Minimal unsatisfiability and deficiency: recent developments

Oliver Kullmann

Computer Science Department
Swansea University

International Workshop on Graph Structure and Satisfiability Testing
July 4, 2016

MU

- $\mathcal{MU}$ is the set of clause-sets, which are unsatisfiable, while removal of any clause renders them satisfiable.
- $n(F)$ is the number of (occurring) variables.
- $c(F)$ is the number of clauses.
- $\delta(F) := c(F) - n(F) \in \mathbb{Z}$ is the **deficiency**.

See Handbook Chapter Kleine Büning and Kullmann [8].

“Tarsi’s Lemma”

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

- Best known proof Aharoni and Linial [1].
- For an overview see the introduction of [14].

Deficiency for MU yields a complexity parameter.

- MU decision poly-time for fixed k (Fleischner, Kullmann, and Szeider [3]).
- Indeed fpt (Szeider [16]).

Degrees

The most basic information about MU is given by some knowledge on the *degrees*.

Literal degrees:

$$\mathrm{ld}_F(x) := |\{C \in F : x \in C\}|$$

Variable degrees:

$$\mathrm{vd}_F(v) := \mathrm{ld}_F(v) + \mathrm{ld}_F(\overline{v}).$$

MU(1) I

$$\mathcal{MU}_{\delta=1} = \{F \in \mathcal{MU} : \delta(F) = 1\}$$

- These are nice formulas, with a surprising number of applications.
- In the SAT world, classification due to Aharoni and Linial [1], Davydov, Davydova, and Kleine Büning [2].
- Indeed, independently equivalent classifications in different areas have been obtained; see [14].

Characterisation becomes MUCH easier, once you know that for all $F \in \mathcal{MU}_{\delta=1}$, $n(F) \neq 0$, there exists $v \in \mathrm{var}(F)$ with $\mathrm{vd}_F(v) \leq 2$.

We express this as

$$\mathrm{VDM}(1) = 2.$$

MU(1) II

Of course, relevant open problems!

- For example concerning the *uniform* elements of $\mathcal{MU}_{\delta=1}$; Hoory and Szeider [6], Gebauer, Szabo, and Tardos [4].
- Uniformity (constant clause-length) features a lot in hypergraph theory, while we work mostly in the unrestricted setting.

MU(2)

$$\mathcal{MU}_{\delta=2} = \{F \in \mathcal{MU} : \delta(F) = 2\}$$

- The basic characterisation is due to Kleine Büning [7].
- This concerns *nonsingular* elements of $\mathcal{MU}_{\delta=2}$ — every variable occurs at least twice positively as well as negatively.

The main open question here is:

Extend the generalisation to all of $\mathcal{MU}_{\delta=2}$ —
in a sense fusing the characterisations obtained for $\delta = 1, 2$.

This is needed for a better understanding of higher deficiencies.

Again, a fundamental step is to to show
$\forall\, F \in \mathcal{MU}_{\delta=2}, n(F) \neq 0\ \exists v \in \mathrm{var}(F) : \mathrm{vd}_F(v) \leq 4.$

I.e., $\mathrm{VDM}(2) = 4$

Min-var-degree

In general (F a clause-set, $\mathcal{C}$ a class of clause-sets):

$$\begin{aligned}
\mu\mathrm{vd}(F) &:= \min_{v \in \mathrm{var}(F)} \mathrm{vd}_F(v) \\
\mu\mathrm{vd}(\mathcal{C}) &:= \max_{F \in \mathcal{C}} \mu\mathrm{vd}(F) \\
\mathrm{VDM}(k) &:= \mu\mathrm{vd}(\mathcal{MU}_{\delta=k}).
\end{aligned}$$

In [9] the fundamental bound

$$\forall\, k \geq 1 : \mathrm{VDM}(k) \leq 2k$$

was shown.

Improving the bound

In [12] the upper bound $\text{VDM}(k) \leq 2k$ was improved to

$$\text{VDM}(k) \leq 1 + k + \log_2(k).$$

- Indeed a precise number-theoretical function $\text{nM}(k)$ yields the upper bound.
- This upper bound is not sharp, and the first deficiency needing a correction is $k = 6$.

The main open problem here is the precise determination of $\text{VDM}(k)$.

The sharpenings we produced unearth interesting aspects of MU; see [14] for further information.

LEAN

Indeed, the upper bound $nM(k)$ is sharp for **lean** clause-sets of deficiency k.

LEAN means: no non-trivial autarkies.

This leads to interesting algorithmic consequences:

If the bound is violated, then there *exists* a non-trivial autarky.

- Indeed, the effect of the autarky reduction can be simulated.
- But to find the autarky itself (the witness) is an open problem!
- See [14] for further information.

Full clauses

An interesting combinatorial quantity for a clause-set F is the *number of full clauses*:

$$\mathrm{fc}(F) := |\{C \in F : \mathrm{var}(C) = \mathrm{var}(F)\}| \in \mathbb{N}_0.$$

We have

$$\mathrm{fc}(F) \le \mu\mathrm{vd}(F).$$

So maximising the number of full clauses yields *lower bounds* on VDM(k).

Let FCM(k) be the maximum of fc(F) for $F \in \mathcal{MU}_{\delta=k}$.

Thus FCM $\le$ VDM.

Hitting clause-sets

Indeed it helps a lot to consider *hitting clause-sets* here:

Every two clauses have a clash.

If a hitting clause-set is unsatisfiable, it is automatically MU.

$\mathrm{VDH}(k), \mathrm{FCH}(k)$ denote the
maximal min-var-degree resp. number of full clauses
for hitting MU.

- We conjecture $\mathrm{VDH} = \mathrm{VDM}$.
- But definitely only $\mathrm{FCH} \leq \mathrm{FCM}$.

Meta-Fibonacci

We show

$$S_2 \leq \text{FCH}$$

for the number-theoretic function S_2.

And indeed we conjecture equality.

- Interesting recursion-theoretic phenomena show up.
- Belong to the field of “meta-Fibonacci” functions (nested recursive calls), as introduced by Hofstadter [5].

The four fundamental quantities

To summarise the first part of the talk:

The quantities $\mathrm{VDM}(k), \mathrm{FCM}(k), \mathrm{VDH}(k), \mathrm{FCH}(k)$
seem interesting beasts: offering a lot of depth
— and good attack points!

The *precise* quantities matter here, and relevant number-theoretical functions appear.

It's part of the fundamental **Finite Patterns Conjecture**:

For every k, $\mathcal{MU}_{\delta=k}$ can be characterised by
finitely many patterns.

The next frontier is $\mathcal{MU}_{\delta=3}$.

DP-reduction

DP-reduction

$$F \rightsquigarrow \mathrm{DP}_v(F)$$

replaces all clauses containing variable v by their (non-tautological) resolvents.

- It it "commutative" ([10, 11]).
- Maintains the hitting property.
- But in general does not maintain MU.

Singular variables

Singular variables occur in one sign only once.

- *Singular DP-reduction* behaves well also for MU.
- Full reduction via singular DP-reduction establishes some kind of "normal form", via confluence and weaker forms ([13]).

The details are intriguing, and many open problems.

Also the other direction, *singular extension*, is of relevance:

- First one characterises the non-singular elements.
- Then one studies their singular extensions.

Clause-factors

In [15] we introduced a new concept for analysing MUs:

Definition

A clause-set F is called a **clause-factor** if F is logically equivalent to a single clause. F is called a **clause-factor of** F' if F is a clause-factor and $F \subseteq F'$.

It is easy to show:

Lemma

F is logically equivalent to a clause C iff the following two conditions hold:

1. $\forall D \in F : C \subseteq D$.
2. $\{D \setminus C : D \in F\}$ *is unsatisfiable.*

Clause-irreducibility

If we have a clause-factor F of F', then we can "factorise" F' into

- the "residue" $\{D \setminus C : D \in F\}$,
- and the "cofactor" $(F' \setminus F) \cup \{C\}$.

This becomes trivial iff $F = \{C\}$ for $F = F'$.

Definition

If a clause-set has no trivial clause-factor, then it is called **clause-irreducible**.

For unsatisfiable hitting clause-sets,

- clause-irreducibility is surprisingly powerful,
- with good structural properties.

It seems an essential tool for classification, to reduce complexity.

Summary and outlook

I Studying the “four fundamental quantities” reveals surprising structures — if you go for the exact determination.

II Reductions are an important tool for understanding MU: first you concentrate on understanding (only) reduced cases, and then you extend.

III The reductions have good properties, and likely there is much more to come.

End

(references on the remaining slides).

For my papers see
http://cs.swan.ac.uk/~csoliver/papers.html.

Bibliography I

[1] Ron Aharoni and Nathan Linial. Minimal non-two-colorable hypergraphs and minimal unsatisfiable formulas. *Journal of Combinatorial Theory, Series A*, 43(2):196–204, November 1986. doi:10.1016/0097-3165(86)90060-9.

[2] Gennady Davydov, Inna Davydova, and Hans Kleine Büning. An efficient algorithm for the minimal unsatisfiability problem for a subclass of CNF. *Annals of Mathematics and Artificial Intelligence*, 23(3-4):229–245, 1998. doi:10.1023/A:1018924526592.

[3] Herbert Fleischner, Oliver Kullmann, and Stefan Szeider. Polynomial–time recognition of minimal unsatisfiable formulas with fixed clause–variable difference. *Theoretical Computer Science*, 289(1):503–516, November 2002. doi:10.1016/S0304-3975(01)00337-1.

Bibliography II

[4] Heidi Gebauer, Tibor Szabo, and Gabor Tardos. The Local Lemma is asymptotically tight for SAT. Technical Report arXiv:1006.0744v3 [math.CO], arXiv.org, April 2016. URL http://arxiv.org/abs/1006.0744.

[5] Douglas R. Hofstadter. *Gödel, Escher, Bach: An eternal golden braid*. Basic Books, 1979. ISBN 0140055797. URL http://www.physixfan.com/wp-content/files/GEBen.pdf. Pdf version with 801 pages, md5sum="0cb32e8ea5dd2485f63842f5acffb3f0 GEBen.pdf".

[6] Shlomo Hoory and Stefan Szeider. Computing unsatisfiable *k*-SAT instances with few occurrences per variable. *Theoretical Computer Science*, 337(1-3):347–359, June 2005. doi:10.1016/j.tcs.2005.02.004.

Bibliography III

[7] Hans Kleine Büning. On subclasses of minimal unsatisfiable formulas. *Discrete Applied Mathematics*, 107(1-3):83–98, 2000. doi:10.1016/S0166-218X(00)00245-6.

[8] Hans Kleine Büning and Oliver Kullmann. Minimal unsatisfiability and autarkies. In Armin Biere, Marijn J.H. Heule, Hans van Maaren, and Toby Walsh, editors, *Handbook of Satisfiability*, volume 185 of *Frontiers in Artificial Intelligence and Applications*, chapter 11, pages 339–401. IOS Press, February 2009. ISBN 978-1-58603-929-5. doi:10.3233/978-1-58603-929-5-339.

[9] Oliver Kullmann. An application of matroid theory to the SAT problem. In *Proceedings of the 15th Annual IEEE Conference on Computational Complexity*, pages 116–124, July 2000. doi:10.1109/CCC.2000.856741. See also TR00-018, Electronic Colloquium on Computational Complexity (ECCC), March 2000.

Bibliography IV

[10] Oliver Kullmann and Horst Luckhardt. Deciding propositional tautologies: Algorithms and their complexity. Preprint, 82 pages, January 1997. URL http://cs.swan.ac.uk/~csoliver/papers.html#Pre199799.

[11] Oliver Kullmann and Horst Luckhardt. Algorithms for SAT/TAUT decision based on various measures. Preprint, 71 pages, February 1999. URL http://cs.swan.ac.uk/~csoliver/papers.html#Pre199799.

[12] Oliver Kullmann and Xishun Zhao. On variables with few occurrences in conjunctive normal forms. In Laurent Simon and Karem Sakallah, editors, *Theory and Applications of Satisfiability Testing - SAT 2011*, volume 6695 of *Lecture Notes in Computer Science*, pages 33–46. Springer, 2011. ISBN 978-3-642-14185-0. doi:10.1007/978-3-642-21581-0_5.

Bibliography V

[13] Oliver Kullmann and Xishun Zhao. On Davis-Putnam reductions for minimally unsatisfiable clause-sets. *Theoretical Computer Science*, 492:70–87, June 2013. doi:10.1016/j.tcs.2013.04.020.

[14] Oliver Kullmann and Xishun Zhao. Bounds for variables with few occurrences in conjunctive normal forms. Technical Report arXiv:1408.0629 [math.CO], arXiv, April 2016. URL http://arxiv.org/abs/1408.0629.

[15] Oliver Kullmann and Xishun Zhao. Unsatisfiable hitting clause-sets with three more clauses than variables. Technical Report arXiv:1604.01288v1 [cs.DM], arXiv, April 2016. URL http://arxiv.org/abs/1604.01288.

[16] Stefan Szeider. Minimal unsatisfiable formulas with bounded clause-variable difference are fixed-parameter tractable. *Journal of Computer and System Sciences*, 69(4):656–674, December 2004. doi:10.1016/j.jcss.2004.04.009.