The combinatorics of Minimal Unsatisfiability

Oliver Kullmann

Computer Science Department
Swansea University
http://cs.swan.ac.uk/~csoliver/

Dagstuhl, November 20, 2012
SAT Interactions
Understanding unsatisfiability

We think it is actually possible to “understand unsatisfiability”, that is, minimal unsatisfiability can be reduced to basic, intuitive patterns.
Outline

1. Introduction
2. Some history
3. Background
4. Saturation
5. Classification of $MU_{\delta=k}$ (small $k$)
6. Understanding unsatisfiability
7. Singular DP-reduction
8. Infrequent variables
9. Conclusion
Aharoni and Linial [1] showed “Tarsi’s Lemma”, that for minimally unsatisfiable CNF $F$ (i.e., $F \in \mathcal{MU}$) we have

$$c(F) \geq n(F) + 1.$$  

- They also characterised “strongly minimally unsatisfiable clause-sets” with $c(F) = n(F) + 1$, i.e., the class $\mathcal{SMU}_{\delta=1} \subset \mathcal{MU}_{\delta=1}$.
- This was revived by Kleine Büning [6], Davydov, Davydova, and Kleine Büning [2], who achieved recognition of $\mathcal{MU}_{\delta=1} = \{ F \in \mathcal{MU} : c(F) = n(F) + 1 \}$ in linear time, and started investigations into the general cases $c(F) = n(F) + k$ for $k \in \mathbb{N}$.
- The main conjecture was that recognition for constant $k$ is in polynomial time.
Remarks on autarkies

1. The conjecture was proven independently by Fleischner/Szeider and by K; see the overview article Kleine Büning and Kullmann [8] in the Handbook of Satisfiability.

2. For me this was an important starting point of systematic investigation into the theory of autarkies (originating from the Monien/Speckenmeyer article with the $k$-SAT bound; see the handbook article).

Autarkies serve here as a tool for detecting (and eliminating) degeneracies.

Autarky systems allow specialised notions of autarkies, for example with associated poly-time algorithms.

Not only Tarsi’s Lemma can be understood in this way (using “matching autarkies”), but also the solution of the Polya Problem by Seymour et al (using “balanced linear autarkies”; see Robertson, Seymour, and Thomas [17], Kullmann [12]).
The deficiency

An important quantity for all these considerations is given by the notion of “deficiency” introduced in Franco and Gelder [5]:

$$\delta(F) := c(F) - n(F)$$

(thus Tarsi’s Lemma states $\delta(F) \geq 1$ for $F \in \mathcal{MU}$).

- For generalised notions of clause-sets (including hypergraphs and non-boolean clause-sets) the notion of deficiency can be adapted.
- The basic phenomenon is that by some (poly-time) autarky reduction we can establish a minimum value for the deficiency (Tarsi’s Lemma for all CNF, and the generalisation of Fisher’s Inequality for (critical) hypergraph 2-colouring from Seymour’s MSc thesis Seymour [18]).
Deficiency continued

- A fundamental question is then whether the layers (constant deficiency) can be recognised in poly-time.
- This was established for the ordinary deficiency (as mentioned above).
- It is an open problem for hypergraph deficiency (the base case can be obtained by autarky techniques and constructivisation from the solution of Polya’s problem; see Kullmann [12]).
- More ambitious programs ask for some form of classifications of the layers.
- For (very small) deficiencies there are results.

Developing general tools for the classification of $\mathcal{MU}(k)$ is the research program Xishun Zhao and I are pursuing.
Clause-sets

- Literals are variables $v$ and their complements $\overline{v}$.
- A clause $C$ is a finite and clash-free set of literals, i.e., $C \cap \overline{C} = \emptyset$.
- A **clause-set** is a finite set of clauses.
- The set of all clause-sets is $\mathcal{CLS}$.

For example

$$\mathcal{F}_2 = \{ \{ v_1, v_2 \}, \{ \overline{v_1}, \overline{v_2} \}, \{ \overline{v_1}, v_2 \}, \{ \overline{v_2}, v_1 \} \}$$

is a clause-set (minimally unsatisfiable, deficiency 2).

Remark: Clause-sets are considered as (precise) combinatorial objects, as generalised hypergraphs.
Applying partial assignments

The application of a partial assignment $\varphi \in \mathcal{P}ASS$ to a clause-set $F \in \mathcal{CLS}$ is denoted by

$$\varphi \ast F \in \mathcal{CLS}.$$ 

Satisfied clauses are removed, then falsified literals. A special clause-set is $\top := \emptyset$, a special clause is $\bot := \emptyset$.

- $F$ is **satisfiable** iff there is $\varphi \in \mathcal{P}ASS$ with $\varphi \ast F = \top$.
- $\top$ is satisfiable.
- $\{\bot\}$ is unsatisfiable.
- More generally, every $F$ with $\bot \in F$ is unsatisfiable.

$$\mathcal{CLS} = \mathcal{SAT} \cup \mathcal{USAT}.$$
Variable degrees

- The set of variables of $F$ is denoted by $\text{var}(F)$.
- The **degree** of a variable $v$ in a clause-set $F$ is the number of its occurrences, denoted by

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

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

$$\mu_{\text{vd}}(F) := \min_{v \in \text{var}(F)} \text{vd}_F(v).$$

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

- An **autarky** for a clause-set $F$ is a partial assignment $\varphi$ which satisfies every clause $C \in F$ it touches (i.e., $\text{var}(\varphi) \cap \text{var}(C) \neq \emptyset$).
- An autarky $\varphi$ for $F$ is **trivial** if $\text{var}(\varphi) \cap \text{var}(F) = \emptyset$.

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

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

For a minimally unsatisfiable clause-set $F$ we have

$$\delta(F) \geq 1,$$

where

- $c(F) := |F|$ is the number of clauses
- $n(F) := |\text{var}(F)|$ is the number of variables,
- $\delta(F) := c(F) - n(F)$ is the déficiency.

This was first shown in Aharoni and Linial [1].

The proof of Aharoni and Linial [1] can be generalised in terms of “matching autarkies” (Kullmann [10]), using the notion of “deficiency”. This yields

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

Recall

$$\mathcal{MU} \subset \text{LEAN}$$

Thus

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

We use here $\mathcal{MU}(k) := \{F \in \mathcal{MU} : \delta(F) = k\}$. 
Resolution and DP-reduction

Clauses $C, D$ are **resolvable** if $C \cap \overline{D} = \{x\}$:

$$C \diamond D := (C \setminus \{x\}) \cup (D \setminus \{\overline{x}\}).$$

**DP-reduction** (or “variable elimination”):

$$\text{DP}_v(F) := \{ C \in F : v \notin \text{var}(C) \} \cup \{ C \diamond D : C, D \in F \land C \cap \overline{D} = \{v\} \}.$$  

$\text{DP}_v(F)$ is semantically the existential quantification of $F$ by $v$:

$$\text{DP}_v(F) \iff \exists v : F.$$
Singular variables

A variable $\nu$ is **singular** for $F$ if
- in one sign it occurs only once,
- while in the other sign it occurs at least once.

**Singular DP-reduction** ("sDP-reduction") is $F \sim DP_{\nu}(F)$ for a singular variable.

sDP-reduction decreases the number of clauses at least by one.
Fliti and Reynaud [4] introduced saturated minimally unsatisfiable clause-sets:

- These are minimally unsatisfiable $F$ such that addition of any literal to any clause renders $F$ satisfiable.
- Every minimally unsatisfiable $F$ can be saturated, without changing the number of clauses or the number of variables.
**MU and SMU**

**Minimal** unsatisfiability: no clause can be removed.

\[ \mathcal{MU} = \{ F \in \text{USAT} \mid \forall C \in F : F \setminus \{ C \} \in \text{SAT} \} . \]

**Saturated** minimal unsatisfiability: no literal can be added.

\[ \mathcal{SMU} := \{ F \in \mathcal{MU} \mid \forall C \in F \forall C' \supset C : (F \setminus \{ C \}) \cup \{ C' \} \in \text{SAT} \} . \]

**Lemma**

\( F \in \mathcal{SMU} \) iff for all \( v \in \text{var}(F) \) and \( \varepsilon \in \{0, 1\} \) holds \( \langle v \rightarrow \varepsilon \rangle \ast F \in \mathcal{MU} \).
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 $SMU \subset MU$.
4. Every $F \in MU$ can be saturated, yielding $F' \in SMU$ with $n(F') = n(F)$ and $c(F') = c(F)$ (thus $\delta(F') = \delta(F)$).
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 $v \in \text{var}(F') = \text{var}(F)$,
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 $\langle v \rightarrow \varepsilon \rangle \ast F$?

This can not happen if the degree of $v$ is minimal!
A simple proof of Tarsi’s Lemma

We prove $\delta(F) \geq 1$ for $F \in \mathcal{MU}$ by induction on $n(F)$:

$n(F) = 0$: $F = \{\bot\}$, thus $\delta(F) = 1$. √

$n(F) > 0$: Saturate $F$, obtain $F' \in S\mathcal{MU}$ with $\delta(F') = \delta(F)$.

Consider $v \in \text{var}(F)$ with minimal variable degree, that is, $vd_F(v) = \text{ld}_F(v) + \text{ld}_F(\overline{v})$ minimal (i.e., $vd_F(v) = \mu vd(F)$).

Consider $\langle v \rightarrow 0 \rangle \ast F \in \mathcal{MU}$.

We have $\text{var}(\langle v \rightarrow 0 \rangle \ast F) = \text{var}(F) \setminus \{v\}$, since $v$ occurs minimally often.

So $\delta(\langle v \rightarrow 0 \rangle \ast F) = \delta(F) - \text{ld}(\overline{v}) + 1 \leq \delta(F)$.

The induction hypothesis is $\delta(\langle v \rightarrow 0 \rangle \ast F) \geq 1$. √
Aharoni and Linial [1] characterised $SMU_{\delta=1} \subset MU_{\delta=1}$.

Davydov et al. [2] characterised $MU_{\delta=1}$ in terms of matrices.

Kullmann [9] added more information:
1. the tree-characterisation of $SMU_{\delta=1}$;
2. $MU_{\delta=1}$ is obtained from $SMU_{\delta=1}$ by removing literal occurrences from the clauses such that no pure literal is created.

Using that the hermitian defect is an upper bound for the deficiency, in Kullmann [11] it is shown that $SMU_{\delta=1}$ is the class of all unsatisfiable clause-sets where every two (different) clauses clash in exactly one literal (corresponding to the DNF tautologies where the hamming distance between the cubes corresponding to two (different) terms is always 1).

In Kullmann [14] these results are generalised to non-boolean clause-sets.
Many more properties of $\mathcal{M}U_{\delta=1}$ are known, and more and more are discovered. We need an ever more precise knowledge on $\mathcal{M}U_{\delta=1}$ in order to characterise $\mathcal{M}U_{\delta=k}$ for higher $k$. 
There is always a 1-singular variable

- The basic fact on $\mathcal{M}U_{\delta=1}$ is that every $F \in \mathcal{M}U_{\delta=1}$, $F \neq \{\bot\}$, contains a variable occurring positively and negatively exactly once.
- A simple proof is given later.
- This is a special case of a singular variable (namely a “1-singular variable”), which more generally only requires that the variable occurs in one polarity only once.
- By DP-reduction such variables can be eliminated, without leaving $\mathcal{M}U$ and without changing the deficiency.
Classification of \( \mathcal{M}U_{\delta=k} \) (small \( k \))

**MU(1)**

1. \( \{ \{ a, b \}, \{ a, \bar{b} \}, \{ \bar{a}, c \}, \{ \bar{a}, \bar{c} \} \} \in S\mathcal{M}U(1) \). Singular: \( b, c \).
2. \( \{ \{ a, b \}, \{ \bar{b} \}, \{ \bar{a}, c \}, \{ \bar{a}, \bar{c} \} \} \in \mathcal{M}U(1) \). Singular: all.
3. \( \{ \{ a, b \}, \{ \bar{b} \}, \{ \bar{a}, c \}, \{ \bar{c} \} \} \in \mathcal{M}\mathcal{M}U(1) \) (“marginal”). 1-singular: all.
Eliminating trivialities

We consider sDP-reduction as eliminating “trivialities”.

- So all of $\mathcal{MU}(1)$ boils down to $\{\bot\}$.
- Intuitively one can understand applying sDP as removing some “$\mathcal{MU}(1)$-hunch”.

A clause-set $F$ is called **non-singular** if it does not contain singular variables.

- The set of singular variables of $F$ is $\text{var}_s(F) \subseteq \text{var}(F)$.
- So $F$ is nonsingular iff $\text{var}_s(F) = \emptyset$.

$$\mathcal{MU}' := \{F \in \mathcal{MU} : \text{var}_s(F) = \emptyset\}.$$
A breakthrough was achieved by Kleine Büning [7] (2000). The elements of $\mathcal{MU}'(2) = S\mathcal{MU}(2)$ are precisely the following clause-sets for $n \geq 2$:

\[
\begin{align*}
    x_1 &\rightarrow x_2, \ x_2 \rightarrow x_3, \ldots, \ x_{n-1} \rightarrow x_n, \ x_n \rightarrow x_1 \\
    \{x_1, \ldots, x_n\}, \\
    \{\overline{x_1}, \ldots, \overline{x_n}\}.
\end{align*}
\]

That is, one cycle, with opposed forced “directions” ($n = 6$):
The knowledge on higher deficiencies is very scarce yet.

- Decision is possible, and we even have fixed-parameter tractability (Fleischner, Kullmann, and Szeider [3], Kullmann [9], Szeider [19], [13, 14]).
- Such decision procedures work by splitting (thereby reducing the deficiency (in some sense)).
- Also the main methods to gain knowledge on $\mathcal{MU}_{\delta\leq k}$ work by induction and splitting.
- $\mathcal{MU}_{\delta=1}$ seems a special case, a “degeneration”.
- A major complication is that splitting may introduce singular variables, and so we need to know also the singular elements of $\mathcal{MU}_{\delta<k}$ rather well.
For unsatisfiable $F \in \text{USAT}$

we want to “understand” its unsatisfiability.

We want to SEE it.

Each $F' \subseteq F$ with $F' \in \text{MU}$ contains one explanation:

- The additional clauses in $F \setminus F'$ can make the contradiction of $F'$ more easily accessible, but do not contribute “another reason”.
- Different $F'$ provide different explanations.

So we have to “explain” unsatisfiability for $F \in \text{MU}$.

Explain $\text{MU}'(k)$ for $k = 1, 2, 3, \ldots$. 

Oliver Kullmann (Swansea)
The Classification Conjecture

**Conjecture** *For every* $k \in \mathbb{N}$ *there are finitely many “patterns”, which explain* $\mathcal{M}U'(k)$ *completely.*

1. $\mathcal{M}U'(1) = \{ \bot \}$
2. $\mathcal{M}U'(2) =$ cycles of length $n \geq 2$ plus “at least one variable is false” plus “at least one variable is true”.
3. $\mathcal{M}U'(3)$: in preparation.
For $F \in \mathcal{MU}$ let

$$sDP(F) := \{ F' \in \mathcal{MU} : F \xrightarrow{sDP} F' \}.$$ 

We want to understand $F \in \mathcal{MU}(k)$. For that we consider $sDP(F)$. Note $sDP(F) \subset \mathcal{MU}(k)$.

Easiest is $|sDP(F)| = 1$ — when does this hold?
The “singularity index”

**Theorem**

For $F \in \mathcal{MU}$ and $F', F'' \in sDP(F)$ we have $n(F') = n(F'')$.

This allows us to define the **singularity index** $si(F) \in \mathbb{N}_0$ as $si(F) := n(F) - n(F')$ for some $F' \in sDP(F)$.

**Corollary**

If $F \in \mathcal{MU}(2)$, then for $F', F'' \in sDP(F)$ we have $F' \cong F''$.

Here $F' \cong F''$ denotes isomorphism.
Confluence

**Theorem**

If $F \in SMU$, then $|sDP(F)| = 1$. 
Confluence modulo isomorphism

**Theorem**
If for $F \in \mathcal{MU}$ we have $sDP(F) \subseteq SMU$, then for $F', F'' \in sDP(F)$ we have $F' \cong F''$.

**Corollary**
If $F \in \mathcal{MU}(2)$, then for $F', F'' \in sDP(F)$ we have $F' \cong F''$. 
We didn’t make full use of the proof that \( \delta(F) \geq 1 \):

1. For a variable with \( \nu vd(\nu) = \mu vd(F) \) we have
   \[ \delta(\langle \nu \rightarrow 0 \rangle \ast F) = \delta(F) - \text{ld}(\nu) + 1. \]
2. Since \( \delta(\langle \nu \rightarrow 0 \rangle \ast F) \geq 1 \), we thus have \( \text{ld}(\nu) \leq \delta(F) \).
3. The same for \( \nu \rightarrow 1 \).

We get thus

\[ \mu vd(F) \leq 2\delta(F) \]

for \( F \in \mathcal{MU} \).
Consider $F \in \mathcal{MU}_{\delta=3}$.

Assume $\mu \vd(F) = 2 \cdot 3 = 6$, and consider $v \in \text{var}(F)$ with $\vd(v) = 6$.

1. Thus $\text{Id}(v) = \text{Id}(\overline{v}) = 3$.
2. Consider $F' := \langle v \rightarrow 0 \rangle \ast F$.
3. We have $\delta(F') = 3 - 3 + 1 = 1$.
4. W.l.o.g. $F \in \mathcal{SMU}$, and thus $F' \in \mathcal{MU}$.
5. So there exists $v' \in F'$ with $\vd_{F'}(v') = 2$.
6. Thus $\vd_{F}(v') \leq 2 + 3 = 5$, contradicting the assumption!

It follows $\mu \vd(F) \leq 5$. 
A general upper bound for $\MU$

In Kullmann and Zhao [15] we show for $F \in \MU$:

$$\muvd(F) \leq nM(\delta(F)),$$

where $nM : \mathbb{N} \rightarrow \mathbb{N}$ can be defined as

$$nM(k) := k + \lfloor \log_2(k + 1 + \lfloor \log_2(k + 1) \rfloor) \rfloor \leq k + 1 + \log_2(k).$$
Non-Mersenne numbers

For \( k \in \mathbb{N} \) we have \( n_M(k) = 2 \) if \( k = 1 \), while else

\[
n_M(k) = \max_{i \in \{2, \ldots, k\}} \min(2 \cdot i, n_M(k - i + 1) + i).
\]

The values \( n_M(k) \) for \( k = 1, \ldots, 26 \) are

\[
\begin{align*}
2 \\
4, 5, 6 \\
8, \ldots, 14 \\
16, \ldots, 30
\end{align*}
\]

So the numbers 3, 7, 15, 31, \ldots, \( 2^m - 1 \), \ldots are skipped (and thus the name).
Main results on variable degrees

For a class \( C \) of clause-sets let

\[
\mu_{vd}(C) := \sup_{F \in C} \mu_{vd}(F).
\]

The main results are:

- For all \( k \in \mathbb{N} \) we have \( \mu_{vd}(\mathcal{EAN}_{\delta=k}) = nM(k) \).
- We introduce
  \[
  nM > nM_1 > nM_2 > \cdots > nM_\omega > nM_{\omega+1}
  \]
  with
  \[
  \mu_{vd}(\mathcal{MU}_{\delta=k}) \leq nM_{\omega+1}(k).
  \]
  ("\( f(k) > g(k) \)" means here that there is at least one argument \( k \) with \( f(k) > g(k) \).)

(\( f \) and \( g \) are functions, and \( \omega \) is the first infinite ordinal.)
The proof idea for $\mu \text{vd}(\mathcal{M}U_{\delta=k}) \leq n\text{M}(k)$

We use induction on the deficiency $k$:

1. Consider $F \in \mathcal{M}U_{\delta=k}$.
2. W.l.o.g. $F \in SMU_{\delta=k}$.
3. Consider $v \in \text{var}(F)$ realising $\mu \text{vd}(F)$.
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 $F' := \langle v \rightarrow \varepsilon \rangle * F \in \mathcal{M}U_{\delta=k-i+1}$.
7. Consider a variable $w$ of minimal degree in $F'$.
8. We have $\text{vd}_F(w) \leq \text{vd}_{F'}(w) + i$.
9. Minimise over $\varepsilon$, maximise over $i$. 

Oliver Kullmann (Swansea)
Combinatorics of MU
Dagstuhl 2012 39 / 44
The bound for $\mathcal{M}U \leq nM(6)$ is not sharp, as shown in [16].

- We have $nM(6) = 9$.
- By a careful study of the combinatorial situation we can prove

$$\mu vd(\mathcal{M}U_{\delta=6}) = 8.$$

The proof

- exploits splitting
- and a rather precise knowledge about the structure of $\mathcal{M}U_{\delta=2}$ and (to a lesser degree) $\mathcal{M}U_{\delta=3}$. 
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, \ldots, we can show $\mu v d(MU_{\delta=0}) = nM(k) - 1$.

These changes to $nM$ by $-1$ are denoted by $nM_1$. 
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

   \[
   \mu vd(F) > k + 1 + \log_2(k) \geq nM(k)
   \]

   there exist a non-trivial autarky — can we **find it** in polynomial time?

2. Is the bound \(\mu vd(\mathcal{MU}_{\delta=k}) \leq nM_{\omega+1}(k)\) sharp, or are there more (and more) improvements?

3. We conjecture \(\mu vd(\mathcal{MU}_{\delta=k}) \geq nM(k) - 1\) for all \(k\).
Summary and outlook

I Quite some remarkable structure has been discovered on the structure of $\mathcal{MU}(k)$.

II We believe there is much more to come!

III The major goal is to first precisely state and then to prove the Finitely-many-patterns conjecture.

IV For the special case of unsatisfiable hitting clause-sets we can precisely state the finiteness-conjecture; more in the talk of Xishun.
End

(references on the remaining slides).


Conclusion

Bibliography II


Bibliography IV


