Preliminary Report on
Input Cover Number as a Metric for Propositional Resolution Proofs

Allen Van Gelder
Computer Science Department
University of California
Santa Cruz, CA, USA

avg@cs.ucsc.edu

http://www.cse.ucsc.edu/∼avg/Papers/

What Property Ensures that Propositional Formulas Are "Hard" for Resolution?

Notation: Formula F has N variables, length L.

- Lowercase n is used as parameter within family, e.g. $\mathrm{PHP}(n+1,n)$.

Main Intuition [Ben-Sasson and Wigderson, 2001]:

If

every (general resolution) refutation of a (propositional CNF) formula F *must* contain a "very weak" derived clause,

then

every (resolution) refutation of F *must* have length superpolynomial in L.

Their meaning of "very weak" (*excess clause width* criterion):

"has at least $\Theta\left(N^{\frac{1}{2}+\varepsilon}\right)$ more literals than widest clause in F"

Often need to transform F_0 with wide clauses into F in 3-CNF.

Excess Clause Width Does Not Distinguish Polynomial from Superpolynomial Resolution Length

GT(n) Family [Krishnamurthy 1985]	Pigeon-Hole Family, PHP($n+1,n$)
It is inconsistent that a partially ordered set of n elements has no maximal element.	It is inconsistent that $(n+1)$ pigeons can be mapped 1–1 into n holes.
$N = n^2$	$N = (n+1)n$
n positive clauses of width $n-1$	$n+1$ positive clauses of width n
About n^3 clauses of lengths 2–3.	About $n^3/2$ clauses of length 2.
$L \approx 3N^{3/2}$	$L \approx N^{3/2}$

Remark: Clause distributions typical of many constraint-satisfaction problems.

EPHP($n+1,n$) requires excess clause width $\Theta(N^{\frac{1}{2}})$ [Ben-Sasson and Wigderson].
PHP($n+1,n$) requires *superpolynomial* refutations [Haken, 1985].

MGT(n) requires excess clause width $\Theta(N^{\frac{1}{2}})$ [Bonet and Galesi, 2001].
GT(n) has *linear* refutations [Stålmarck, 1996].

Is There a Better Notion of "Very Weak"?

Input Distance [Van Gelder, 2005]:

- Count only literals that do not occur in any *one* input clause to measure weakness of derived clause D; i.e., $\Delta(D) = \max_{C \in F} |D - C|$.
- Motivation: Ensure that simply rederiving long clauses of original formula does not qualify as "very weak".
- Tolerates formulas with long clauses.

Theorem: $\mathrm{PHP}(n+1,n)$ requires input distance $\Delta \geq \Theta(N^{\frac{1}{2}})$.

That is, after transforming to $\mathrm{EPHP}(n+1,n)$, you not only need to derive a long clause, but it has to be *substantially different* from any long clause in the original $\mathrm{PHP}(n+1,n)$.

Bad news: $\mathrm{GT}(n)$ (apparently) also requires input distance $\Delta \geq \Theta(N^{\frac{1}{2}})$.

New Try for Better Notion of “Very Weak”: Input Cover Number

Input Cover Number of derived clause D, $\kappa(D)$, is the minimum number of input clauses needed to “cover” D’s literals.

- I.e., minimize $|G|$ such that $G \subseteq F$ and $D \subseteq \bigcup G$.
- Motivation: Ensure that simply rederiving long clauses of original formula does not qualify as “very weak”.
- Tolerates formulas with long clauses.

Theorem: PHP$(n+1,n)$ requires input cover number $\kappa \geq \Theta(N^{\frac{1}{2}})$.

Theorem: GT(n) requires input cover number $\kappa \geq \Theta(1)$, e.g., 2 or 3.

Bad news: Input cover number can be “manipulated” by transforming the formulas from their natural representations.

- E.g., MGT(n) requires input cover number $\kappa \geq \Theta(N^{\frac{1}{2}})$.

Future Work: Refine the definition of input cover number so that it is immune to such changes of form.