Mycielski graphs and PR proofs

Emre Yolcu Xinyu Wu Marijn J. H. Heule

Carnegie Mellon University

23rd International Conference on
Theory and Applications of Satisfiability Testing

eyolcu@cs.cmu.edu

Problem

- **Mycielski graphs:**
  Triangle-free graphs M_k with arbitrarily high chromatic number.
  M_k is not properly colorable using $k-1$ colors.

Problem

- **Mycielski graphs:**
  Triangle-free graphs M_k with arbitrarily high chromatic number.
  M_k is not properly colorable using $k - 1$ colors.
- **Propagation redundancy (**PR**):**
  "Interference-based" propositional proof system.
  Generalizes the commonly-used DRAT.

Problem

- **Mycielski graphs:**
  Triangle-free graphs M_k with arbitrarily high chromatic number.
  M_k is not properly colorable using $k-1$ colors.

- **Propagation redundancy (**PR**):**
  "Interference-based" propositional proof system.
  Generalizes the commonly-used DRAT.

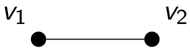
Asked by Donald Knuth:

> *For a family of formulas encoding the colorability of the Mycielski graphs, are there small* PR *proofs without using new variables?*

Mycielski graphs

Mycielski graph $\mu(G)$ of $G = (V, E)$ is constructed as follows:

1. Include G in $\mu(G)$ as a subgraph.

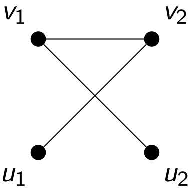


eyolcu@cs.cmu.edu

Mycielski graphs

Mycielski graph $\mu(G)$ of $G = (V, E)$ is constructed as follows:

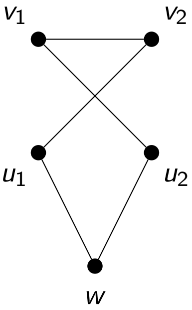
1. Include G in $\mu(G)$ as a subgraph.
2. For each $v_i \in V$, add a vertex $u_i \in U$ adjacent to all of $N_G(v_i)$.



Mycielski graphs

Mycielski graph $\mu(G)$ of $G = (V, E)$ is constructed as follows:

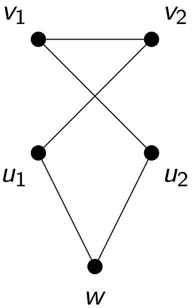
1. Include G in $\mu(G)$ as a subgraph.
2. For each $v_i \in V$, add a vertex $u_i \in U$ adjacent to all of $N_G(v_i)$.
3. Add a vertex w adjacent to all of U.



Mycielski graphs

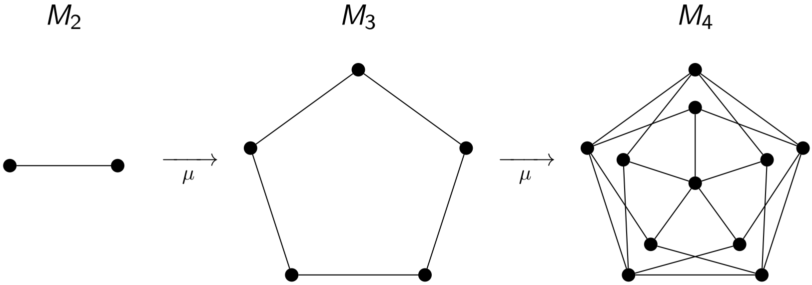
Mycielski graph $\mu(G)$ of $G = (V, E)$ is constructed as follows:

1. Include G in $\mu(G)$ as a subgraph.
2. For each $v_i \in V$, add a vertex $u_i \in U$ adjacent to all of $N_G(v_i)$.
3. Add a vertex w adjacent to all of U.



- Unless G has a triangle $\mu(G)$ does not have a triangle.
- $\mu(G)$ has chromatic number one higher than G.

Mycielski graphs



M_k has $\Theta(2^k)$ vertices and $\Theta(3^k)$ edges.

Main references

Marijn J. H. Heule, Benjamin Kiesl, and Armin Biere (2019)
Strong extension-free proof systems

Sam Buss and Neil Thapen (2019)
DRAT and propagation redundancy proofs without new variables

Proof complexity

Interested in proofs (refutations) of unsatisfiable formulas F.

Proof complexity

Interested in proofs (refutations) of unsatisfiable formulas F.

Definition

A valid *clausal proof* of F is a sequence $(C_1, \omega_1), \ldots, (C_N, \omega_N)$ where, defining $F_i := F \wedge \bigwedge_{j=1}^{i} C_j$,

- each clause C_i preserves satisfiability, i.e. is redundant wrt F_{i-1},
- the redundancy of C_i is decidable in polynomial time given ω_i,
- $C_N = \bot$ (empty clause).

Reverse unit propagation (RUP)

- Partial assignments are conjunctions of literals.
  Example: $x \mapsto 1$, $y \mapsto 0$, $z \mapsto 1$ is denoted by $x \wedge \overline{y} \wedge z$.

Reverse unit propagation (RUP)

- Partial assignments are conjunctions of literals.
  Example: $x \mapsto 1$, $y \mapsto 0$, $z \mapsto 1$ is denoted by $x \wedge \overline{y} \wedge z$.

- $\overline{C}$ denotes the assignment corresponding to $\bigwedge_{p \in C} \overline{p}$.

eyolcu@cs.cmu.edu

Reverse unit propagation (RUP)

- Partial assignments are conjunctions of literals.
  Example: $x \mapsto 1$, $y \mapsto 0$, $z \mapsto 1$ is denoted by $x \wedge \overline{y} \wedge z$.

- $\overline{C}$ denotes the assignment corresponding to $\bigwedge_{p \in C} \overline{p}$.

- Unit propagation: satisfy a unit clause, restrict formula, repeat.

Reverse unit propagation (RUP)

- Partial assignments are conjunctions of literals.
  Example: $x \mapsto 1$, $y \mapsto 0$, $z \mapsto 1$ is denoted by $x \wedge \overline{y} \wedge z$.
- $\overline{C}$ denotes the assignment corresponding to $\bigwedge_{p \in C} \overline{p}$.
- Unit propagation: satisfy a unit clause, restrict formula, repeat.
- C is a RUP inference from F if unit propagation refutes $F \wedge \overline{C}$.

Reverse unit propagation (RUP)

- Partial assignments are conjunctions of literals.
  Example: $x \mapsto 1$, $y \mapsto 0$, $z \mapsto 1$ is denoted by $x \wedge \overline{y} \wedge z$.
- $\overline{C}$ denotes the assignment corresponding to $\bigwedge_{p \in C} \overline{p}$.
- Unit propagation: satisfy a unit clause, restrict formula, repeat.
- C is a RUP inference from F if unit propagation refutes $F \wedge \overline{C}$.
- $F \vDash_1 H$ means each clause $C \in H$ is a RUP inference from F.
  We say that F *implies* H *by unit propagation*.

PR proof system

Definition

Let F be a formula, C a clause, and $\alpha = \overline{C}$.

C is *propagation redundant* with respect to F if there exists an assignment ω such that

$$\omega \text{ satisfies } C \qquad \text{and} \qquad F|_\alpha \vDash_1 F|_\omega.$$

PR proof system

Definition

Let F be a formula, C a clause, and $\alpha = \overline{C}$.

C is *propagation redundant* with respect to F if there exists an assignment ω such that

$$\omega \text{ satisfies } C \qquad \text{and} \qquad F|_\alpha \vdash_1 F|_\omega.$$

Intuitively, PR clauses allow us to argue that satisfying assignments can be assumed to have certain properties. This can be seen as capturing **"without loss of generality"** arguments.

PR variants

Formula F, proof π, clause–witness pair (C_i, ω_i)

- SPR: Require $\text{dom}(\omega_i) = \text{dom}(\alpha_i)$.
- PR$^-$: Restrict C_i to only include variables appearing in F.
- DPR: Allow deletion of a previous clause in π or F.

PR variants

Formula F, proof π, clause–witness pair (C_i, ω_i)

- SPR: Require $\mathrm{dom}(\omega_i) = \mathrm{dom}(\alpha_i)$.
- PR^-: Restrict C_i to only include variables appearing in F.
- DPR: Allow deletion of a previous clause in π or F.

Definition

For a PR inference, its *discrepancy* is $|\mathrm{dom}(\omega) \setminus \mathrm{dom}(\alpha)|$.

Theorem (Buss and Thapen, 2019)

A PR *proof of length* N *and discrepancy* $\leq \delta$ *can be converted into an* SPR *proof of length* $O(2^\delta N)$ *without additional variables.*

eyolcu@cs.cmu.edu

Proof complexity of PR

With new variables, PR can simulate Extended Resolution.

Proof complexity of PR

With new variables, PR can simulate Extended Resolution.

Without new variables, SPR^- was shown to admit short proofs of:

- pigeonhole principle
- bit pigeonhole principle
- parity principle
- clique-coloring principle
- Tseitin tautologies

Proof complexity of PR

With new variables, PR can simulate Extended Resolution.

Without new variables, SPR^- was shown to admit short proofs of:

- pigeonhole principle
- bit pigeonhole principle
- parity principle
- clique-coloring principle
- Tseitin tautologies

Without new variables, strictly stronger than RAT: it is known that RAT^- does not simulate SPR^-.

Proof complexity of PR

With new variables, PR can simulate Extended Resolution.

Without new variables, SPR^- was shown to admit short proofs of:

- pigeonhole principle
- bit pigeonhole principle
- parity principle
- clique-coloring principle
- Tseitin tautologies

Without new variables, strictly stronger than RAT: it is known that RAT^- does not simulate SPR^-.

Currently, there are no known lower bounds for PR^- (or SPR^-).

Mycielski graph formulas

Encoding graph coloring: Given graph $G = (V, E)$ and k colors.

$$\bigvee_{c \in [k]} x_c \qquad \text{for each } x \in V$$

$$\overline{x_c} \vee \overline{y_c} \qquad \text{for each } xy \in E,\ c \in [k]$$

Mycielski graph formulas

Encoding graph coloring: Given graph $G = (V, E)$ and k colors.

$$\bigvee_{c \in [k]} x_c \qquad \text{for each } x \in V$$

$$\overline{x_c} \vee \overline{y_c} \qquad \text{for each } xy \in E,\ c \in [k]$$

$\text{MYC}_k \equiv$ "M_k is $(k-1)$-colorable."

Results

Let $N = \Theta(3^k k)$ be the length of MYC_k.

Theorem

MYC_k *has* DSPR^- *proofs of length* $O(N \log N)$ *and constant width.*

Results

Let $N = \Theta(3^k k)$ be the length of MYC_k.

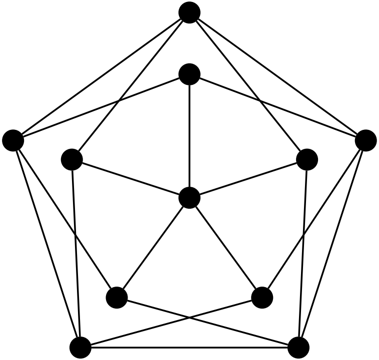
Theorem

MYC_k *has* DSPR^- *proofs of length* $O(N \log N)$ *and constant width.*

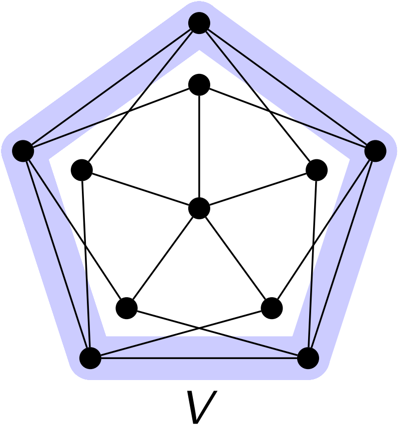
Theorem

MYC_k *has* PR^- *proofs of length* $O(N^{\log_3 2}(\log N)^2)$*, constant width, and maximum discrepancy* $\Theta(2^k)$*.*

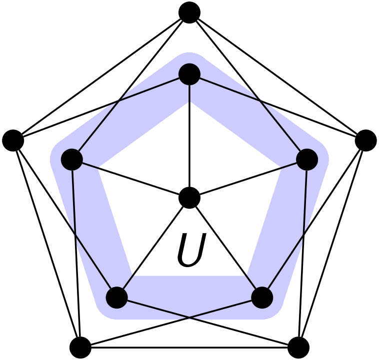
Proof outline



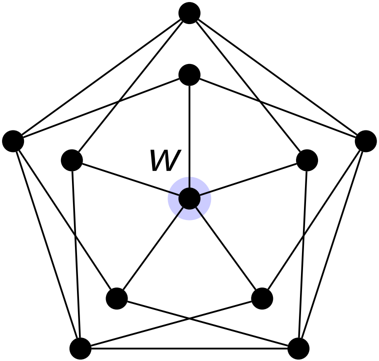
Proof outline



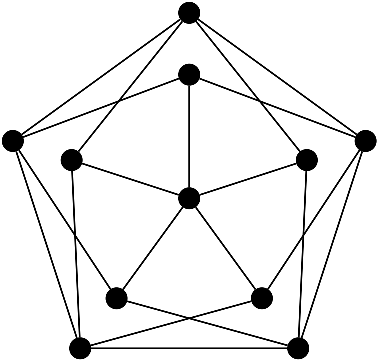
Proof outline



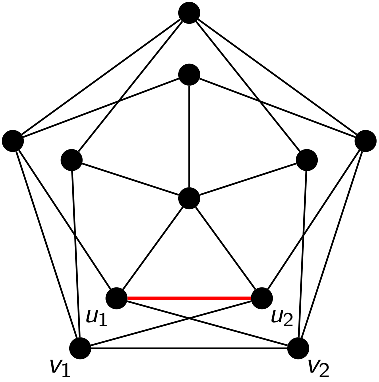
Proof outline



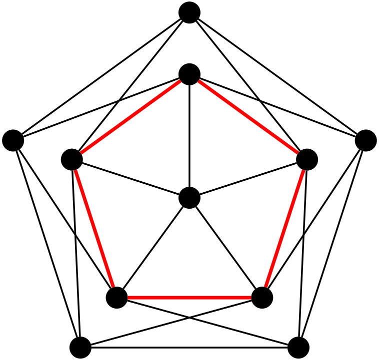
Proof outline



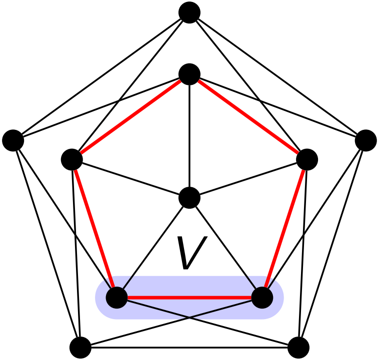
Proof outline



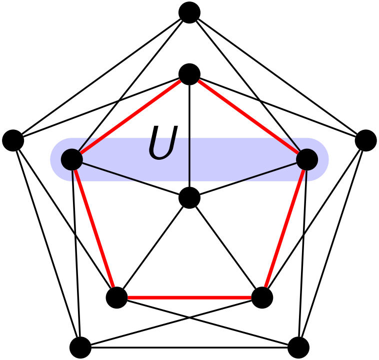
Proof outline



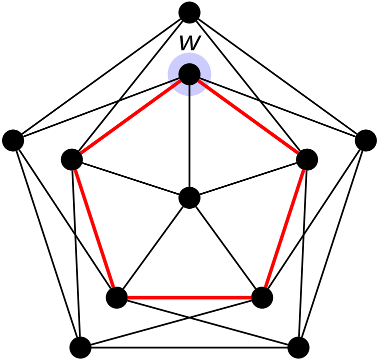
Proof outline



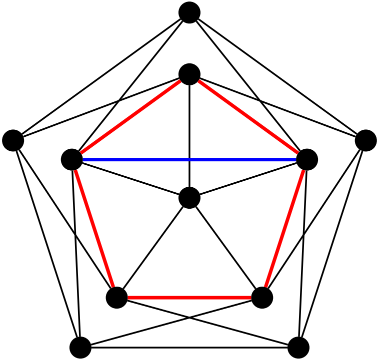
Proof outline



Proof outline



Proof outline



DPR$^-$ proof

Partition the vertices of M_k into $V \cup U \cup \{w\}$.

Let E_{k-1} denote the edge set of the $(k-1)$th Mycielski graph.

Let $n_k = |V| = |U|$.

DPR^{-} proof

Partition the vertices of M_k into $V \cup U \cup \{w\}$.
Let E_{k-1} denote the edge set of the $(k-1)$th Mycielski graph.
Let $n_k = |V| = |U|$.

1. Introduce the blocked clauses

$$\begin{aligned} &\overline{v_{i,c}} \vee \overline{v_{i,c'}} && \text{for each } i \in [n_k] \\ &\overline{u_{i,c}} \vee \overline{u_{i,c'}} && \text{for each } i \in [n_k] \\ &\overline{w_{c}} \vee \overline{w_{c'}} && \end{aligned}$$

   for each $c, c' \in [k-1]$ such that $c < c'$.

eyolcu@cs.cmu.edu

DPR$^-$ proof

2. Introduce the PR clauses

$$\overline{v_{i,c}} \lor \overline{u_{i,c'}} \lor w_c \qquad \text{for each } i \in [n_k] \text{ and for each } c, c' \in [k-1],\ c \neq c'.$$

DPR$^-$ proof

2. Introduce the PR clauses

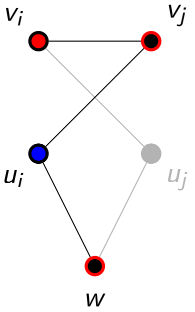
$$\overline{v_{i,c}} \vee \overline{u_{i,c'}} \vee w_c \qquad \text{for each } i \in [n_k] \text{ and}$$
$$\text{for each } c, c' \in [k-1],\ c \neq c'.$$

DPR$^-$ proof

2. Introduce the PR clauses

$$\overline{v_{i,c}} \vee \overline{u_{i,c'}} \vee w_c \qquad \text{for each } i \in [n_k] \text{ and for each } c, c' \in [k-1],\ c \neq c'.$$

$$\alpha = v_{i,c} \wedge u_{i,c'} \wedge \overline{w_c}$$



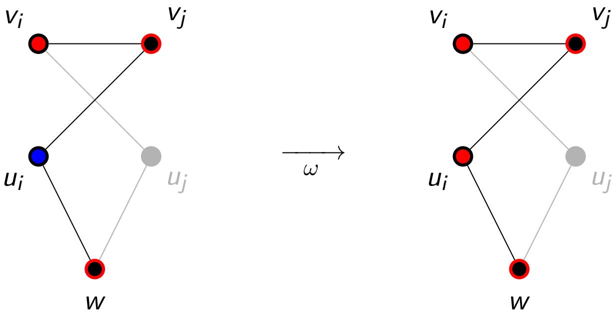
eyolcu@cs.cmu.edu

DPR^{-} proof

2. Introduce the PR clauses

$$\overline{v_{i,c}} \vee \overline{u_{i,c'}} \vee w_c \qquad \text{for each } i \in [n_k] \text{ and for each } c, c' \in [k-1],\ c \neq c'.$$

$$\alpha = v_{i,c} \wedge u_{i,c'} \wedge \overline{w_c} \qquad \omega = v_{i,c} \wedge \overline{u_{i,c'}} \wedge u_{i,c} \wedge \overline{w_c}$$



eyolcu@cs.cmu.edu

DPR$^-$ proof

3. Introduce RUP inferences

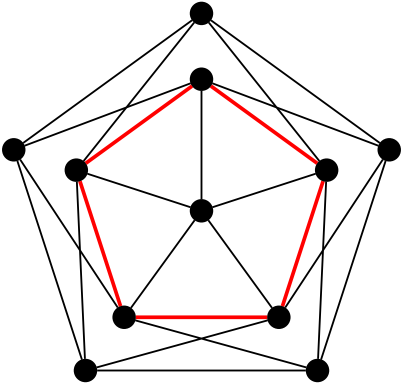
$$\overline{u_{i,c}} \vee \overline{u_{j,c}} \vee \overline{v_{i,c'}} \qquad \text{for each } i, j \text{ such that } v_i v_j \in E_{k-1} \text{ and for each } c, c' \in [k-1],\ c \neq c'.$$

4. Introduce RUP inferences

$$\overline{u_{i,c}} \vee \overline{u_{j,c}} \qquad \text{for each } i, j \text{ such that } v_i v_j \in E_{k-1} \text{ and for each } c \in [k-1].$$

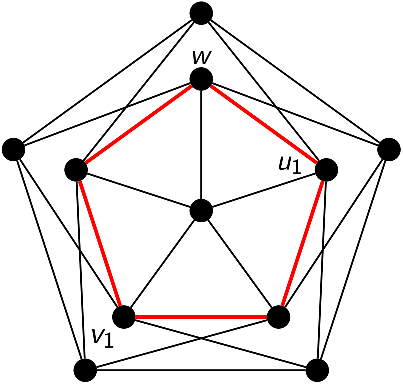
DPR$^-$ proof

Effect of adding the last set of RUP inferences on M_4. We obtain a subgraph isomorphic to M_3:



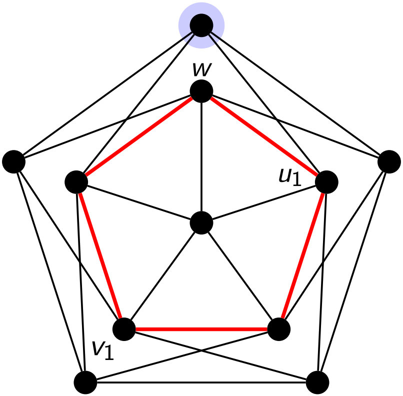
DPR$^-$ proof

Effect of adding the last set of RUP inferences on M_4. We obtain a subgraph isomorphic to M_3:



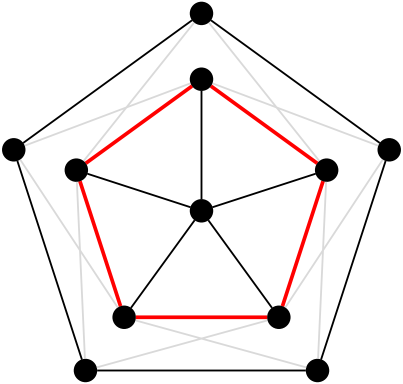
DPR$^-$ proof

Effect of adding the last set of RUP inferences on M_4. We obtain a subgraph isomorphic to M_3:



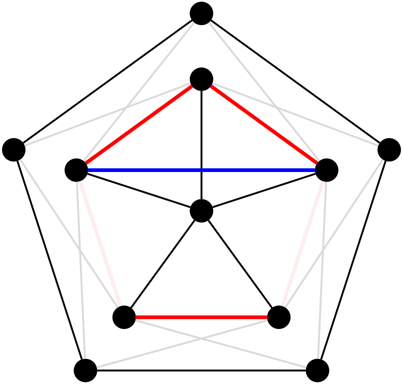
DPR$^-$ proof

5. Next, we delete the clauses introduced in steps 2, 3, and the clauses corresponding to the edges between U and V.



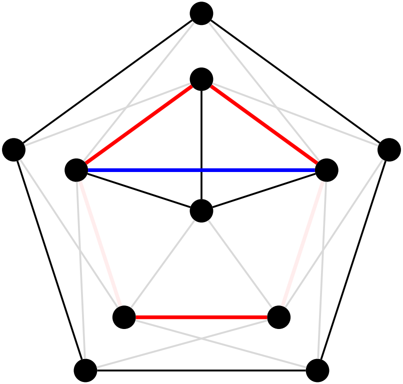
DPR$^-$ proof

6. Then we inductively repeat steps 2–5: introducing clauses and deleting the intermediate ones for subgraphs of decreasing order.



DPR^{-} proof

7. After obtaining the k-clique, we delete all the edges leaving the clique. This detaches the clique from the rest of the graph.



DPR$^-$ proof

Finally, we concatenate a PR$^-$ proof of the pigeonhole principle to derive the empty clause.

Proof has length $O(3^k k^2)$ and PR inferences have discrepancy ≤ 2.
$\Longrightarrow$ There exists a DSPR$^-$ proof of length $O(3^k k^2)$.

Letting $N = \Theta(3^k k)$ denote the length of MYC$_k$, the proof is of quasilinear length $O(N \log N)$.

PR^- proof

At a high level the idea remains the same.

1. Introduce the blocked clauses as before.
2. Introduce the PR inferences as before.

PR$^-$ proof

At a high level the idea remains the same.

1. Introduce the blocked clauses as before.
2. Introduce the PR inferences as before.
3. It becomes possible to infer the following clauses via PR.

$$\overline{u_{i,c}} \vee \overline{v_{i,c'}} \qquad \text{for each } i \in [n_k] \text{ and for each } c, c' \in [k-1],\ c \neq c'.$$

At this step, $u_{i,c} \leftrightarrow v_{i,c}$ is implied via unit propagation. Due to the edge $v_i v_j$, the edge $u_i u_j$ is also implied via unit propagation.

PR^- proof

At a high level the idea remains the same.

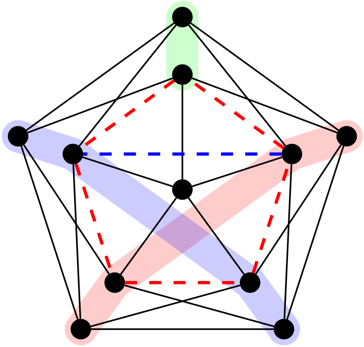
1. Introduce the blocked clauses as before.
2. Introduce the PR inferences as before.
3. It becomes possible to infer the following clauses via PR.

$$\overline{u_{i,c}} \vee \overline{v_{i,c'}} \qquad \text{for each } i \in [n_k] \text{ and}$$
$$\qquad\qquad\qquad \text{for each } c, c' \in [k-1],\ c \neq c'.$$

   At this step, $u_{i,c} \leftrightarrow v_{i,c}$ is implied via unit propagation. Due to the edge $v_i v_j$, the edge $u_i u_j$ is also implied via unit propagation.
4. Inductively repeat steps 2–3 for each subgraph as before.

PR$^-$ proof

Equivalent vertices and implied edges on M_4:



PR^- proof

5. At the end, a k-clique is implied via unit propagation. We conclude with a PR^- proof of the pigeonhole principle.

Proof has length $S = O(2^k k^2)$. Letting $N = \Theta(3^k k)$ denote the length of MYC_k, the proof is of sublinear length $O(N^{\log_3 2}(\log N)^2)$.

Maximum discrepancy is $\Omega(S/(\log S)^2)$ and the existence of a small SPR^- proof for the Mycielski graph formulas remains open.

Extended MYC_k formulas

Denote by

- `BC`: the blocked clauses that we add in step 1,
- `PR`: the PR clauses that we add inductively in step 2,
- `R1`: the RUP inferences that we add inductively in step 3,
- `R2`: the RUP inferences that we add inductively in step 4.

Consider versions of MYC_k where we cumulatively include more of the redundant clauses.

Example: $\mathsf{MYC}_k+\mathtt{PR}$ includes the redundant clauses from `BC` and `PR`.

CDCL performance

k	MYC_k	BC	PR	R1	R2
5	0.07	0.04	0.03	0.01	0.00
6	29.53	24.51	1.17	0.03	0.01
7	—	—	26.80	0.28	0.02
8	—	—	1503	1.33	0.19
9	—	—	—	22.99	0.88
10	—	—	—	196.18	12.88

Finding proofs for $\text{MYC}_k+\texttt{PR}$

At a high level, the method we implemented is as follows.

1. Remove the clause with the largest number of resolution candidates until the formula becomes satisfiable.
2. Sample satisfying assignments using a local search solver and find pairs of literals (cubes) that do not appear together in any of the assignments.
3. Partition the list of cubes into pieces, use parallel workers to perform incremental solving with a limit on the number of conflicts allowed. Aggregate a list of refuted cubes.
4. Run CDCL on the conjunction of the original formula with the negations of all the refuted cubes.

Results on finding proofs

k	time to cubes	conflict limit	#workers	time to solve
8	2m 15s	100	1	2m 50s
			12	44.4s
9	10m 37s	100	1	38m 40s
			12	7m 4s
10	35m 18s	100	1	11h 37m
			12	1h 55m

Open questions

- Are there small, constant-width SPR^- proofs of the Mycielski graph formulas?
- What is a potentially difficult principle for PR^-?
- Is there a resource trade-off for PR^- proofs involving the maximum discrepancy?
- Is PR^- (or SPR^-) width-automatizable? Are there proof search heuristics that perform well on the typical practical benchmarks?
- Are there other generic heuristics that can help CDCL handle the $\mathsf{MYC}_k + \mathtt{PR}$ formulas?