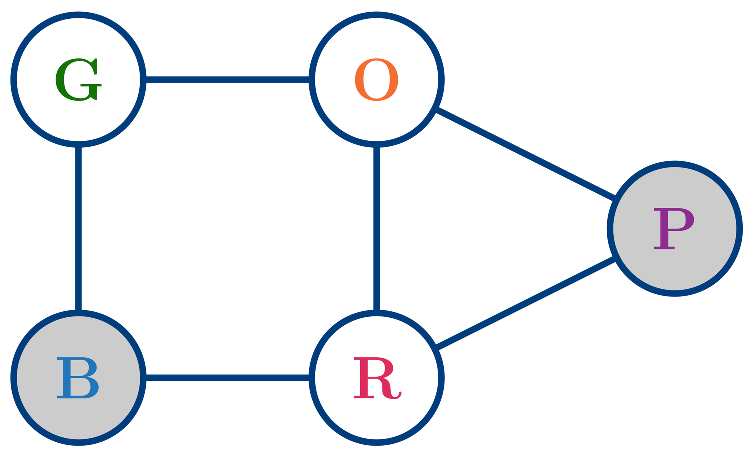
Solving the Identifying Code Set Problem with Grouped Independent Support

Anna L.D. Latour[1], Arunabha Sen[2], Kuldeep S. Meel[1]

1. Generalised Identifying Code Set (GICS) problem

$\Gamma := (V, E)$
$D := \{B, P\}$



Closed 1-neighbourhoods of (sets of) nodes:

$$N_1^+(v) := \{v\} \cup N_1(v) \quad \text{(for } v \in V\text{)}$$
$$N_1^+(U) := \bigcup_{v \in U} N_1^+(v) \quad \text{(for } U \subseteq V\text{)}$$

Given a set of nodes $D \subseteq V$, we define the **signature** of another set of nodes $U \subseteq V$ as $s_U := \langle S_U^0, S_U^1 \rangle$.

U	$S_U^0 := D \cap U$	$S_U^1 := D \cap N_1^+(U)$
$\varnothing$	$\varnothing$	$\varnothing$
$\{B\}$	$\{B\}$	$\{B\}$
$\{G\}$	$\varnothing$	$\{B\}$
$\{O\}$	$\varnothing$	$\{P\}$
$\{R\}$	$\varnothing$	$\{B, P\}$
$\{P\}$	$\{P\}$	$\{P\}$
$\{B, G\}$	$\{B\}$	$\{B\}$
$\{B, O\}$	$\varnothing$	$\{P\}$
$\vdots$	$\vdots$	$\vdots$

D is a **generalised identifying code set (GICS)** [KCL1998, SGBZG2019] of $\langle \Gamma, k \rangle$ if each $U \subseteq V$ with $|U| \leq k$, has a unique signature s_U.

GICS problem: minimise $|D|$.

2. Previous state of the art

Encode the problem into an **integer-linear program (ILP)** and solve with off-the-shelve MIP solver CPLEX [PBPBS2020].

- Checking if a candidate is a solution: **polytime**.
- Returns **cardinality-minimal** solution.
- **Problem**: encoding has $O\left(\binom{|V|}{k}^2\right)$ constraints.

3. Independent Support

Solution $\sigma : X \mapsto \{0, 1\}$ maps variables to truth values.

Example formula: $F(X) := (x_1 \vee x_2) \leftrightarrow x_3$
Projection set: $I := \{x_1, x_2\}$

	x_1	x_2	x_3
σ_1	1	1	1
σ_2	1	0	1
σ_3	0	1	1
σ_4	0	0	0

$$|Sol_{\downarrow I}(F)| = |Sol(F)|$$

I is an **independent support** [CFMSV2014] of $F(X)$.

Key property: an independent support preserves the cardinality of its solution set after projection.

4. Grouped Independent Support

Extension of *independent support*.

Given: $F(X)$ and a partition $\mathcal{G}$ of variables X.

$\mathcal{I} \subseteq \mathcal{G}$ is a **grouped independent support (GIS)** of $F(X)$ if $\bigcup_{G \in \mathcal{I}} G$ is an independent support of $F(X)$.

5. New Approach

Encode GICS problem as **CNF** $F(Z)$ with partition $\mathcal{G}$ of Z. Use new solver **gismo** to find a GIS for $\langle F(Z), \mathcal{G} \rangle$.

- Checking if a candidate is a solution: **co-NP**.
- **Gismo** returns **set-minimal** solution.
- Encoding has $O(k \cdot |V| + |E|)$ clauses (**linear!**).

by reducing to a

computationally harder

problem, we can

exponentially decrease

the encoding size, and

solve much larger

instances

6. Reduction of GICS to GIS

Variable **groups**:

$$\mathcal{G} := \{G_v := \{x_v, y_v\} \mid v \in V\}$$
$$= \{G_B, G_G, G_O, G_R, G_P\}$$

Constraints:

$$F_{\text{detection}} := \bigwedge_{v \in V} \left(y_v \leftrightarrow \bigvee_{u \in N_1^+(v)} x_u \right)$$
$$F_{\text{cardinality},k} := \sum_{v \in V} x_v \leq k$$

Transform to **CNF**:

$$F_k(X \cup Y, A) := F_{\text{detection}} \wedge F_{\text{cardinality},k}$$

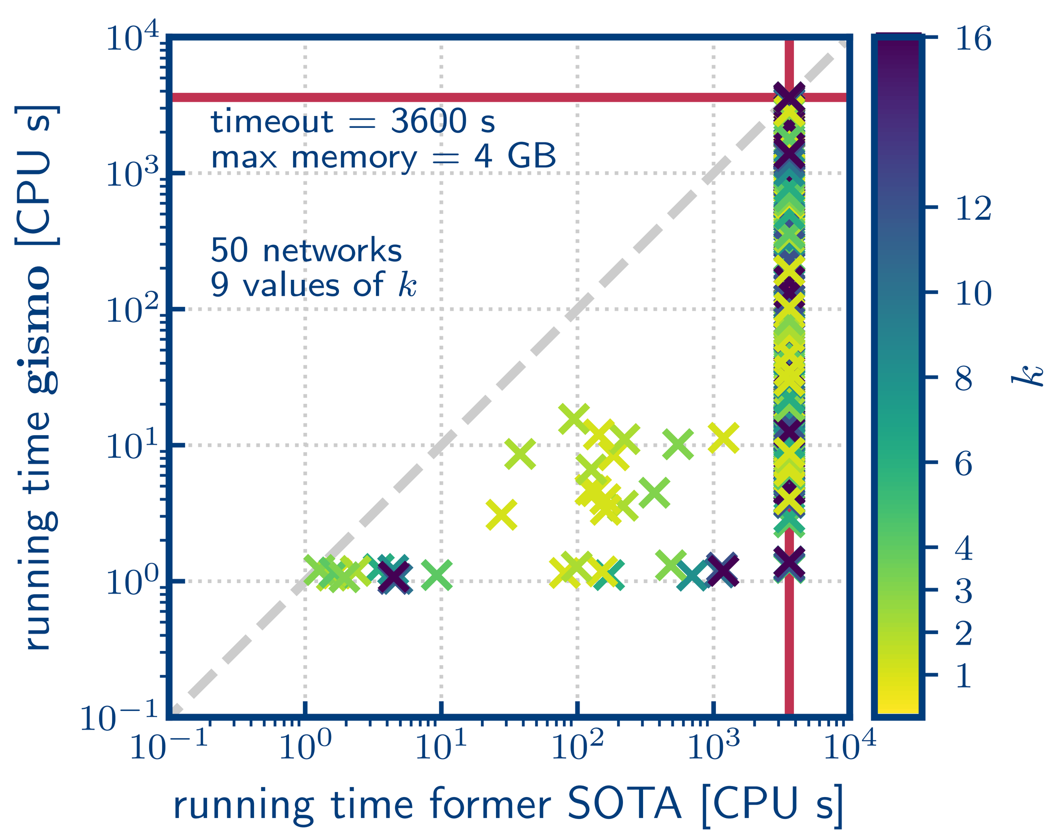
Number of clauses is **linear** in problem size:

$$O(k \cdot |V| + |E|)$$

7. Results

Our experiments show the following:

- Model size **scales linearly** with problem size.
- **Gismo** solves $8\times$ **more instances** than previous state of the art.
- **Gismo** is **2–6 times faster** in terms of PAR2, and up to $520\times$ **faster** in terms of median running time.
- For the majority of instances, **gismo**'s solution is **at most** 10% **larger** than that of the state of the art.
- **Gismo** solves $43\times$ **larger instances** than previous state of the art, and for **larger values of** k.



Reduction of GICS to GIS (example)

Solutions of $F_1(X \cup Y, A)$ projected on GIS $\mathcal{I} = \{G_B := \{x_B, y_B\}, G_P := \{x_P, y_P\}\} \subseteq \mathcal{G}$:

Observations:

- **Bijective** relation between set of solutions and set of signatures.
- All projected solutions are **unique**.
- Hence, all **signatures** are unique.

	X (S_U^0)					Y (S_U^1)						
U	x_B	x_G	x_O	x_R	x_P	y_B	y_G	y_O	y_R	y_P	S_U^0	S_U^1
$\varnothing$	0	0	0	0	0	0	0	0	0	0	$\varnothing$	$\varnothing$
$\{B\}$	1	0	0	0	0	1	1	0	1	0	$\{B\}$	$\{B\}$
$\{G\}$	0	1	0	0	0	1	1	1	0	0	$\varnothing$	$\{B\}$
$\{O\}$	0	0	1	0	0	0	1	1	1	1	$\varnothing$	$\{P\}$
$\{R\}$	0	0	0	1	0	1	0	1	1	1	$\varnothing$	$\{B, P\}$
$\{P\}$	0	0	0	0	1	0	0	1	1	1	$\{P\}$	$\{P\}$

[1] NUS National University of Singapore
[2] Arizona State University

This work was supported in part by National Research Foundation Singapore under its NRF Fellowship Programme [NRF-NRFFAI1-2019-0004], Ministry of Education Singapore Tier 2 grant MOE-T2EP20121-0011, and Ministry of Education Singapore Tier 1 Grant [R-252-000-B59-114]. The computational work for this project was performed on resources of the National Supercomputing Centre, Singapore www.nscc.sg.

paper (preprint) **code** (and more!)

contact & more info

github.com/meelgroup/gismo
arxiv.org/abs/2306.15693
www.annalatour.nl
latour@nus.edu.sg