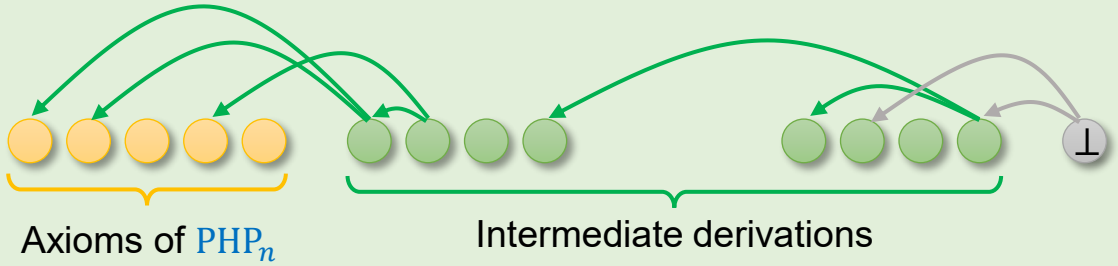
Finding Bugs in Short Proofs: The Metamathematics of Resolution Lower Bounds

Jiawei Li, Yuhao Li, ***Hanlin Ren***

Motivating question: How hard is it to prove lower bounds in proof complexity?

Haken's Theorem (1985)

Any resolution refutation for the pigeonhole principle PHP_n requires size $\geq 1.01^n$.



Resolution rule

$$\frac{C \vee \ell \qquad D \vee \bar{\ell}}{C \vee D}$$

Weakening rule

$$\frac{C}{C \vee D}$$

The Refuter Problem: $\mathrm{Refuter}(\mathrm{Haken})$

Input: a circuit C encoding a claimed resolution proof Π of PHP, but the proof size is $\leq 1.01^n$.
($C(i)$ = clause written at the i-th line & how it was derived)

Output: an index i such that step i of Π is an **invalid derivation** (a "bug").

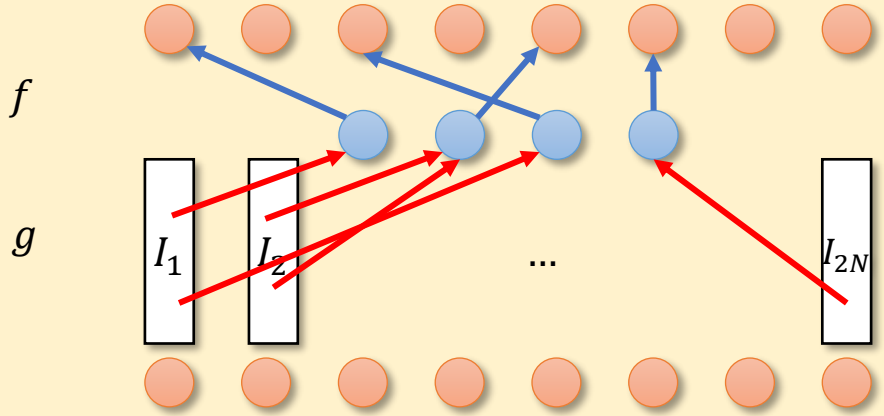
Haken's theorem $\Longrightarrow$ this problem is total (in the class $\mathbf{TFNP}$)!

Q: What is the complexity of this TFNP problem?

Our new class: $\mathrm{rwPHP}(\mathbf{PLS})$

Input: $f\colon [N] \to [2N]$, $2N$ **PLS** instances $I_1, I_2, \ldots, I_{2N}$; each solution ans of I_y carries an outgoing edge $g_{y,ans} \in [N]$.

Output: some $y \in [2N]$ and a valid solution ans of I_y such that $f(g_{y,ans}) \neq y$.



Hint 1: This is a "randomized" version of **PLS**

Hint 2: **PLS** = resolution; $\mathrm{rwPHP}(\mathbf{PLS})$ is strictly stronger than resolution

Main Theorem: $\mathbf{Refuter}(\mathbf{Haken})$ is $\mathrm{rwPHP}(\mathbf{PLS})$-complete.

Moreover, $\mathrm{rwPHP}(\mathbf{PLS})$-hardness does not depend on the hard tautology being PHP_n.

Key takeaway (informal):

There is a proof system P such that any system weaker than P cannot prove resolution lower bounds. Moreover, P is strictly stronger than resolution itself!