Solving the Identifying Code Set Problem with Grouped Independent Support

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

[1]School of Computing, National University of Singapore

[2]Computer Science and Engineering Faculty, Arizona State University

ModRef @ CP, 27 August 2023, Toronto, Canada

Motivation

Problem

There will always be a problem whose encoding is too big.

Motivation

Problem

There will always be a problem whose encoding is too big.

Solution

Sacrifice some desiderata (*e.g.*, theoretical guarantees).

Motivation

Problem

There will always be a problem whose encoding is too big.

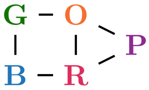
Solution

Sacrifice some desiderata (*e.g.*, theoretical guarantees).

Question

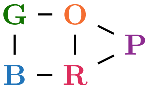
Which trade-offs can we make for an exponentially more succinct encoding?

Main contributions



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

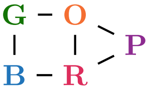
Main contributions



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

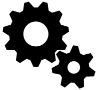
An extension of the independent support of a Boolean formula: **Grouped Independent Support (GIS)**.

Main contributions



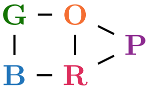
A **case study** that reduces the NP-hard **generalised identifying code set (GICS)** problem to the computationally harder **GIS** problem.

An extension of the independent support of a Boolean formula: **Grouped Independent Support (GIS)**.



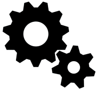
A new solver, **gismo**, for finding a grouped independent support.

Main contributions

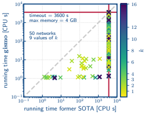


A **case study** that reduces the NP-hard **generalised identifying code set (GICS)** problem to the computationally harder **GIS** problem.

An extension of the independent support of a Boolean formula: **Grouped Independent Support (GIS)**.



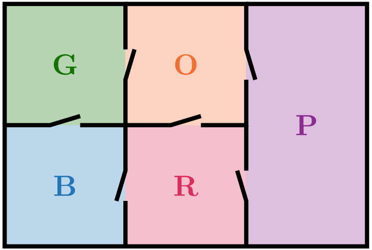
A new solver, **gismo**, for finding a grouped independent support.



Experiments that **demonstrate the effectiveness** of reducing GICS to GIS and solving with **gismo**.

Problem (1/3)

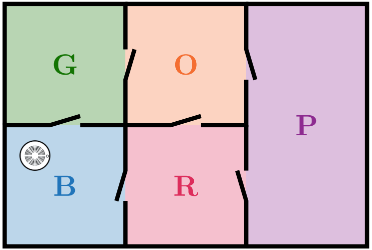
Problem (1/3)



Problem (1/3)

	t_0	t_1
B		

🔥	signature	
	t_0	t_1

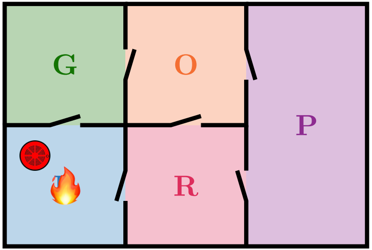


Problem (1/3)

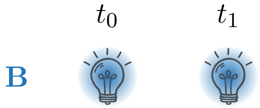
	t_0	t_1
B		

signature

🔥	t_0	t_1
$\{B\}$	$\{B\}$	

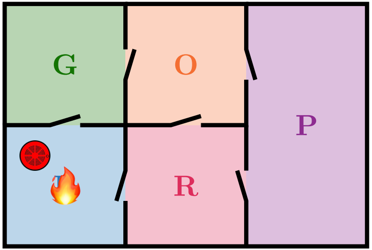


Problem (1/3)

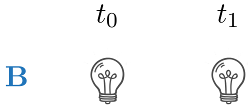


signature

🔥	t_0	t_1
{B}	{B}	{B}

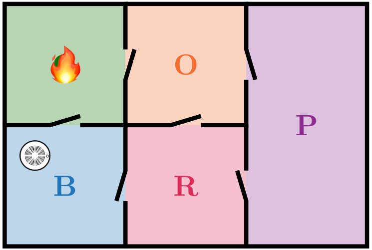


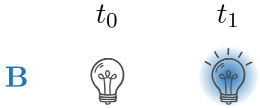
Problem (1/3)



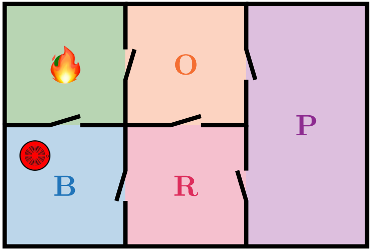
signature

🔥	t_0	t_1
$\{B\}$	$\{B\}$	$\{B\}$
$\{G\}$	$\varnothing$	





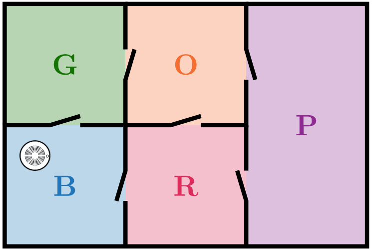
	signature	
🔥	t_0	t_1
{B}	{B}	{B}
{G}	∅	{B}



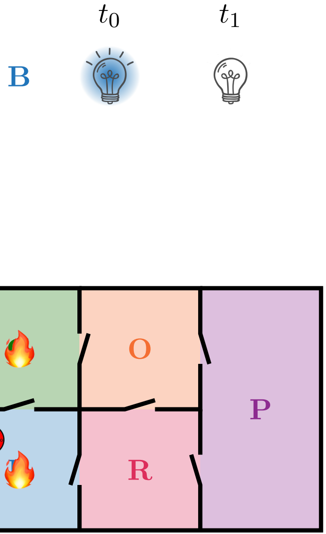
Problem (1/3)

signature

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

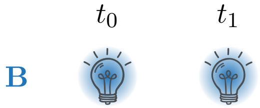


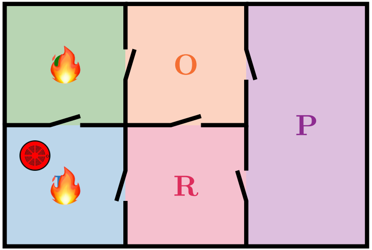
Problem (1/3)



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

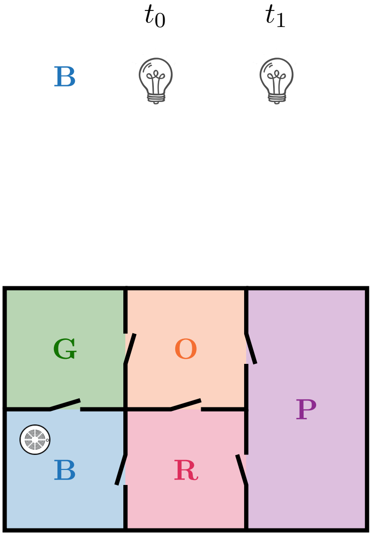
Problem (1/3)





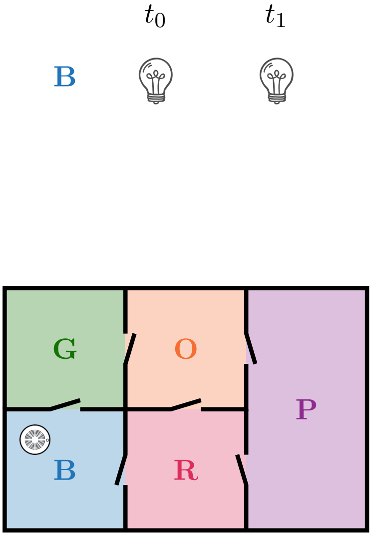
	signature	
🔥	t_0	t_1
$\{B\}$	$\{B\}$	$\{B\}$
$\{G\}$	$\varnothing$	$\{B\}$
$\{O\}$	$\varnothing$	$\varnothing$
$\{R\}$	$\varnothing$	$\{B\}$
$\{P\}$	$\varnothing$	$\varnothing$
$\{B, G\}$	$\{B\}$	$\{B\}$

Problem (1/3)



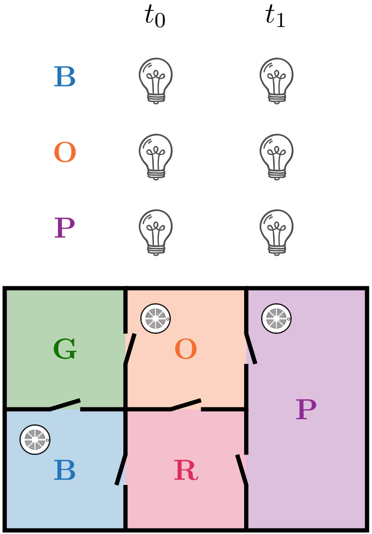
	signature	
🔥	t_0	t_1
$\{B\}$	$\{B\}$	$\{B\}$
$\{G\}$	$\varnothing$	$\{B\}$
$\{O\}$	$\varnothing$	$\varnothing$
$\{R\}$	$\varnothing$	$\{B\}$
$\{P\}$	$\varnothing$	$\varnothing$
$\{B, G\}$	$\{B\}$	$\{B\}$
$\{B, O\}$	$\{B\}$	$\{B\}$
$\vdots$	$\vdots$	$\vdots$
$\{O, R, P\}$	$\varnothing$	$\{B\}$
$\vdots$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

Problem (1/3)



	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}
⋮	⋮	⋮
∅	∅	∅

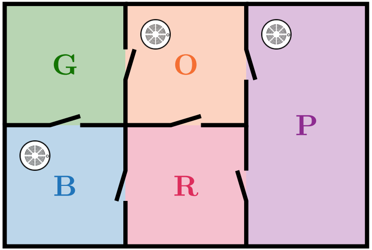
Problem (2/3)



	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 (3/3)

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.

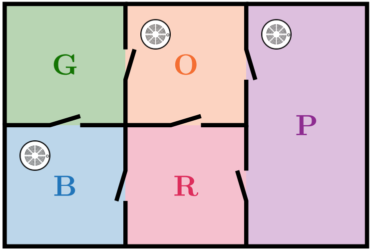


	signature	
🔥	t_0	t_1
$\{B\}$	$\{B\}$	$\{B\}$
$\{G\}$	$\varnothing$	$\{B, O\}$
$\{O\}$	$\{O\}$	$\{O, P\}$
$\{R\}$	$\varnothing$	$\{B, O, P\}$
$\{P\}$	$\{P\}$	$\{O, P\}$
$\{B, G\}$	$\{B\}$	$\{B, O\}$
$\{B, O\}$	$\{B, O\}$	$\{B, O\}$
$\vdots$	$\vdots$	$\vdots$
$\{O, R, P\}$	$\{O, P\}$	$\{B, O, P\}$
$\vdots$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

Problem (3/3)

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 $|D|$

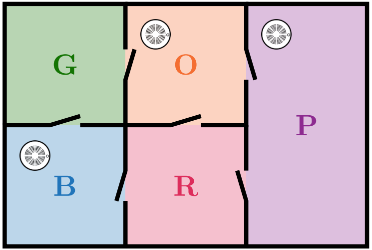


	signature	
🔥	t_0	t_1
{B}	{B}	{B}
{G}	∅	{B, O}
{O}	{O}	{O, P}
{R}	∅	{B, O, P}
{P}	{P}	{O, P}
{B, G}	{B}	{B, O}
{B, O}	{B, O}	{B, O}
⋮	⋮	⋮
{O, R, P}	{O, P}	{B, O, P}
⋮	⋮	⋮
∅	∅	∅

Problem (3/3)

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 $|D|$



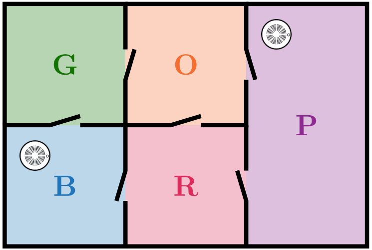
🔥	signature t_0	t_1
$\{B\}$	$\{B\}$	$\{B\}$
$\{G\}$	$\varnothing$	$\{B, O\}$
$\{O\}$	$\{O\}$	$\{O, P\}$
$\{R\}$	$\varnothing$	$\{B, O, P\}$
$\{P\}$	$\{P\}$	$\{O, P\}$
$\varnothing$	$\varnothing$	$\varnothing$

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

Problem (3/3)

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 $|D|$



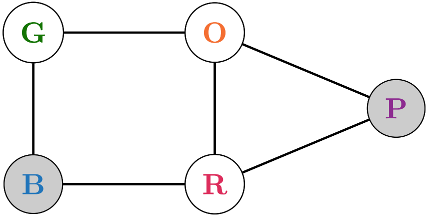
	signature	
🔥	t_0	t_1
$\{\mathbf{B}\}$	$\{\mathbf{B}\}$	$\{\mathbf{B}\}$
$\{\mathbf{G}\}$	$\varnothing$	$\{\mathbf{B}\}$
$\{\mathbf{O}\}$	$\varnothing$	$\{\mathbf{P}\}$
$\{\mathbf{R}\}$	$\varnothing$	$\{\mathbf{B}, \mathbf{P}\}$
$\{\mathbf{P}\}$	$\{\mathbf{P}\}$	$\{\mathbf{P}\}$
$\varnothing$	$\varnothing$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

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

Problem (3/3)

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 $|D|$



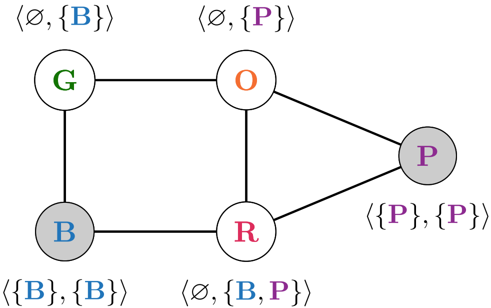
	signature	
🔥	t_0	t_1
$\{B\}$	$\{B\}$	$\{B\}$
$\{G\}$	$\varnothing$	$\{B\}$
$\{O\}$	$\varnothing$	$\{P\}$
$\{R\}$	$\varnothing$	$\{B, P\}$
$\{P\}$	$\{P\}$	$\{P\}$
$\varnothing$	$\varnothing$	$\varnothing$

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

Problem (3/3)

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.

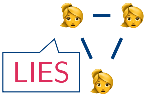
Problem: minimise $|D|$



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

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).

Satellite image: Flaticon.com

Solving the GICS problem

Former state of the art

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

1. Encode problem in integer-linear program (ILP).
2. Solve with CPLEX.

Solving the GICS problem

Former state of the art

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

1. Encode problem in integer-linear program (ILP).
2. Solve with CPLEX.

- **Exponential** $\#$ constraints ($O\left(\binom{|V|}{k}\right)$).

Solving the GICS problem

Former state of the art

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

1. Encode problem in integer-linear program (ILP).
2. Solve with CPLEX.

- **Exponential** $\#$ constraints ($O\left(\binom{|V|}{k}\right)$).

New approach (contribution)

1. Reduce GICS problem to finding a minimal **grouped independent support** (GIS).

Solving the GICS problem

Former state of the art

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

1. Encode problem in integer-linear program (ILP).
2. Solve with CPLEX.

- **Exponential** $\#$ constraints $\left(O\left(\binom{|V|}{k}\right)\right)$.

New approach (contribution)

1. Reduce GICS problem to finding a minimal **grouped independent support** (GIS).

- **Linear** $\#$ clauses $(O(k \cdot |V| + |E|))$.

Solving the GICS problem

Former state of the art

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

1. Encode problem in integer-linear program (ILP).
2. Solve with CPLEX.

- **Exponential** # constraints ($O\left(\binom{|V|}{k}\right)$).
- Checking if candidate is a solution: **polytime**.

New approach (contribution)

1. Reduce GICS problem to finding a minimal **grouped independent support** (GIS).

- **Linear** # clauses ($O\left(k \cdot |V| + |E|\right)$).
- Checking if candidate is a GIS: **co-NP**.

Solving the GICS problem

Former state of the art

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

1. Encode problem in integer-linear program (ILP).
2. Solve with CPLEX.

- **Exponential** $\#$ constraints ($O\left(\binom{|V|}{k}\right)$).
- Checking if candidate is a solution: **polytime**.
- **Cardinality-minimal** solution D.

New approach (contribution)

1. Reduce GICS problem to finding a minimal **grouped independent support** (GIS).
2. Use **gismo** to find a GIS.

- **Linear** $\#$ clauses ($O(k \cdot |V| + |E|)$).
- Checking if candidate is a GIS: **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

Background: Propositional Logic

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

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

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

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

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

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

	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: $I := \{x_1, x_2\}$

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

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 $F(X)$.

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

Grouped Independent Support (GIS)

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

Grouped Independent Support (GIS)

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

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

Grouped Independent Support (GIS)

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

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

Example 1

$\mathcal{G}_1 := \{\{x_1, x_2\}, \{x_3\}\}$

$\mathcal{I}_1 = \{\{x_1, x_2\}\}$ is a grouped independent support of $\langle F(X), \mathcal{G}_1 \rangle$.

$\bigcup_{G \in \mathcal{I}_1} G = \{x_1, x_2\}$ is an independent support of $F(X)$.

Grouped Independent Support (GIS)

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

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

Example 1

$\mathcal{G}_1 := \{\{x_1, x_2\}, \{x_3\}\}$

$\mathcal{I}_1 = \{\{x_1, x_2\}\}$ is a grouped independent support of $\langle F(X), \mathcal{G}_1 \rangle$.

$\bigcup_{G \in \mathcal{I}_1} G = \{x_1, x_2\}$ is an independent support of $F(X)$.

Example 2

$\mathcal{G}_2 := \{\{x_1\}, \{x_2, x_3\}\}$

Grouped Independent Support (GIS)

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

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

Example 1

$\mathcal{G}_1 := \{\{x_1, x_2\}, \{x_3\}\}$

$\mathcal{I}_1 = \{\{x_1, x_2\}\}$ is a grouped independent support of $\langle F(X), \mathcal{G}_1 \rangle$.

$\bigcup_{G \in \mathcal{I}_1} G = \{x_1, x_2\}$ is an independent support of $F(X)$.

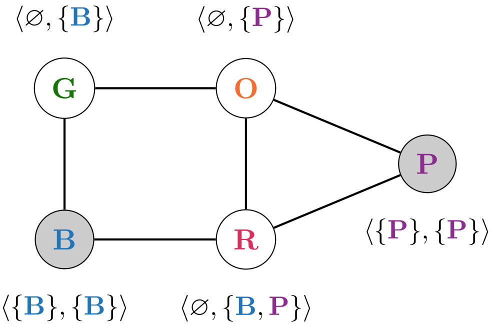
Example 2

$\mathcal{G}_2 := \{\{x_1\}, \{x_2, x_3\}\}$

$\mathcal{I}_2 = \{\{x_1\}, \{x_2, x_3\}\}$ is a grouped independent support of $\langle F(X), \mathcal{G}_2 \rangle$.

$\bigcup_{G \in \mathcal{I}_2} G = \{x_1, x_2, x_3\}$ is an independent support of $F(X)$.

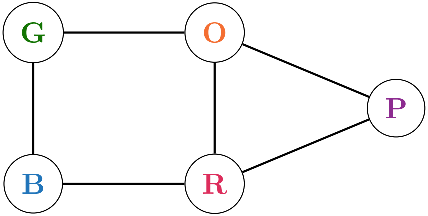
Contribution: Reduction of GICS to GIS



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 group.
- One variable group for each node.
- Use **gismo** to find minimal GIS.
- Groups in GIS correspond to nodes in D.

Example

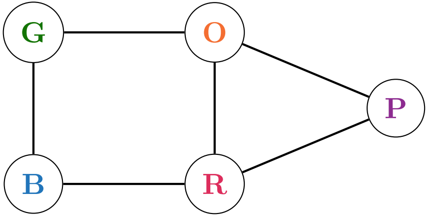


Two variables per node, *e.g.*,

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

Partition: $\mathcal{G} := \{G_v := \{x_v, y_v\} \mid v \in V\}$
$$= \{G_{\mathbf{B}}, G_{\mathbf{G}}, G_{\mathbf{O}}, G_{\mathbf{R}}, G_{\mathbf{P}}\}$$

Example



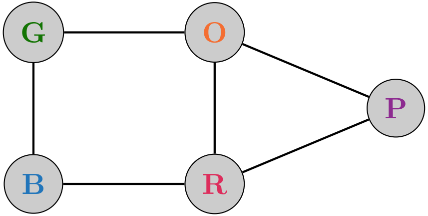
Two variables per node, *e.g.*,

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

Partition: $\mathcal{G} := \{G_v := \{x_v, y_v\} \mid v \in V\}$
$= \{G_{\mathbf{B}}, G_{\mathbf{G}}, G_{\mathbf{O}}, G_{\mathbf{R}}, G_{\mathbf{P}}\}$

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

Example



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

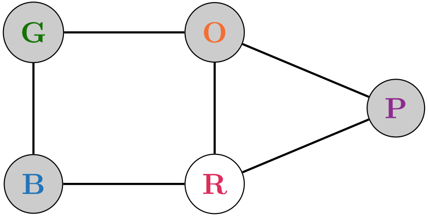
- $x_{\mathbf{B}}$ models at t_0
- $y_{\mathbf{B}}$ models at t_1

Partition: $\mathcal{G} := \{G_v := \{x_v, y_v\} \mid v \in V\}$

$= \{G_{\mathbf{B}}, G_{\mathbf{G}}, G_{\mathbf{O}}, G_{\mathbf{R}}, G_{\mathbf{P}}\}$

	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	$\{\mathbf{B}\}$	$\{\mathbf{B}, \mathbf{G}, \mathbf{R}\}$
$\{\mathbf{G}\}$	0	1	0	0	0	1	1	1	0	0	$\{\mathbf{G}\}$	$\{\mathbf{B}, \mathbf{G}, \mathbf{O}\}$
$\{\mathbf{O}\}$	0	0	1	0	0	0	1	1	1	1	$\{\mathbf{O}\}$	$\{\mathbf{G}, \mathbf{O}, \mathbf{P}, \mathbf{R}\}$
$\{\mathbf{R}\}$	0	0	0	1	0	1	0	1	1	1	$\{\mathbf{R}\}$	$\{\mathbf{B}, \mathbf{O}, \mathbf{P}, \mathbf{R}\}$
$\{\mathbf{P}\}$	0	0	0	0	1	0	0	1	1	1	$\{\mathbf{P}\}$	$\{\mathbf{O}, \mathbf{P}, \mathbf{R}\}$

Example



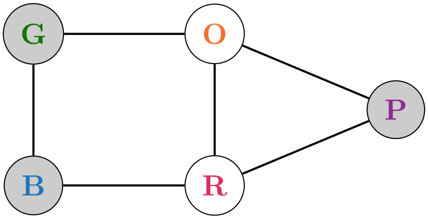
Two variables per node, *e.g.*,

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

Partition: $\mathcal{G} := \{G_v := \{x_v, y_v\} \mid v \in V\}$
$= \{G_{\mathbf{B}}, G_{\mathbf{G}}, G_{\mathbf{O}}, G_{\mathbf{R}}, G_{\mathbf{P}}\}$

🔥	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}
{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}	0	0	0	1	0	1	0	1	1	1	$\varnothing$	{B, O, P}
{P}	0	0	0	0	1	0	0	1	1	1	{P}	{O, P}

Example



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

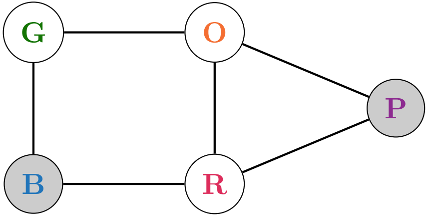
- $x_{\mathbf{B}}$ models  at t_0
- $y_{\mathbf{B}}$ models  at t_1

Partition: $\mathcal{G} := \{G_v := \{x_v, y_v\} \mid v \in V\}$

$= \{G_{\mathbf{B}}, G_{\mathbf{G}}, G_{\mathbf{O}}, G_{\mathbf{R}}, G_{\mathbf{P}}\}$

	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	$\{\mathbf{B}\}$	$\{\mathbf{B}, \mathbf{G}\}$
$\{\mathbf{G}\}$	0	1	0	0	0	1	1	1	0	0	$\{\mathbf{G}\}$	$\{\mathbf{B}, \mathbf{G}\}$
$\{\mathbf{O}\}$	0	0	1	0	0	0	1	1	1	1	$\varnothing$	$\{\mathbf{G}, \mathbf{P}\}$
$\{\mathbf{R}\}$	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	$\{\mathbf{P}\}$	$\{\mathbf{P}\}$

Example



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

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

Partition: $\mathcal{G} := \{G_v := \{x_v, y_v\} \mid v \in V\}$
$= \{G_{\mathbf{B}}, G_{\mathbf{G}}, G_{\mathbf{O}}, G_{\mathbf{R}}, G_{\mathbf{P}}\}$

	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

Results

Size

Largest network ($|V|$):

	encoded	**solved**
SOTA	494	494
gismo	$227\,320$	$21\,363$
improvement	$460\times$	$43\times$

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

Results

Size

Largest network ($|V|$):

	encoded	**solved**
SOTA	494	494
gismo	$227\,320$	$21\,363$
improvement	$460\times$	$43\times$

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

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

Results

Size

Largest network ($|V|$):

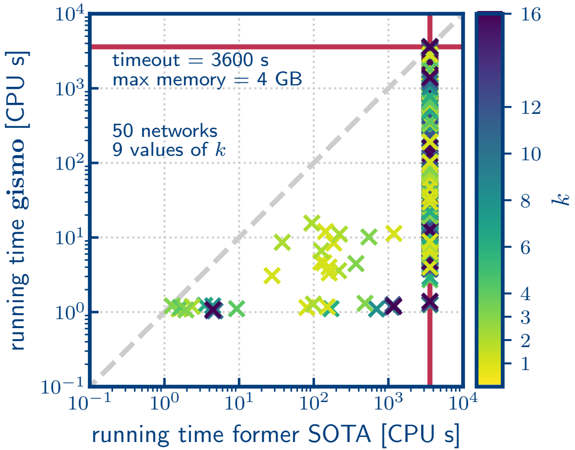
	encoded	solved
SOTA	494	494
gismo	$227\,320$	$21\,363$
improvement	$460\times$	$43\times$

SOTA: $k = 1$

gismo: for all tested k.

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

Time



Solving the Identifying Code Set Problem with Grouped Independent Support

Authors: **Anna L.D. Latour**, Arunabha Sen and Kuldeep S. Meel
Presented at: IJCAI 2023 and **ModRef @ CP 2023**

full paper

www.ijcai.org/proceedings/2023/219

gismo

github.com/meelgroup/gismo

more info

www.annalatour.nl/publications

Solving the Identifying Code Set Problem with Grouped Independent Support

Authors: **Anna L.D. Latour**, Arunabha Sen and Kuldeep S. Meel
Presented at: IJCAI 2023 and **ModRef @ CP 2023**

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

full paper	**gismo**	**more info**
www.ijcai.org/proceedings/2023/219	github.com/meelgroup/gismo	www.annalatour.nl/publications

References I

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.

Padoa, A. (1901). “Essai d’une théorie algébrique des nombres entiers, précédé d’une introduction logique à une theorie déductive quelconque”. In: *Bibliothèque du Congrès international de philosophie* 3, 309—365.

Background: Independent Support

Independent Support (Chakraborty, Fremont, Meel, Seshia, and Vardi 2014)

Given a Boolean formula $F(X)$ on Boolean variables X.
A set $I \subseteq X$ is an *independent support* of F if the following holds:

$\sigma := X \mapsto \{0,1\}$ is an assignment of truth values to variables X

$$\forall\, \sigma_1, \sigma_2 \in Sol(F) \,.\, \left(\sigma_{1\downarrow I} = \sigma_{2\downarrow I}\right) \Rightarrow \left(\sigma_1 = \sigma_2\right)$$

set of solutions of $F(X)$

solution σ_2 projected on $I \subseteq X$

Grouped Independent Support: more formally

Grouped Independent Support

Given a Boolean formula $F(Z, A)$ with $Z \cap A = \varnothing$ and a partitioning $\mathcal{G}$ of Z into non-empty sets. The subset $\mathcal{I} \subseteq \mathcal{G}$ is a *grouped independent support (GIS)* of $\langle F, \mathcal{G} \rangle$ if the following holds:

solution σ_2 projected on $sup(\mathcal{I}) \subseteq Z$

$$\forall \sigma_1, \sigma_2 \in Sol(F). \left(\sigma_{1\downarrow sup(\mathcal{I})} = \sigma_{2\downarrow sup(\mathcal{I})}\right) \Rightarrow \left(\sigma_{1\downarrow Z} = \sigma_{2\downarrow Z}\right)$$

the *support* of $\mathcal{I}$ is $sup(\mathcal{I}) := \bigcup_{G \in \mathcal{I}} G$

solution σ_2 projected on Z

The gismo Algorithm (more detail)

Input: Formula $F(Z)$ with partitioning of Z into $\mathcal{G}$, a time limit τ.
Output: GIS $\mathcal{I} \subseteq \mathcal{G}$.

$E \leftarrow \{e_i \mid z_i \in Z\}$
Initialise $\phi\left(Z, A, \hat{Z}\right)$
$Q \leftarrow sup\left(\mathcal{G}\right)$
$\mathcal{I} \leftarrow \varnothing$
for $G \in \mathcal{G}$ **do**
    $Q \leftarrow Q \setminus G$
    $C \leftarrow Q \cup sup\left(\mathcal{I}\right)$
    $\xi \leftarrow \bigwedge_{z_i \in C}^{m} e_i$
    for $z \in G$ **do**
        $\psi \leftarrow \phi \wedge \xi \wedge z \wedge \neg \hat{z}$
        $sat \leftarrow \text{CHECKSAT}(\psi, \tau)$
        if sat **then**
            $\mathcal{I} \leftarrow \mathcal{I} \cup \{G\}$
            break
return $\mathcal{I}$

How to check for definability?

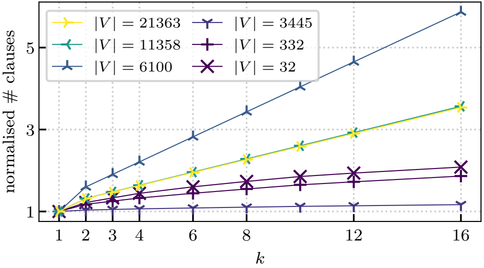
Define indicator variables $E := \{e_i \mid z_i \in Z\}$.
Define the invariant $C \leftarrow Q \cup sup\,(\mathcal{I})$ (always a GIS of $\langle F(Z), \mathcal{G}\rangle$)
Define $\xi := \bigwedge_{z_i \in C}^{m} e_i$.

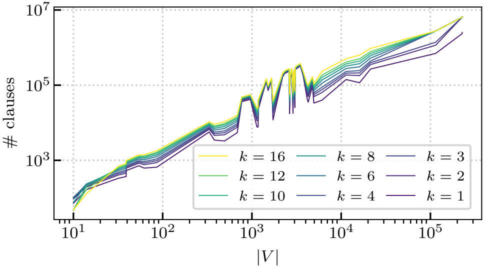
Capture definability (Padoa 1901):

$$\psi\left(Z, \hat{Z}, z\right) := F(Z) \wedge F\left(Z \mapsto \hat{Z}\right) \wedge \bigwedge_{j=1}^{m} e_j \rightarrow (z_j \leftrightarrow \hat{z}_j) \wedge \xi \wedge z \wedge \neg \hat{z}.$$

- Copy of Z: $\hat{Z}$
- Testing z: z and $z \wedge \neg \hat{z}$
- Original formula: $F(Z)$
- Copy of formula with fresh variables: $F\left(Z \mapsto \hat{Z}\right)$
- Forces $z_j = \hat{z}_j$ if True: e_j
- Forced equalities: ξ

How do the CNF models scale?





How do the CNF models compare to the ILP models?