# Minimal variable-degrees for $\mathcal{LEAN}$ and $\mathcal{MU}$


Oliver Kullmann[1] Xishun Zhao[2]

[1]Computer Science, Swansea University, UK
http://cs.swan.ac.uk/~csoliver

[2]Institute of Logic and Cognition, Sun Yat-sen University, China


SAT 2011
Ann Arbor, June 20, 2011

# The basic notions I

- A *clause* $C$ is a (finite) clash-free set of literals.
- A *clause-set* $F$ is a (finite) set of clauses.
- The set of variables of $F$ is denoted by $\mathbf{var}(F)$.
- The **degree** of a variable $v$ in a clause-set $F$ is the number of its occurrences, denoted by

$$\mathbf{vd}_F(v) := |\{C \in F : v \in C\}| + |\{C \in F : \overline{v} \in C\}| \in \mathbb{N}_0.$$

The **minimal variable degree** of clause-set $F$ is

$$\mu\mathbf{vd}(F) := \min_{v \in \mathrm{var}(F)} \mathrm{vd}_F(v).$$

(A variable realising the min-var-degree of $F$ can be considered a “weak spot” of $F$.)

# The basic notions II

- A *partial assignment* $\varphi$ is a map $\varphi : V \to \{0, 1\}$ for some (finite) set $V$ of variables.
- Application of a partial assignment $\varphi$ to a clause-set $F$ is denoted by $\varphi * F$.
- An **autarky** for a clause-set $F$ is a partial assignment $\varphi$ which satisfies every clause $C \in F$ it touches (i.e., $\mathrm{var}(\varphi) \cap \mathrm{var}(C) \neq \emptyset$).
- An autarky $\varphi$ for $F$ is *trivial* if $\mathrm{var}(\varphi) \cap \mathrm{var}(F) = \emptyset$.

A **lean clause-set** is a clause-set $F$ which has only the trivial autarky.

The class of all lean clause-sets is $\mathcal{LEAN}$.

# The basic notions III

- A **minimally unsatisfiable clause-set** is a clause-set $F$ which is unsatisfiable, while removal of any clause renders it satisfiable.
- The class of all minimally unsatisfiable clause-sets is denoted by $\mathcal{MU}$.

The **deficiency** of a clause-set is

$$\boldsymbol{\delta(F)} := c(F) - n(F)$$

the difference of the number of clauses $c(F) := |F|$ and the number of variables $n(F) := |\mathrm{var}(F)|$.
Recall

$$\mathcal{MU} \subset \mathcal{LEAN}$$
$$\forall F \in \mathcal{LEAN} : \delta(F) \geq 1.$$

# The main results

We study the function **nM** : $\mathbb{N} \to \mathbb{N}$ ("non-Mersenne") with

$$k + \lfloor \log_2(k+1) \rfloor \leq \text{nM}(k) \leq k + 1 + \lfloor \log_2(k) \rfloor.$$

E.g. $\text{nM}(1) = 2$, $\text{nM}(2) = 4$, $\text{nM}(3) = 5$.

For a class $\mathcal{C}$ of clause-sets let

$$\boldsymbol{\mu}\textbf{vd}(\boldsymbol{\mathcal{C}}) := \sup_{F \in \mathcal{C}} \mu\text{vd}(F).$$

The main results are:

- For all $k \in \mathbb{N}$ we have $\mu\text{vd}(\mathcal{LEAN}_{\delta=k}) = \text{nM}(k)$.
- We introduce
  $\text{nM} > \text{nM}_1 > \text{nM}_2 > \cdots > \text{nM}_\omega > \text{nM}_{\omega+1}$ with
  $\mu\text{vd}(\mathcal{MU}_{\delta=k}) \leq \text{nM}_{\omega+1}(k)$.

# Open problems



The open problems are whether the bound for $\mathcal{LEAN}$ can be made efficient, and whether the bound for $\mathcal{MU}$ can be further improved:

1. For a clause-set $F$ with $k := \delta(F)$ and

$$\mu\mathrm{vd}(F) > k + 1 + \log_2(k) \geq \mathrm{nM}(k)$$

   there exist a non-trivial autarky — can we **find it** in polynomial time?
2. Is the bound $\mu\mathrm{vd}(\mathcal{MU}_{\delta=k}) \leq \mathrm{nM}_{\omega+1}(k)$ sharp, or are there more (and more) improvements?
3. We conjecture $\mu\mathrm{vd}(\mathcal{MU}_{\delta=k}) \geq \mathrm{nM}(k) - 1$.

# Potential applications

- If the autarky can be found efficiently, then it might be interesting for SAT solving.
- The upper bound on the min-var-degree for $\mathcal{MU}$ could be used for finding minimally unsatisfiable sub-clause-sets more efficiently.

Last, but not least, these results and the underlying methods are steps towards

the **classification** of $\mathcal{MU}_{\delta=k}$
for all deficiencies $k = 1, 2, \ldots$

(The situation for $k = 1$ is very well known, for $k = 2$ quite well (in both cases important work by Hans Kleine Büning), and for bigger $k$ we are building up knowledge.)

# Outline

# Saturation

Minimally unsatisfiable clause-sets are unsatisfiable clause-sets which are quite close to satisfiable clause-sets:

1. We can push it even further by weakening the clauses, i.e., adding literals (of course, not introducing new variables).
2. If this is no longer possible, then we arrive at **saturated minimally unsatisfiable clause-sets**.
3. The class is $\boldsymbol{\mathcal{SMU}} \subset \mathcal{MU}$.
4. Every $F \in \mathcal{MU}$ can be **saturated**, yielding $F' \in \mathcal{SMU}$ with $n(F') = n(F)$ and $c(F') = c(F)$ (thus $\delta(F') = \delta(F)$).

For $F \in \mathcal{MU}$ we have

$$F \in \mathcal{SMU} \Leftrightarrow \forall v \in \mathrm{var}(F) \, \forall \varepsilon \in \{0, 1\} : \langle v \to \varepsilon \rangle * F \in \mathcal{MU}.$$

# The fundamental observations

The basic technique here to prove properties of $F \in \mathcal{MU}$ is

1. first saturate $F$, obtaining $F'$;
2. now use splitting on a variable $v \in \mathrm{var}(F') = \mathrm{var}(F)$,
3. together with induction on the deficiency.

In order to get the deficiency down, $v$ must not be **singular**, that is, $v$ must occur in both signs at least twice.

What if besides $v$ also other variables vanish in $\langle v \to \varepsilon \rangle * F$ ?

This can not happen if the degree of $v$ is **minimal** !

# The proof idea for $\mu\mathrm{vd}(\mathcal{MU}_{\delta=k}) \leq \mathrm{nM}(k)$

We use induction on the deficiency $k$:

1. Consider $F \in \mathcal{MU}_{\delta=k}$.
2. W.l.o.g. $F \in \mathcal{SMU}_{\delta=k}$.
3. Consider $v \in \mathrm{var}(F)$ realising $\mu\mathrm{vd}(F)$.
4. Consider $\varepsilon \in \{0, 1\}$.
5. Let $i$ be the degree of *literal* $\overline{v}/v$ in case of $\varepsilon = 0/1$.
6. So $F' := \langle v \to \varepsilon \rangle * F \in \mathcal{MU}_{\delta=k-i+1}$.
7. Consider a variable $w$ of minimal degree in $F'$.
8. We have $\mathrm{vd}_F(w) \leq \mathrm{vd}_{F'}(w) + i$.
9. Minimise over $\varepsilon$, maximise over $i$.

# Non-Mersenne numbers

For $k \in \mathbb{N}$ let $\mathrm{nM}(k) := 2$ if $k = 1$, while else

$$\mathrm{nM}(k) := \max_{i \in \{2,\ldots,k\}} \min(2 \cdot i, \mathrm{nM}(k - i + 1) + i).$$

The values $\mathrm{nM}(k)$ for $k = 1, \ldots, 26$ are

$$2$$
$$4, 5, 6$$
$$8, \ldots, 14$$
$$16, \ldots, 30$$

So the numbers $3, 7, 15, 31, \ldots, 2^m - 1, \ldots$ are skipped (and thus the name).

# The surplus

For a set $V$ of variables by $\boldsymbol{F[V]}$ we denote the restriction of $F$ to $V$:

- remove all clauses $C$ with $\mathrm{var}(C) \cap V = \emptyset$, and
- remove from the remaining clauses all literals with variables not in $V$.

The **surplus** $\boldsymbol{\sigma(F)} \in \mathbb{Z}$ is the minimal deficiency over all such non-trivial restrictions:

$$\sigma(F) := \min_{\emptyset \neq V \subseteq \mathrm{var}(F)} \delta(F[V]).$$

We have

1. $\sigma(F) \leq \delta(F)$
2. $\mu\mathrm{vd}(F) \geq \sigma(F) + 1$
3. $\sigma(F) \geq 1$ iff $F$ is matching lean (has no "matching autarky", a special case of autarkies).

# The generalisation





We can generalise and sharpen the MU-bound:

For all $F \in \mathcal{LEAN}$ we have $\mu\mathrm{vd}(F) \leq \mathrm{nM}(\sigma(F))$.

As already stated, this bound is sharp (to appear in [Kullmann and Zhao, 2011a]), even when considering the deficiency:

$$\forall k \in \mathbb{N} : \mu\mathrm{vd}(\mathcal{LEAN}_{\delta=k}) = \mathrm{nM}(k).$$

# How to find the autarky?

Consider a clause-set $F$ with deficiency $k$:

- If $k \leq 0$ then $F$ has a matching autarky (which can be found in polynomial time). So assume $k \geq 1$.
- If now every variable of $F$ occurs at least $\mathrm{nM}(k) + 1$ times, then $F$ has a non-trivial autarky.
- Can we find such an autarky in polynomial time?

The proof for the existence of the autarky seems only to offer a simple brute-force algorithm.

# Deficiency 6



The bound for $\mathcal{MU}$ however is not sharp, as shown in [Kullmann and Zhao, 2011a].

- We have $\mathrm{nM}(6) = 9$.
- By a careful study of the combinatorial situation we can prove
$$\mu\mathrm{vd}(\mathcal{MU}_{\delta=6}) = 8.$$

The proof

- exploits splitting
- and a rather precise knowledge about the structure of $\mathcal{MU}_{\delta=2}$ and (to a lesser degree) $\mathcal{MU}_{\delta=3}$.

# What follows

Using the general splitting structure, now

for all deficiencies $k = 2^m - m + 1$, $m \geq 3$, that is, besides 6 also $13, 28, 59, \dots$, we can show $\mu\text{vd}(\mathcal{MU}_{\delta=k}) = \text{nM}(k) - 1$.

These changes to nM by $-1$ are denoted by $\mathbf{nM_1}$.

# Deficiency 14



Generalising the idea, we can show that also for $k = 14$ we have to subtract 1, that is,

$$\mu\text{vd}(\mathcal{MU}_{\delta=14}) = \text{nM}(14) - 1 = 17.$$

And then again it follows that for all deficiencies $k = 2^m - m + 2$, $m \geq 4$, we have

$$\mu\text{vd}(\mathcal{MU}_{\delta=k}) = \text{nM}(k) - 1.$$

These further changes to nM by $-1$ are denoted by $\mathbf{nM_2}$.

# And so on

Generalising further we arrive at the improved bounds $\mathrm{nM}_3, \mathrm{nM}_4, \dots,$ defined as follows:

1. $\mathrm{nM}_0 := \mathrm{nM}$
2. for $i > 0$ let $\mathrm{nM}_i$ be $\mathrm{nM}_{i-1}$, except that for $k = 2^m - m + i$, $m \geq 2 + i$, we set $\mathrm{nM}_i(k) := \mathrm{nM}_{i-1}(k) - 1 = \mathrm{nM}(k) - 1$.

And

$$\mathbf{nM}_{\omega}(\boldsymbol{k}) := \min_{i \in \mathbb{N}} \mathrm{nM}_i(k).$$

We have for all $k \in \mathbb{N}$

$$\mu\mathrm{vd}(\mathcal{MU}_{\delta=k}) \leq \mathrm{nM}_{\omega}(k)$$

where $\mathrm{nM}_{\omega}(k) \geq \mathrm{nM}(k) - 1$.

# That's not it



We originally thought that would be it, but, alas, we have $\mathrm{nM}_\omega(15) = \mathrm{nM}(k) = 19$, while we can show, using new ideas

$$\mu\mathrm{vd}(\mathcal{MU}_{\delta=15}) = \mathrm{nM}(15) - 1.$$

- This time, this improvement has no further consequences for higher $k$.
- We call the obtained new function $\mathrm{nM}_{\omega+1}$.
- "Likely" that's not it?!

# Summary

I We study the minimal var-degree of lean clause-sets in the deficiency.

II For this class we provide the sharp bound $\mathrm{nM}(k)$.

III If this bound is not respected, then there exists an autarky — can we find it efficiently?

IV For $\mathcal{MU}$ the situation is more complicated, and we provided an infinite sequence of improvements.

V Can we also here determine the precise bound, or does it get more and more complicated?

# Bibliography I


Kleine Büning, H. and Kullmann, O. (2009).
Minimal unsatisfiability and autarkies.
In Biere, A., Heule, M. J., van Maaren, H., and Walsh, T., editors, *Handbook of Satisfiability*, volume 185 of *Frontiers in Artificial Intelligence and Applications*, chapter 11, pages 339–401. IOS Press.

Kullmann, O. and Zhao, X. (2011a).
Bounds for variables with few occurrences in conjunctive normal forms.
In preparation.

# Bibliography II


Kullmann, O. and Zhao, X. (2011b).
On variables with few occurrences in conjunctive normal forms.
In Simon, L. and Sakallah, K., editors, *Theory and Applications of Satisfiability Testing - SAT 2011*, volume LNCS 6695 of *Lecture Notes in Computer Science*, pages 33–46. Springer.
ISBN-13 978-3-642-14185-0.

End