On the size of irredundant propagation complete CNF formulas

Petr Savický
Institute of Computer Science,
Czech Academy of Sciences

Liblice, September 2023

Contents

Preliminary remarks:

- main result: on **irredundant** **propagation complete** CNF formulas
- more general context: arbitrary CNF formulas which are **irredundant** with respect to **unit clause propagation (ucp)**

Contents:

1. ucp-equivalence and ucp-irredundancy of CNF formulas
2. the ratio of the size of a ucp-irredundant and a smallest ucp-equivalent formula
3. an **upper bound** n^2 on this ratio
4. a **lower bound** $\Omega(n/\ln n)$ on this ratio (main result)
5. some parts of the proof (using known results on covering numbers for uniform hypergraphs)

Ucp-equivalence

In some contexts, we need to distinguish CNF formulas according to the strength of **unit clause propagation (ucp)** on them (learned clauses in a SAT solver).

Definition

Formulas φ_1 and φ_2 are **ucp-equivalent**, if for every partial assignment α, unit propagation derives either the same consistent set of literals from $\varphi_1 \wedge \alpha$ and $\varphi_2 \wedge \alpha$ or a contradiction in both cases.

ucp-equivalent $\Rightarrow$ **equivalent** (α can be a total assignment).

algorithmic problem: For a given formula, find a ucp-equivalent formula as small as possible.
a base-line: consider irredundant formulas

Irredundancy w.r.t. unit clause propagation

Bordeaux and Marques-Silva, 2012:
minimal PC formula using the notion of **absorbed clause**.

a straightforward generalization:

Definition

Formula φ is **ucp-irredundant** if (equivalent conditions):

- none of the clauses $C \in \varphi$ is absorbed by $\varphi \setminus \{C\}$ or
- for every $C \in \varphi$, φ and $\varphi \setminus \{C\}$ are not ucp-equivalent.

The notions are efficient in the following sense:

- test of absorbed clause is in **poly-time**
- ucp-equivalence can be tested in **poly-time**
- any formula can be made ucp-irredundant in **poly-time**

Dual-rail encoding (a remark)

For any CNF formula φ, the Horn formula $\mathrm{DR}(\varphi)$ **represents unit clause propagation** on φ.

Idea: for every literal l introduce variable $[\![l]\!]$ (l is derived)

If $\varphi = C_1 \wedge \ldots \wedge C_m$ on variables $x_1, \ldots, x_n$, then

$$\mathrm{DR}(\varphi) = \mathrm{DR}(C_1) \wedge \ldots \wedge \mathrm{DR}(C_m) \wedge \bigwedge_{i=1}^{n} (\neg[\![x_i]\!] \vee \neg[\![\neg x_i]\!])$$

where for a clause $C = l_1 \vee \ldots \vee l_k$,

$$\mathrm{DR}(C) = \bigwedge_{i=1}^{k} \left(\bigwedge_{j \neq i} [\![\neg l_j]\!] \rightarrow [\![l_i]\!] \right)$$

The size of ucp-irredundant formulas

Formulas φ_1 and φ_2 are **ucp-equivalent** if and only if the Horn formulas $\mathrm{DR}(\varphi_1)$ and $\mathrm{DR}(\varphi_2)$ are **equivalent**.

An **irredundant Horn formula** is at most n times larger than a minimal equivalent formula.
$\Rightarrow$ an upper bound $2n^2$ for **ucp-irredundant formulas**

Proposition

Let φ be ucp-irredundant and let φ^* be a minimal formula ucp-equivalent to φ, both of n variables, then

$$|\varphi| \leq n^2 |\varphi^*|$$

$$\|\varphi\| \leq n^2 \|\varphi^*\|$$

A proof of the upper bound

Assume, φ and φ^* are ucp-equivalent and ucp-irredundant.

For every $l \in C \in \varphi^*$, unit clause propagation derives l from $\varphi^* \wedge \mathrm{neg}(C \setminus l)$ and, hence, also from $\varphi \wedge \mathrm{neg}(C \setminus l)$.

Let φ' be the subset φ needed for this. Clearly, $|\varphi'| \leq n\|\varphi^*\|$.

Using the assumption, we obtain $\varphi' = \varphi$, so we have

$$|\varphi| \leq n\|\varphi^*\|$$

and, hence,

$$|\varphi| \leq n\|\varphi^*\| \leq n^2|\varphi^*|$$

$$\|\varphi\| \leq n|\varphi| \leq n^2\|\varphi^*\|$$

Question: Can the upper bounds be improved and how much?

The result

Theorem

For every large enough n, there is a ucp-irredundant formula φ^ℓ of n variables, such that

$$|\varphi^\ell| = \Omega(n/\ln n)|\varphi^*|$$

$$\|\varphi^\ell\| = \Omega(n/\ln n)\|\varphi^*\|$$

where φ^* is a minimal formula ucp-equivalent to φ^ℓ.

The proof uses PC formulas

A formula φ is **propagation complete (PC)**, if it is ucp-equivalent to the set of all its prime implicates.

Formulas φ^ℓ and φ^* will be representations of a specific symmetric definite Horn function.

A symmetric definite Horn function

Let n and $k = n/2 + O(1)$ be fixed, let X be a set of n variables.

Let $\Psi_{n,k}$ be the set of all definite Horn clauses of length $k + 1$ over the n variables X.

$\Psi_{n,k}$ is the set of **all prime implicates** of a symmetric definite Horn function which is also a key-Horn function.

Goal: Find prime ucp-irredundant formulas φ^ℓ (large) and φ^* (small) which are ucp-equivalent to $\Psi_{n,k}$.

φ^ℓ and φ^* will be **hypergraph formulas** defined as

$$\theta(H) = \{\mathrm{neg}(e \setminus \{x\}) \cup \{x\} \mid e \in H\}$$

where H is a suitable $(k + 1)$-uniform hypergraph.

Hypergraph formulas ucp-equivalent to $\Psi_{n,k}$

For a $(k+1)$-uniform hypergraph H on X and
a set $A \subseteq X$, $|A| = k-1$, consider the graph on $X \setminus A$

$$G(H, A) = \{\{x_i, x_j\} \mid A \cup \{x_i, x_j\} \in H\}$$

Definition

A $(k+1)$-uniform hypergraph H has **connected restrictions**, if for all $A \subseteq X$, $|A| = k-1$, the undirected graph $G(H, A)$ is connected on $X \setminus A$.

Theorem

$\theta(H)$ is **ucp-equivalent** *to $\Psi_{n,k}$ if and only if H has* **connected restrictions**.

A large irredundant formula

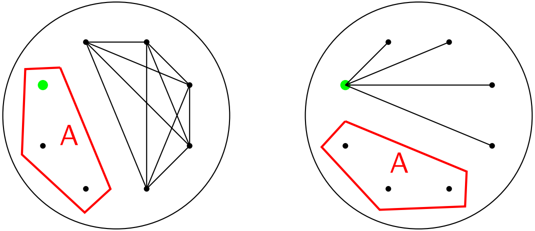
Let $\varphi^\ell = \theta(H)$, where $H = \{B \subseteq X \mid x_1 \in B, |B| = k+1\}$.

Lemma

φ^ℓ is ucp-irredundant and ucp-equivalent to $\Psi_{n,k}$.

Proof:

$G(H, A)$ is connected for every $A \Rightarrow H$ has connected restrictions.



By case inspection, we get that φ is ucp-irredundant.

Covering designs

Construction of a small hypergraph H with connected restrictions (used for φ^*) uses known bounds on the size of covering designs.

A **covering design** with parameters $(n, k+1, k)$ is any system of $(k+1)$-sets of an n-set covering every k-set.

The size of a smallest such system will be denoted $C(n, k+1, k)$.

By result [Erdős, Spencer, 1974] or [Sidorenko, 1997], we have

$$\frac{1}{k+1}\binom{n}{k} \leq C(n, k+1, k) \leq \frac{O(\ln n)}{k+1}\binom{n}{k}$$

Size of a hypergraph with connected restrictions

Theorem

For a large enough n and $k = n/2 + O(1)$, there is a hypergraph H^* with connected restrictions of size at most $6C(n, k+1, k)$.

Sketch of proof:

Let H_0 be a covering design with parameters $(n, k+1, k)$.

For every set A, $|A| = k-1$, the graph $G(H_0, A)$ has non-zero degree for all vertices (variables $X \setminus A$).

Union of 3 random permutations $\Rightarrow$ a connected graph with high probability.

Using this for all sets A, $|A| = k-1$ requires to use 5 random permutations and some additional hyperedges, so we get

$$|H^*| \leq (5 + o(1))|H_0| \leq 6|H_0| = 6C(n, k+1, k)$$

Summary of the bounds

Formula $\theta(H^*)$ is ucp-equivalent to $\Psi_{n,k}$ and has size at most $6(k+1)C(n, k+1, k)$.

Theorem

For every n and $k = n/2 + O(1)$, there is a ucp-irredundant formula φ^ℓ of size

$$|\varphi^\ell| = (k+1)\binom{n-1}{k} = \Theta(n)\binom{n}{k}$$

and such that a smallest ucp-equivalent formula φ^* has size

$$|\varphi^*| \leq 6(k+1)C(n, k+1, k) = O(\ln n)\binom{n}{k}$$

In particular, $|\varphi^\ell| = \Omega(n/\ln n)|\varphi^*|$.

Thank you for attention.