The complexity of Boolean formula minimization

Dave Buchfuhrer   Chris Umans

July 7, 2008

Formula Minimization Problem

- You wish to compute some function f

Formula Minimization Problem

- You wish to compute some function f
- So you create a formula F computing f
- You want F small

Formal Definition

Problem (Minimum Equivalent Expression)

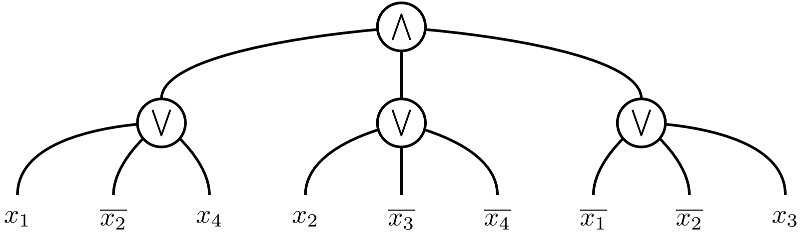
Given a formula F and an integer k, is there a formula F' equivalent to F of size at most k?

- Size is defined to be the number of occurrences of input variables in the formula.
- In Σ_2^P

Example

Problem (Minimum Equivalent Expression)

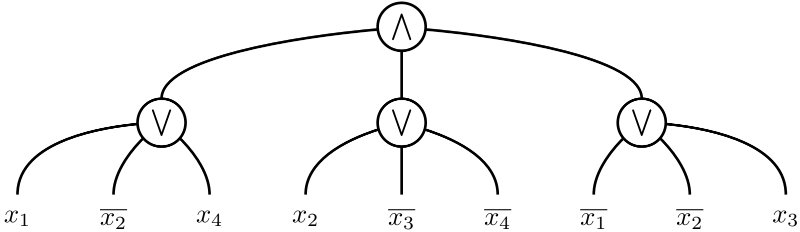
Given a formula F and an integer k, is there a formula F' equivalent to F of size at most k?



Example

Problem (Minimum Equivalent Expression)

Given a formula F and an integer k, is there a formula F' equivalent to F of size at most k?

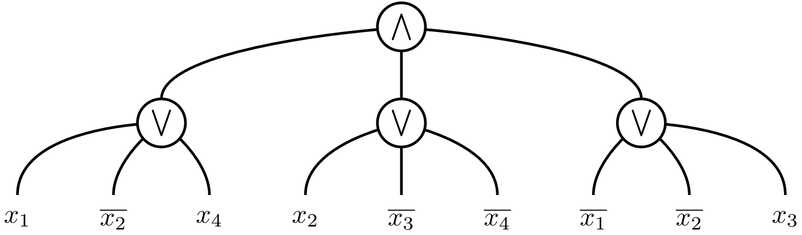


- The size of this formula is 9

Example

Problem (Minimum Equivalent Expression)

Given a formula F and an integer k, is there a formula F' equivalent to F of size at most k?



- The size of this formula is 9
- What's special about $k = 0$ here?

History of the Problem

- Defined in the early 70's by Meyer and Stockmeyer, inspired the Polynomial Hierarchy
- Clearly coNP-hard
- Proven $\mathrm{P}_{||}^{\mathrm{NP}}$-hard in 1997 (Hemaspaandra and Wechsung)
- DNF version proven Σ_2^{P}-complete in 1999 (Umans)
- We show that Minimum Equivalent Expression is Σ_2^P-complete under Turing reductions, both for unrestricted formulas and for formula restricted to any fixed depth $d \geq 3$

Why is it hard?

- In the hard direction of the reduction, we need a formula lower bound
- Circuit and formula lower bounds are hard
- We make use of very simple lower bounds

Outline

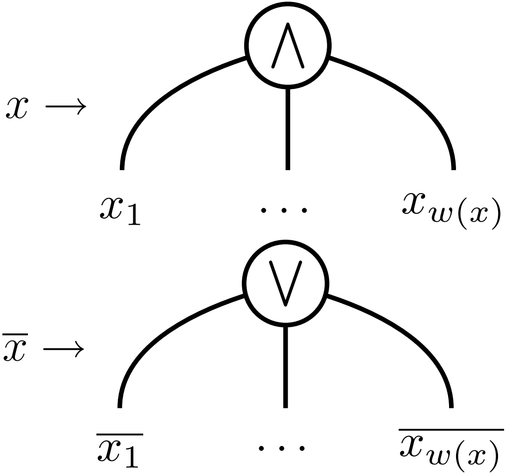
Weighting

- It would be convenient to count each variable x as having weight $w(x)$
- Can this be done without changing the problem definition?

Weighting

- It would be convenient to count each variable x as having weight $w(x)$
- Can this be done without changing the problem definition?
- Our idea: replace x with $x_1 \wedge \cdots \wedge x_{w(x)}$

Variable Weighting



The Results of Weighting

- We start with a formula F computing $f\left(x^{(1)}, \ldots, x^{(n)}\right)$
- We end with a formula F' computing

$$f' = f\left(x_1^{(1)} \wedge \cdots \wedge x_{w\left(x^{(1)}\right)}^{(1)}, \ldots, x_1^{(n)} \wedge \cdots \wedge x_{w\left(x^{(n)}\right)}^{(n)}\right)$$

The Results of Weighting

- We start with a formula F computing $f\left(x^{(1)}, \ldots, x^{(n)}\right)$
- We end with a formula F' computing

$$f' = f\left(x_1^{(1)} \wedge \cdots \wedge x_{w\left(x^{(1)}\right)}^{(1)}, \ldots, x_1^{(n)} \wedge \cdots \wedge x_{w\left(x^{(n)}\right)}^{(n)}\right)$$

Lemma

The minimum formula F' equivalent to F after expanding the weights is at least as large as the minimum weighted formula for F.

Weighting Lemma

Lemma

The minimum formula F' equivalent to F after expanding the weights is at least as large as the minimum weighted formula for F.

Proof

- Each variable x becomes $x_1 \wedge \cdots \wedge x_{w(x)}$ in the expanded form

Weighting Lemma

Lemma

The minimum formula F' equivalent to F after expanding the weights is at least as large as the minimum weighted formula for F.

Proof

- Each variable x becomes $x_1 \wedge \cdots \wedge x_{w(x)}$ in the expanded form
- Take the i^* such that x_{i^*} occurs least frequently of all x_i in F'
- Restrict $x_i = \text{True}$ for $i \neq i^*$ to arrive at F''
- Under this restriction, $x_1 \wedge \cdots \wedge x_{w(x)}$ becomes x_{i^*}

Weighting Lemma

Lemma

The minimum formula F' equivalent to F after expanding the weights is at least as large as the minimum weighted formula for F.

Proof

- Each variable x becomes $x_1 \wedge \cdots \wedge x_{w(x)}$ in the expanded form
- Take the i^* such that x_{i^*} occurs least frequently of all x_i in F'
- Restrict $x_i = \mathrm{True}$ for $i \neq i^*$ to arrive at F''
- Under this restriction, $x_1 \wedge \cdots \wedge x_{w(x)}$ becomes x_{i^*}
- F'' is equivalent to F
- F' has as many x_i as $w(x)$ times the number of x_{i^*} in F'' □

Modified Succinct Set Cover

Problem (Modified Succinct Set Cover)

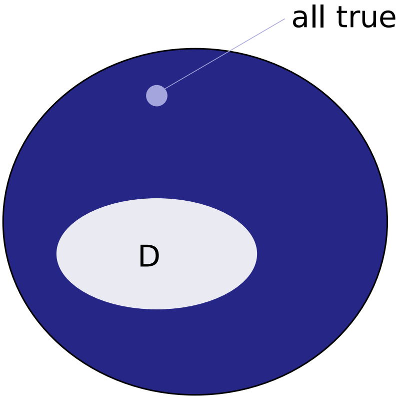
Given a DNF formula D, variables $x_1, \ldots, x_n$ and an integer k, where D is a formula on variables $x_1, \ldots, x_n, v_1, \ldots, v_n$, is there a set I of size at most k such that

$$D \vee \bigvee_{i \in I} \overline{x_i} \equiv D \vee \bigvee_{i=1}^{n} \overline{x_i} \equiv \bigvee_{i=1}^{m} \overline{v_i} \vee \bigvee_{i=1}^{n} \overline{x_i}?$$

- Basically, we want to know how many $\overline{x_i}$ are necessary to cover the assignments not accepted by D, other than the all true assignment
- Slight modification of problem used to prove DNF version Σ_2^P-complete (Umans 1999)

Succinct Set Cover Visualized

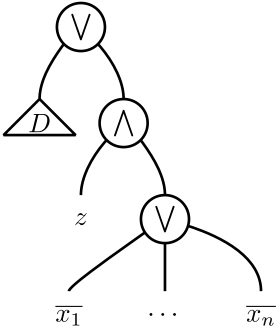
- Each set $\overline{x_i}$ covers half of all points
- None cover the all true point
- How many are necessary to cover the dark blue region?



$\{\text{true}, \text{false}\}^{m+n}$

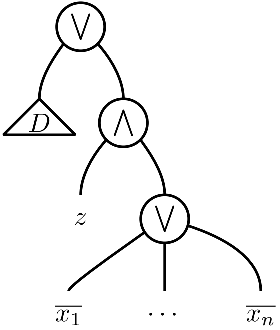
Overview

- We start with Modified Succinct Set Cover instance $\langle D, x_1, \ldots, x_n, k \rangle$
- We create the Minimum Equivalent Expression instance with formula $D \vee (z \wedge \bigvee_i \overline{x_i})$ and size target $|\widehat{D}|_w + w(z) + k$
- Finding $|\widehat{D}|_w$ necessitates a Turing reduction



Overview

- When z is false, the formula becomes simply D
- When z is true, it "unlocks" a portion computing the set cover



The Easy Direction

The question asked by reduction

Is there a formula for $D \vee (z \wedge \bigvee_i \overline{x_i})$ of size $|\widehat{D}|_w + w(z) + k$?

Easy direction: The Modified Succinct Set Cover is positive, so

$$D \vee \bigvee_{i \in I} \overline{x_i} \equiv D \vee \bigvee_{i=1}^{n} \overline{x_i}$$

which gives us the formula

$$\widehat{D} \vee \left(z \wedge \bigvee_{i \in I} \overline{x_i} \right) \equiv D \vee \left(z \wedge \bigvee_{i=1}^{n} \overline{x_i} \right)$$

of size $|\widehat{D}|_w + w(z) + k$

The Hard Direction

The question asked by reduction

Is there a formula for $D \vee (z \wedge \bigvee_i \overline{x_i})$ of size $|\widehat{D}|_w + w(z) + k$?

- If the Modified Succinct Set Cover instance is negative, there shouldn't be a small equivalent formula

The Hard Direction

The question asked by reduction

Is there a formula for $D \vee (z \wedge \bigvee_i \overline{x_i})$ of size $|\widehat{D}|_w + w(z) + k$?

- If the Modified Succinct Set Cover instance is negative, there shouldn't be a small equivalent formula
- We show that one such formula must be $\widehat{D} \vee (z \wedge \bigvee_{i \in I} \overline{x_i})$

The Hard Direction

The question asked by reduction

Is there a formula for $D \vee (z \wedge \bigvee_i \overline{x_i})$ of size $|\widehat{D}|_w + w(z) + k$?

- If the Modified Succinct Set Cover instance is negative, there shouldn't be a small equivalent formula
- We show that one such formula must be $\widehat{D} \vee (z \wedge \bigvee_{i \in I} \overline{x_i})$
- z is weighted such that $2w(z) > |\widehat{D}|_w + w(z) + k$

The Hard Direction

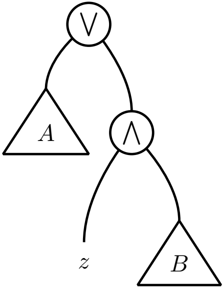
The question asked by reduction

Is there a formula for $D \vee (z \wedge \bigvee_i \overline{x_i})$ of size $|\widehat{D}|_w + w(z) + k$?

- If the Modified Succinct Set Cover instance is negative, there shouldn't be a small equivalent formula
- We show that one such formula must be $\widehat{D} \vee (z \wedge \bigvee_{i \in I} \overline{x_i})$
- z is weighted such that $2w(z) > |\widehat{D}|_w + w(z) + k$
- The position of the z is proven through case analysis and requires slight modifications to the reduction

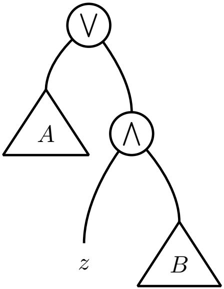
Consequences of z Position

- Given this positioning, $A \equiv D$ and B computes the set cover
- If we weight all variables other than the x_i by more than k, B can only contain x_i variables



Consequences of z Position

- Given this positioning, $A \equiv D$ and B computes the set cover
- If we weight all variables other than the x_i by more than k, B can only contain x_i variables



Lemma

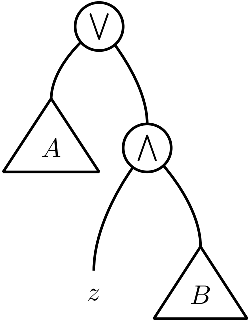
A minimum formula accepting a set S but not the all true assignment is of the form $\bigvee_i \overline{x_i}$

Wrapping it up

The question asked by reduction

Is there a formula for $D \vee (z \wedge \bigvee_i \overline{x_i})$ of size $|\widehat{D}|_w + w(z) + k$?

- $A \equiv D$, so $|A|_w \geq |\widehat{D}|_w$
- So size is at least $|\widehat{D}|_w + w(z) + |B|_w$
- As shown above, $|B|_w \leq k$ only if the Modified Succinct Set Cover instance is positive



Open Problems

- Can the Σ_2^P-completeness result be shown without Turing reductions?
- What is the complexity of approximation?
- What is the complexity when minimizing circuits rather than formulas?

Open Problems

- Can the Σ_2^P-completeness result be shown without Turing reductions?
- What is the complexity of approximation?
- What is the complexity when minimizing circuits rather than formulas?

A full version of this paper is available at
`http://www.cs.caltech.edu/~dave/papers/`