}$

Ensuring the z works

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

Three things must occur before the previous form is achieved

- z occurs only once
- z occurs positively
- z is located under a second-level AND gate
- The other sub-formula is $\overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_\ell}}$

Ensuring the z works

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

Three things must occur before the previous form is achieved

- z occurs only once
- z occurs positively
- z is located under a second-level AND gate
- The other sub-formula is $\overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_\ell}}$

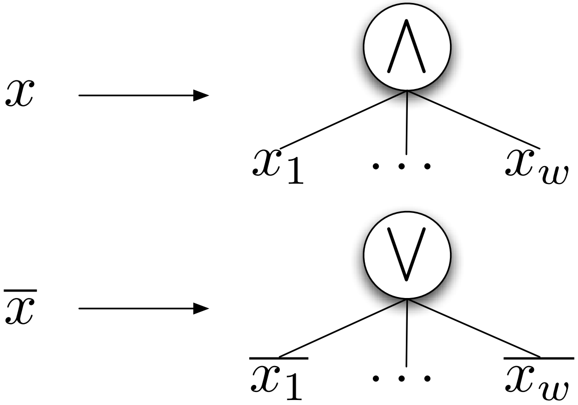
The Reduction

Simply adding z to the formula is not enough to ensure all of these conditions, so we will add elements to the reduction as needed. As is usual with these types of reductions, the positive case is trivial, so we will focus on the negative.

Outline

1. Problem Definition
2. The Reduction
   - Modified Succinct Set Cover
   - The Idea
3. The Proof
   - Ensuring one z
   - Ensuring Positive z
   - Ensuring z under 2^{nd}-level AND
   - Other sub-formula is $\overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_\ell}}$

Weighting


x
$\wedge$
$x_1 \quad \cdots \quad x_w$
$\overline{x}$
$\vee$
$\overline{x_1} \quad \cdots \quad \overline{x_w}$

Weighting

Lemma

The size of the minimum weighted formula for F is a lower bound on the size of the minimum expanded formula for F.

Proof.

For each x, choose the x_i that occurs least frequently in the expanded formula, restrict all others to true. □

Weighting

Lemma

The size of the minimum weighted formula for F is a lower bound on the size of the minimum expanded formula for F.

Proof.

For each x, choose the x_i that occurs least frequently in the expanded formula, restrict all others to true. □

Choosing the Weight for z

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

- Is the min formula of size at most $|\hat{D}| + k + w(z)$?
- Choose $w(z) = |\hat{D}| + k + 1$
- Is the min formula of size at most $2|\hat{D}| + 2k + 1$?
- $2w(z) = 2|\hat{D}| + 2k + 2$

Choosing the Weight for z

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

- Is the min formula of size at most $|\hat{D}| + k + w(z)$?
- Choose $w(z) = |\hat{D}| + k + 1$
- Is the min formula of size at most $2|\hat{D}| + 2k + 1$?
- $2w(z) = 2|\hat{D}| + 2k + 2$

Choosing the Weight for z

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

- Is the min formula of size at most $|\hat{D}| + k + w(z)$?
- Choose $w(z) = |\hat{D}| + k + 1$
- Is the min formula of size at most $2|\hat{D}| + 2k + 1$?
- $2w(z) = 2|\hat{D}| + 2k + 2$

Choosing the Weight for z

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

- Is the min formula of size at most $|\hat{D}| + k + w(z)$?
- Choose $w(z) = |\hat{D}| + k + 1$
- Is the min formula of size at most $2|\hat{D}| + 2k + 1$?
- $2w(z) = 2|\hat{D}| + 2k + 2$

Outline

1. Problem Definition
2. The Reduction
   - Modified Succinct Set Cover
   - The Idea
3. The Proof
   - Ensuring one z
   - **Ensuring Positive z**
   - Ensuring z under 2^{nd}-level AND
   - Other sub-formula is $\overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_\ell}}$

z Must Occur Positively

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

- Since negations are pushed to the bottom, F is monotonic in its literals
- F accepts more when z is true
- If z only occurred negatively, F would accept less when z is true

Outline

z Cannot Occur Under Top-Level OR

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

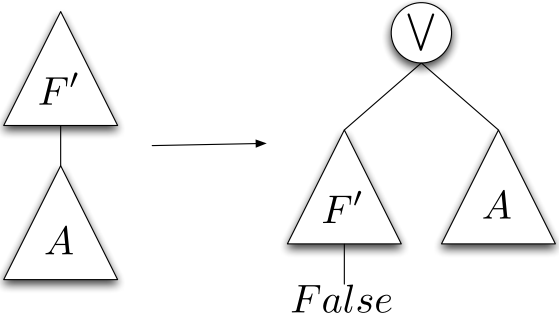
If z occured under the top-level OR, the formula would accept the all true assignment

A Quick Digression...

Lemma

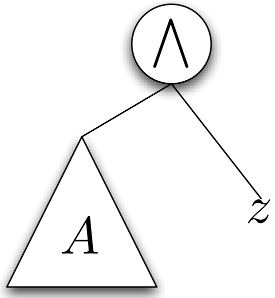
Suppose that a formula F has sub-formula A. If A implies F, then there is an equivalent formula of the same size as F in which A occurs directly under the top OR gate.

A Proof in a Picture



z Cannot Occur Under Low-Level AND

Suppose that z occurs under a low-level AND gate



Low-Level AND Continued

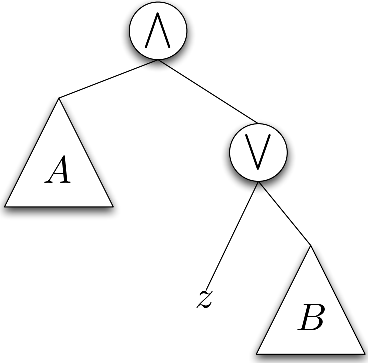
The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

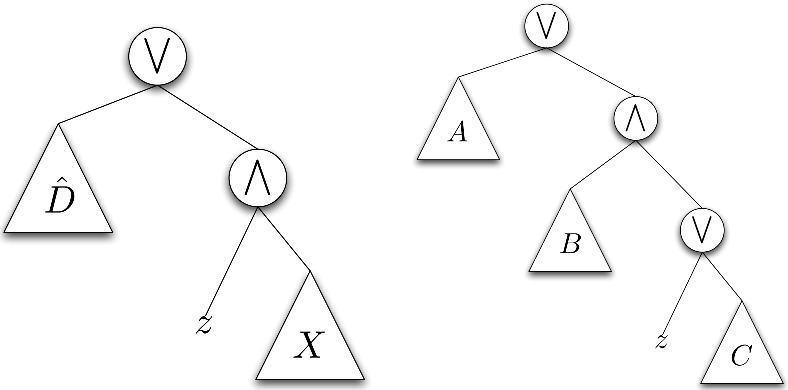
- If A accepts all true, it may as well accept everything
- If A does not accept all true, we can move the entire sub-formula up

z Cannot Occur Under Low-Level OR

If z is under a low-level OR, the proof goes the same way

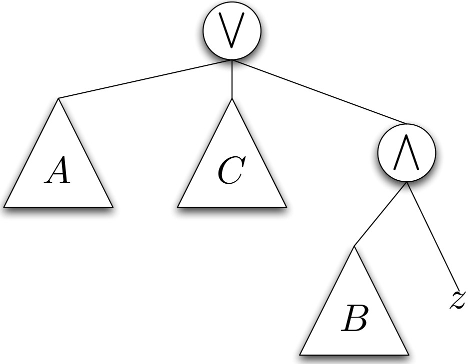


z Cannot Occur Under 3^{rd}-Level OR



C Must Accept All True

If C doesn't accept all true, it implies F, so we can move C up



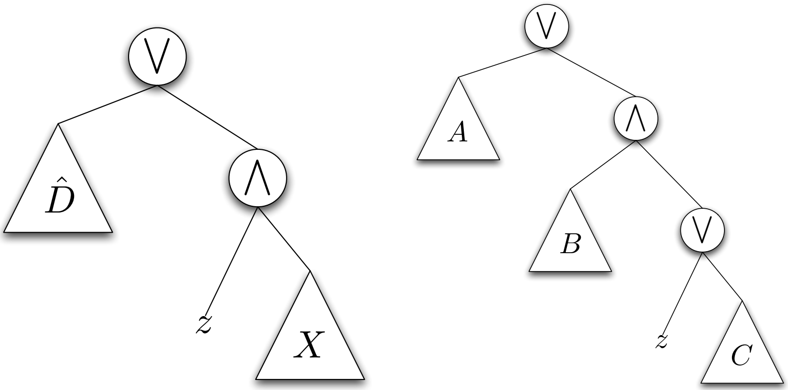
C At Least As Big As X

- C must accept the all true
- C must not accept anything else that D does not
- Thus, $\overline{C}$ could be substituted for X

C At Least As Big As X

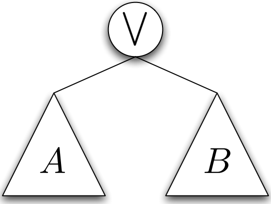
- C must accept the all true
- C must not accept anything else that D does not
- Thus, $\overline{C}$ could be substituted for X

C At Least As Big As X



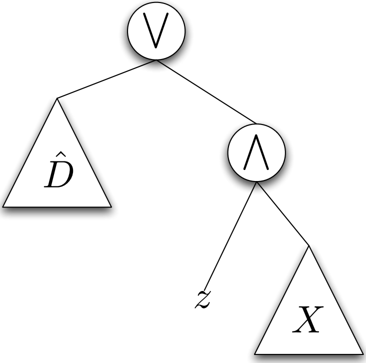
$A \vee B$ Depends on Every Variable

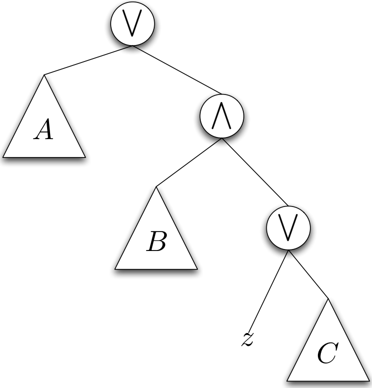
If we substitute $z = \text{true}$, we get



which must accept everything but all true, and thus depends on all variables

Adding Dummy Variables





Dummy Result

Once the dummy variables have been added, the formula is of the below form

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n} \vee \overline{y_1} \vee \cdots \vee \overline{y_{w(y)}})]$$

Outline

Minimum Formula Accepting Enough

Lemma

The minimum formula that does not accept the all true and does accept at least some other fixed set of assignments is of the form $\overline{v_1} \vee \cdots \vee \overline{v_n}$.

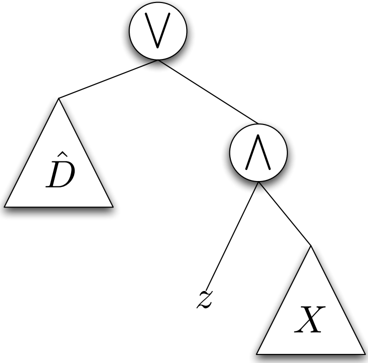
Proof of Lemma

Proof.

Let F be a smallest such formula. Suppose that there is a smaller formula F'. F' has fewer variables . Take the disjunction of their negations, and there is some required assignment it doesn't cover. All the variables F' depends on must be true in this assignment, so to cover it, F' must accept the all true assignment. □

Applying the Lemma

X must accept everything not accepted by D other than the all true



Ensuring Only x_is

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

- Recall that we were given a DNF D and a collection of variables $x_1, \ldots, x_n$
- We don't want variables other than $x_1, \ldots, x_n$ occurring in X
- To assure this, we weight all other variables by $n + 1$

Ensuring Only x_is

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

- Recall that we were given a DNF D and a collection of variables $x_1, \ldots, x_n$
- We don't want variables other than $x_1, \ldots, x_n$ occurring in X
- To assure this, we weight all other variables by $n + 1$

Ensuring Only x_is

The Formula F

$$D \vee [z \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n})]$$

- Recall that we were given a DNF D and a collection of variables $x_1, \ldots, x_n$
- We don't want variables other than $x_1, \ldots, x_n$ occurring in X
- To assure this, we weight all other variables by $n + 1$

The Entire Reduction

We start with a DNF D, variables $x_1, \ldots, x_n$ and an integer k.
We create the formula

$$D' \vee [(z_1 \wedge \cdots \wedge z_{2v+k+n+1}) \wedge (\overline{x_1} \vee \cdots \vee \overline{x_n} \vee \overline{y_1} \vee \cdots \vee \overline{y_{v+n}})],$$

where D' is equal to D, but with all variables other than the x_i weighted by $n+1$, and v is the size of a minimum formula for D'.
The new formula has an equivalent formula of size at most $4v + 2k + 2n + 1$ iff there is a formula of the form $D \vee \overline{x_{i_1}} \vee \cdots \vee \overline{x_{i_k}}$ accepting everything but the all true assignment.

Conclusions

- Main Points
  - Lower bounds based on number of dependent variables can be powerful
  - Indicator variables can be used to dictate the form of a formula
- Future Work
  - Find a many to one reduction
  - Determine the complexity of
    - the non-succinct version
    - the circuit version
    - approximation

Conclusions

- Main Points
  - Lower bounds based on number of dependent variables can be powerful
  - Indicator variables can be used to dictate the form of a formula
- Future Work
  - Find a many to one reduction
  - Determine the complexity of
    - the non-succinct version
    - the circuit version
    - approximation