Solving larger instances by reducing to a computationally harder problem

Anna L.D. Latour[1,2], Arunabha Sen[3] and Kuldeep S. Meel[1,4]

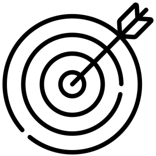
[1]School of Computing, National University of Singapore
[2]Department of Software Technology, Delft University of Technology
[3]Computer Science and Engineering Faculty, Arizona State University
[4]Department of Computer Science, University of Toronto

27 Nov 2024, SIGAlgo symposium

Motivation

Problem

- There will always be an instance whose encoding is too big.
- Smaller encoding does not always lead to faster solving.

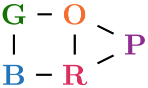


Goal

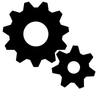
Make encoding exponentially more succinct without sacrificing speed.

Target image from www.freepik.com.

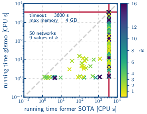
Main contributions



A **case study** that reduces the NP-hard **generalised identifying code set (GICS)** problem to the computationally harder **independent support** problem.



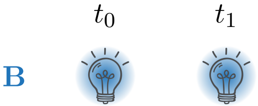
A new solver: **gismo**.

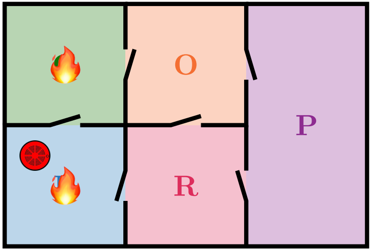


Experiments that **demonstrate the effectiveness** of our reduction and **gismo**.

Problem (1/4)

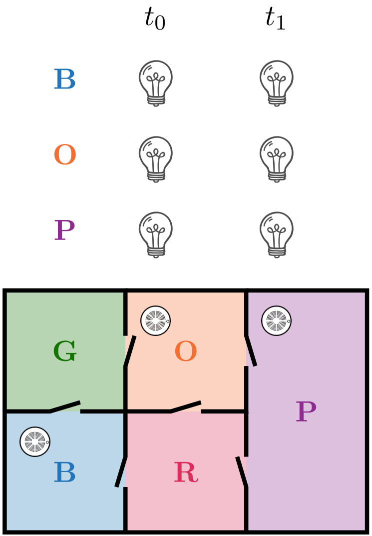
Problem (2/4)





	signature	
🔥	t_0	t_1
{B}	{B}	{B}
{G}	∅	{B}
{O}	∅	∅
{R}	∅	{B}
{P}	∅	∅
{B, G}	{B}	{B}
{B, O}	{B}	{B}
⋮	⋮	⋮
{O, R, P}	∅	{B}
⋮	⋮	⋮
∅	∅	∅

Problem (3/4)

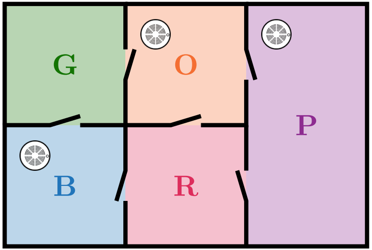


	signature	
🔥	t_0	t_1
{B}	{B}	{B}
{G}	$\varnothing$	{B, O}
{O}	{O}	{O, P}
{R}	$\varnothing$	{B, O, P}
{P}	{P}	{O, P}
{B, G}	{B}	{B, O}
{B, O}	{B, O}	{B, O}
⋮	⋮	⋮
{O, R, P}	{O, P}	{B, O, P}
⋮	⋮	⋮
$\varnothing$	$\varnothing$	$\varnothing$

Problem (4/4)

The set of rooms with a detector, D, is called a **generalised identifying code set (GICS)** (Karpovsky, Chakrabarty, and Levitin 1998) for positive integer k if each set of at most k fires has a unique signature.

Problem: minimise $|D|$



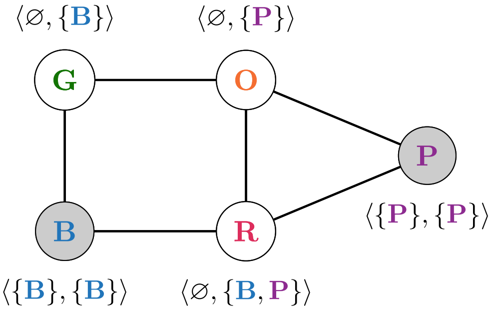
	signature	
🔥	t_0	t_1
$\{B\}$	$\{B\}$	$\{B\}$
$\{G\}$	$\varnothing$	$\{B\}$
$\{O\}$	$\{O\}$	$\{O, P\}$
$\{R\}$	$\varnothing$	$\{B, O, P\}$
$\{P\}$	$\{P\}$	$\{O, P\}$
$\varnothing$	$\varnothing$	$\varnothing$

Example:
$k = 1,\ D = \{B, O, P\}$

Problem (4/4)

The set of rooms with a detector, D, is called a **generalised identifying code set (GICS)** (Karpovsky, Chakrabarty, and Levitin 1998) for positive integer k if each set of at most k fires has a unique signature.

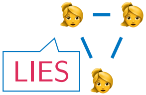
Problem: minimise $|D|$



🔥	signature t_0	t_1
$\{\mathbf{B}\}$	$\{\mathbf{B}\}$	$\{\mathbf{B}\}$
$\{\mathbf{G}\}$	$\varnothing$	$\{\mathbf{B}\}$
$\{\mathbf{O}\}$	$\varnothing$	$\{\mathbf{P}\}$
$\{\mathbf{R}\}$	$\varnothing$	$\{\mathbf{B}, \mathbf{P}\}$
$\{\mathbf{P}\}$	$\{\mathbf{P}\}$	$\{\mathbf{P}\}$
$\varnothing$	$\varnothing$	$\varnothing$

Example:
$k = 1,\ D = \{\mathbf{B}, \mathbf{P}\}$

Applications of Identifying Code Sets



Identifying sources of misinformation (Basu and Sen 2021a).

Identifying criminals in social networks (Basu and Sen 2021b).

Satellite deployment (Sen, Goliber, Basu, Zhou, and Ghosh 2019), (Latour, Sen, Basu, Zhou, and Meel 2024).

Satellite image: Flaticon.com

Solving the GICS problem

Former state of the art

(Padhee, Biswas, Pal, Basu, and Sen 2020)

Encode as integer-linear program (ILP).

- #constraints **exponential** in k.
- Checking if candidate is a solution: **polytime**.
- **Cardinality-minimal** solution D.

New approach (contribution)

Reduce GICS problem to finding a minimal **independent support** (IS).

- #clauses **linear** in k.
- Checking if candidate is an IS: **co-NP**.
- **Set-minimal** solution D.

Background: Propositional Logic

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

Example: $F(X) := (x_1 \vee x_2) \leftrightarrow x_3$

	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

Projection set: $S := \{x_1, x_3\}$

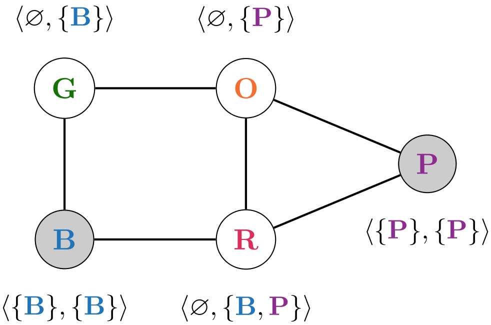
$$|Sol_{\downarrow S}(F)| \leq |Sol(F)|$$

	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

Projection set: $I := \{x_1, x_2\}$ is an **independent support** (Chakraborty, Fremont, Meel, Seshia, and Vardi 2014) of $F(X)$.

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

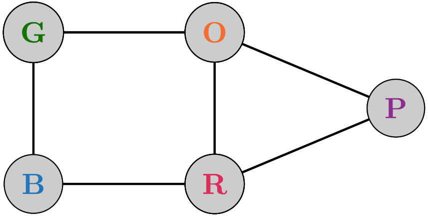
Contribution: Reduction of GICS to GIS



Our method

- Encode GICS in CNF formula
  - each solution corresponds to the signature s_U of a $U \subseteq V$ with $|U| \leq k$;
  - **linear** size.
- Two variables per node, *grouped*.
- Independent support encodes solution D.

Example

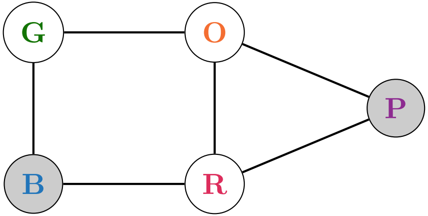


Two variables per node, *e.g.*,

- $x_{\mathbf{B}}$ models  at t_0
- $y_{\mathbf{B}}$ models  at t_1

	X (S_U^0, at t_0)					Y (S_U^1, at t_1)						
	$x_{\mathbf{B}}$	$x_{\mathbf{G}}$	$x_{\mathbf{O}}$	$x_{\mathbf{R}}$	$x_{\mathbf{P}}$	$y_{\mathbf{B}}$	$y_{\mathbf{G}}$	$y_{\mathbf{O}}$	$y_{\mathbf{R}}$	$y_{\mathbf{P}}$	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, R}
{G}	0	1	0	0	0	1	1	1	0	0	{G}	{B, G, O}
{O}	0	0	1	0	0	0	1	1	1	1	{O}	{G, O, P, R}
{R}	0	0	0	1	0	1	0	1	1	1	{R}	{B, O, P, R}
{P}	0	0	0	0	1	0	0	1	1	1	{P}	{O, P, R}

Example



Two variables per node, *e.g.*,

- $x_{\mathbf{B}}$ models 💡 at t_0
- $y_{\mathbf{B}}$ models 💡 at t_1

	X (S_U^0, at t_0)					Y (S_U^1, at t_1)						
🔥	$x_{\mathbf{B}}$	$x_{\mathbf{G}}$	$x_{\mathbf{O}}$	$x_{\mathbf{R}}$	$x_{\mathbf{P}}$	$y_{\mathbf{B}}$	$y_{\mathbf{G}}$	$y_{\mathbf{O}}$	$y_{\mathbf{R}}$	$y_{\mathbf{P}}$	S_U^0	S_U^1
$\varnothing$	0	0	0	0	0	0	0	0	0	0	$\varnothing$	$\varnothing$
$\{\mathbf{B}\}$	1	0	0	0	0	1	1	0	1	0	$\{\mathbf{B}\}$	$\{\mathbf{B}\}$
$\{\mathbf{G}\}$	0	1	0	0	0	1	1	1	0	0	$\varnothing$	$\{\mathbf{B}\}$
$\{\mathbf{O}\}$	0	0	1	0	0	0	1	1	1	1	$\varnothing$	$\{\mathbf{P}\}$
$\{\mathbf{R}\}$	0	0	0	1	0	1	0	1	1	1	$\varnothing$	$\{\mathbf{B}, \mathbf{P}\}$
$\{\mathbf{P}\}$	0	0	0	0	1	0	0	1	1	1	$\{\mathbf{P}\}$	$\{\mathbf{P}\}$

Results

Size

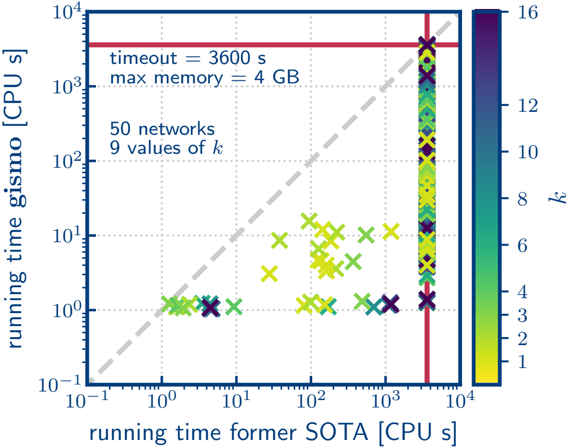
Largest network ($|V|$):

	encoded	solved
SOTA	494	494
gismo	227 320	21 363
improvement	460×	43×

SOTA: $k = 1$
gismo: for all tested k.

Majority of instances: cardinality of solution close or equal to optimum.

Time



Reducing to a **computationally harder problem**
allows us to model and
solve much larger problem instances.

full paper

www.ijcai.org/proceedings/2023/219

gismo

github.com/meelgroup/gismo

more info

www.annalatour.nl/publications

References I

- Latour, Anna L. D., Arunabha Sen, Kaustav Basu, Chenyang Zhou, and Kuldeep S. Meel (2024). "The Cardinality of Identifying Code Sets for Soccer Ball Graph with Application to Remote Sensing (preprint)". In: *arXiv*. arXiv: 2407.14120 [cs.AI]. URL: https://arxiv.org/abs/2407.14120.
- Basu, Kaustav and Arunabha Sen (2021a). "Epidemiological Model Independent Misinformation Source Identification". In: Proceedings of the ICWSM Workshops.
- — (2021b). "Identifying individuals associated with organized criminal networks: A social network analysis". In: *Soc. Networks* 64, pp. 42–54.
- Padhee, Malhar, Reetam Sen Biswas, Anamitra Pal, Kaustav Basu, and Arunabha Sen (2020). "Identifying Unique Power System Signatures for Determining Vulnerability of Critical Power System Assets". In: *SIGMETRICS Perform. Evaluation Rev.* 47.4, pp. 8–11.
- Sen, Arunabha, Victoria Horan Goliber, Kaustav Basu, Chenyang Zhou, and Sumitava Ghosh (2019). "On upper and lower bounds of identifying code set for soccer ball graph with application to satellite deployment". In: *ICDCN*. ACM, pp. 307–316.
- Chakraborty, Supratik, Daniel J. Fremont, Kuldeep S. Meel, Sanjit A. Seshia, and Moshe Y. Vardi (2014). "Distribution-Aware Sampling and Weighted Model Counting for SAT". In: *AAAI*. AAAI Press, pp. 1722–1730.
- Karpovsky, Mark G., Krishnendu Chakrabarty, and Lev B. Levitin (1998). "On a New Class of Codes for Identifying Vertices in Graphs". In: *IEEE Trans. Inf. Theory* 44.2, pp. 599–611.