Searching for Universal Graphs with SAT

April, 2026

Tianwei Zhang[1] Stefan Szeider[1] Carlos Ansótegui[2]

[1]TU Wien, Algorithm and Complexity Group

[2]Universitat de Lleida, Lleida, Spain

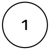
Overview

Smallest k-Universal Graphs

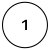
- A k-Universal graph is a graph that contains all graphs with k vertices as **induced subgraphs**.
- The notion was first introduced by German mathematician **Rado** [Rado, 1964].
- Fix a k, we want know the **smallest** size $f(k)$ of a k-universal graph and **the number of** such graphs $F(k)$ modulo isomorphisms.

Smallest 1-Universal Graphs

$$\mathcal{G}(1) = 1$$

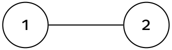


$$f(1) = 1, F(1) = 1$$

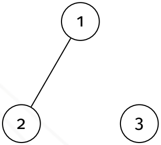


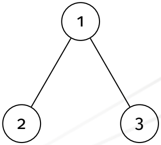
Smallest 2-Universal Graphs

$$\mathcal{G}(2) = 2$$



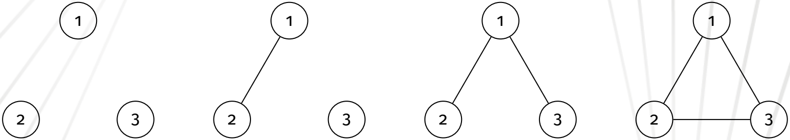
$$f(2) = 3, F(2) = 2$$



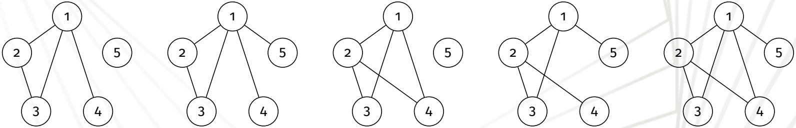


Smallest 3-Universal Graphs

$$\mathcal{G}(3) = 4$$

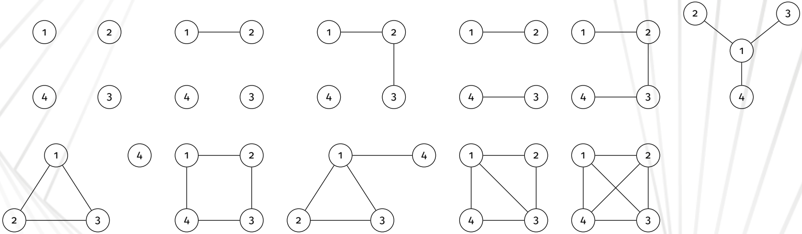


$$f(3) = 5, F(3) = 5$$



Smallest 4-Universal Graphs

$$\mathcal{G}(4) = 11$$



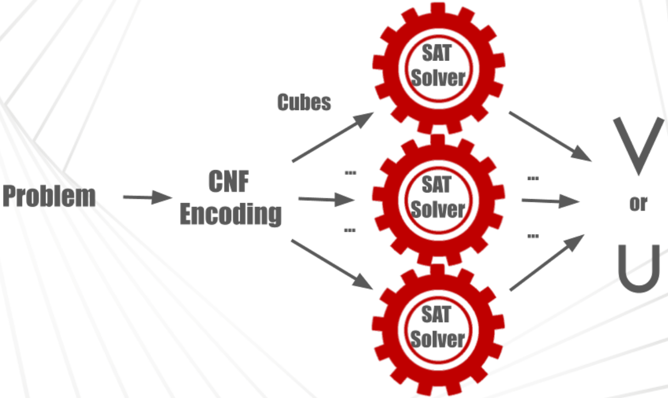
$$f(4) = 8, F(4) = 438$$

Exponential Explotion

- As k increases, both the **number of pattern** graphs and potential **target graphs** grow drastically, quickly beyond the capability of the **generate-and-test** method.
- $\mathcal{G}(1) = 1$, $\mathcal{G}(2) = 2$, $\mathcal{G}(4) = 11$, $\mathcal{G}(5) = 34$, $\mathcal{G}(6) = 156$, $\mathcal{G}(7) = 1044$, ... $\mathbf{G}(\mathbf{17})$ is a number with **27 digits**! [OEIS Foundation Inc., , A000088]
- Luckily, we have powerful **modern SAT solvers** to help.

Method Overview

For each k, we start from a **theoretical lower bound**, and search for a k-universal graph of size n with **increasing** n until we find a universal graph.



Result Overview

A table of known values for $f(k)$ and $F(k)$. The two numbers in **bold** are our **contribution**. Here, $\mathcal{G}(k)$ is the setof graphs of size k modulo isomorphism.

k	$\lvert\mathcal{G}(k)\rvert$	$f(k)$	$F(k)$
1	1	1	1
2	2	3	2
3	4	5	5
4	11	8	438
5	34	10	22
6	156	14	**264662**
7	1044	**18**	–

Theoretical Lower Bound

Lower Bound

Theorem

$f(k) \geq 2k-1$ for all k and $f(k) \geq 2k + 2\lfloor \frac{k}{2} \rfloor - 4$ for all $k \geq 4$.

Corollary

A k universal graph with a disjoint pair of K_k and I_k has at least $2k$ vertices for all k and $f(k) \geq 2k + 2\lfloor \frac{k}{2} \rfloor - 3$ vertices for all $k \geq 4$.

Corollary

$f(1) \geq 1$, $f(2) \geq 3$, $f(3) \geq 5$, $f(4) \geq 8$, $f(5) \geq 10$, $f(6) \geq 14$ and $f(7) \geq 16$.

Proof.

- A universal graph must contain K_k and I_k, and these two overlap by one at most. Thus $f(k) \geq 2k - 1$.
- Define $H := K_{\lfloor \frac{k}{2} \rfloor, \lfloor \frac{k}{2} \rfloor + (k \bmod 2)}$. When $k \geq 4$, H is distinct from K_k and I_k. There is a clique in H that does not intersect K_k. This clique intersects I_k by at most one vertex. This means a clique of size $\lfloor \frac{k}{2} \rfloor - 1$ is disjoint with I_k and K_k. Symmetrically, considering $\overline{H}$ gives us that an independent set of size $\lfloor \frac{k}{2} \rfloor - 1$ is disjoint with I_k and K_k. The latter clique and independent set intersect by at most one vertex. QED.

□

Encoding

Formula

To look for a k-universal graph with size n:

$$\mathrm{Uni}_n^k(\mathcal{H}, \mathrm{Temp}) := \bigwedge_{H \in \mathcal{G}(k)/\mathcal{H}} \Big(\mathrm{IndSub}(H, n) \wedge \mathrm{SymBreak}(H, n)\Big) \wedge \mathrm{Fix}(\mathrm{Temp})$$

- $\mathrm{IndSub}(H, n)$ says that pattern graph H is an **induced subgraph** in the target graph of size n.
- $\mathrm{SymBreak}(H, n)$: imposes a **partial order** among the vertices in H to break symmetries.
- $\mathcal{H}$ a small selection of k graphs. $\mathrm{Fix}(\mathrm{Temp})$ **hardcodes** these graphs according to a template Temp.

Variables

- Top-right half of the $n \times n$ adjacency matrix for G with entries $e_{i,j}$ for all $1 \leq i < j \leq n$.
  - ✓ $e_{i,j}$ is true when there is an edge between i and j.
- For each $H \in \mathcal{G}(k)/\mathcal{H}$: a binary $k \times n$ matrix with entries $m_{H,i,j}$.
  - ✓ $m_{H,i,j}$ is true when vertex i in H maps to vertex j in G.
- Auxiliary variables.

IndSub(H, n)

IndSub(H, n) express that the $k \times n$ matrix m_H is an **injective mapping** that **respects edges/non-edges**.

- **Exactly one** $m_{H,i,j}$ is true in each **row**.
- **At most one** $m_{H,i,j}$ is true in each **column**.
- If both $m_{H,i,j}$ and $m_{H,i',j'}$ are true, then there is an edge between j and j' in G if and only if there is an edge between i and j' in H.

Ladder Encoding for Cardinality

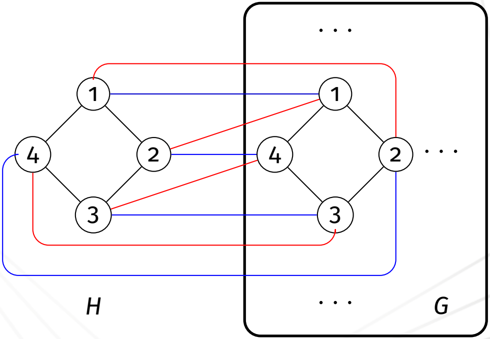
We use the **ladder encoding** [Ansótegui and Manya, 2004, Gent and Nightingale, 2004] for the cardinality constraints 'exactly one' and 'at most one'.

- For both constraints over $x_1, x_2, \ldots, x_m$, we use **auxiliary variables** $y_1, y_2, \ldots, y_m$ and encode them so that y_i is true if and only if exactly one of $x_1, x_2, \ldots, x_i$ is true.
- This encoding has size **linear** to the number of input, and **propagates well** for CDCL solvers.
- the auxiliary variables here also comes into play in symmetry breaking.

Symmetry Breaking

Symmetries in Mappings

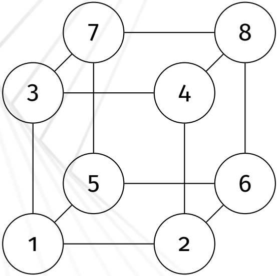
Symmetries arise when a pattern graph has non-trivial automorphisms. This causes **different mappings** to result in **the same** induced subgraph.



Partial Order

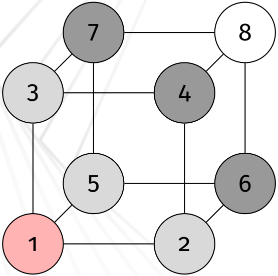
- To remove these symmetries, we impose a **partial order** over the vertices of a pattern graph.
- We write $u < v$ to mean that vertex u is mapped to a **smaller** vertex in G than v does.
- $a_{H,i,j}$ are the **auxiliary variables** used in the ladder encoding for the cardinality constraint that exactly $m_{H,i,j}$ in each row of m_H is true.
- $a_{H,i,j}$ is true if and only if vertex i in H maps to vertices among $1, 2, \ldots, j$ in G.
- For vertices u and v in H, $u < v := \bigwedge_{j \in [n]} a_{H,v,j} \longrightarrow a_{H,u,j}$.

Partial Order: Example



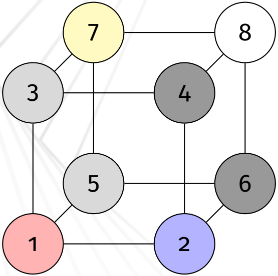
- Starting orbits: $(1\,2\,3\,4\,5\,6\,7\,8)$
- Fix 1.
  Add $1 < 2, 1 < 3, 1 < 4, 1 < 5, 1 < 6, 1 < 7, 1 < 8$.
  Orbits: $(1)\ (2\,3\,5)\ (4\,6\,7)\ (8)$.
- Fix 2.
  Add $2 < 3, 2 < 5$.
  Orbits: $(1)\ (2)\ (3\,5)\ (4\,6)\ (7)\ (8)$.
- Fix 3.
  Add $3 < 5$. Orbits: $(1)\ (2)\ (3)\ (4)\ (5)\ (6)\ (7)\ (8)$.
  Ending orbits.
- Apply transitive reduction on the collection of inequalities to get the final result $\{1 < 2, 2 < 3, 3 < 5\}$

Partial Order: Example



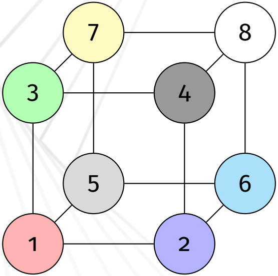
- Starting orbits: $(1\,2\,3\,4\,5\,6\,7\,8)$
- Fix 1.
  Add $1 < 2, 1 < 3, 1 < 4, 1 < 5, 1 < 6, 1 < 7, 1 < 8$.
  Orbits: $(1)\ (2\,3\,5)\ (4\,6\,7)\ (8)$.
- Fix 2.
  Add $2 < 3, 2 < 5$.
  Orbits: $(1)\ (2)\ (3\,5)\ (4\,6)\ (7)\ (8)$.
- Fix 3.
  Add $3 < 5$. Orbits: $(1)\ (2)\ (3)\ (4)\ (5)\ (6)\ (7)\ (8)$.
  Ending orbits.
- Apply transitive reduction on the collection of inequalities to get the final result $\{1 < 2, 2 < 3, 3 < 5\}$

Partial Order: Example



- Starting orbits: (1 2 3 4 5 6 7 8)
- Fix 1.
  Add $1 < 2, 1 < 3, 1 < 4, 1 < 5, 1 < 6, 1 < 7, 1 < 8$.
  Orbits: (1) (2 3 5) (4 6 7) (8).
- Fix 2.
  Add $2 < 3, 2 < 5$.
  Orbits: (1) (2) (3 5) (4 6) (7) (8).
- Fix 3.
  Add $3 < 5$. Orbits: (1) (2) (3) (4) (5) (6) (7) (8).
  Ending orbits.
- Apply transitive reduction on the collection of inequalities to get the final result $\{1 < 2, 2 < 3, 3 < 5\}$

Partial Order: Example



- Starting orbits: (1 2 3 4 5 6 7 8)
- Fix 1.
  Add $1 < 2, 1 < 3, 1 < 4, 1 < 5, 1 < 6, 1 < 7, 1 < 8$.
  Orbits: (1) (2 3 5) (4 6 7) (8).
- Fix 2.
  Add $2 < 3, 2 < 5$.
  Orbits: (1) (2) (3 5) (4 6) (7) (8).
- Fix 3.
  Add $3 < 5$. Orbits: (1) (2) (3) (4) (5) (6) (7) (8).
  Ending orbits.
- Apply transitive reduction on the collection of inequalities to get the final result $\{1 < 2, 2 < 3, 3 < 5\}$

Properties of Encoding

This Encoding is

- concise:
  - ✓ We can show that the reduced collection inequalities forms a **forest** over that vertices of H, and therefore contains at most $(k-1)$ inequalities. Therefore the size of $\mathrm{SymBreak}(H, n)$ is at most $n \times (k-1)$;
- complete:
  - ✓ We can also show that for any subset $S \subseteq V(G)$ that induces H, there is a **unique mapping** from H to S that witnesses it.

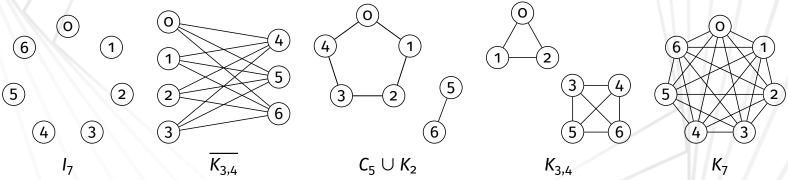
Templates

Templates

- **Fixing** some of the edge variables for G can **divide** the search space so we can run the **easier subproblems** with the SAT solver **in parallel**.
- Problem: number of subproblems grow **exponentially** as the number of fixed variables increase.
- Solution: consider scenarios of how **highly symmetric** pattern graphs **overlap** with each other in the patter graph. Better if the pattern graphs can overlap with each other **as little as possible**.
- For example, there are only **two ways** K_k and I_k can overlap and they can only overlap by **at most one vertex**.
- Fix these pattern graphs in the target graph. The result of this is a graph that does not have every possible edge/non-edge defined. We call such a graph a **template**.

Symmetric Pattern Graphs

- $k \leq 6$: K_k and I_k.
- $k = 7, n = 16$: add $K_3, 4$ and $\overline{K_{3,4}}$.
- $k = 7, n = 17, 18$: add $C_5 \cup K_2$.



Generating Templates

- See templates as graph with **two edge coloring** (edge/non-edge as two different colors) and use the **canonical 2-layered** translation to normal graphs.
- Encode the condition that each **selected pattern graph** is an induced subgraph in the template in a similar way, with the additional constraint that the edges within these induced subgraphs are **the only edges defined**.

Generating Templates (Cont.d)

- Give the encoding to **SMS** (SAT Modulo Symmetries) [Kirchweger and Szeider, 2021], a tool developed specifically for enumerating graphs without isomorphic copies.
- The encoding is **small enough** (only a couple of pattern graphs) and the solutions are **few enough** (highly symmetric pattern graphs) that SMS generates all non-isomorphic templates quickly.
- **Amazingly**, for $k = 7$, $n = 17$, we have only 68950 templates even though on average about 85 edges are defined in each template.

Results

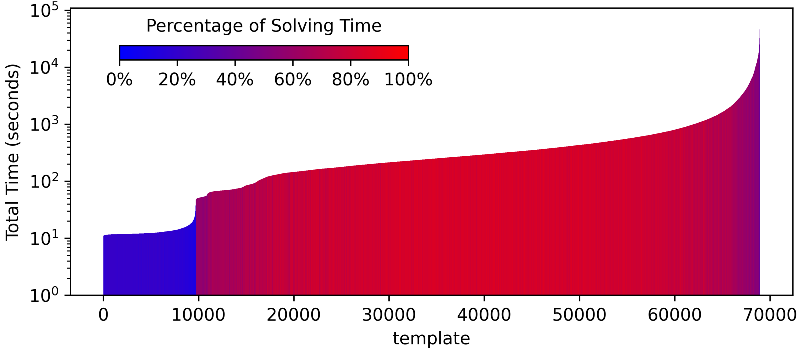
Computational Time

Time spent in computing $F(k)$ when $n = f(k)$ and in showing $n < f(k)$ otherwise.

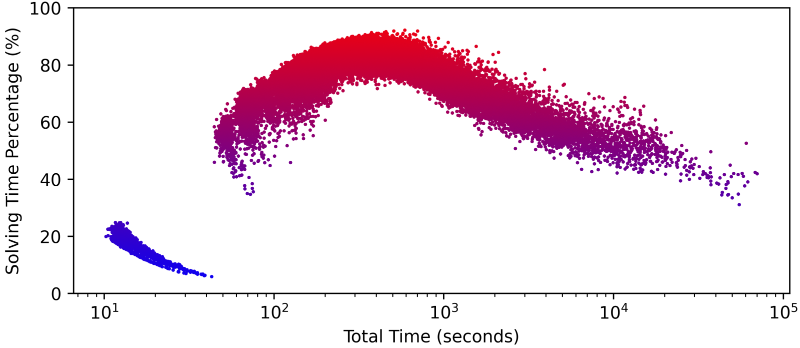
k	1	2	3	4	5	6	7		
n	1	3	5	8	10	14	16	17	18
answer	yes	yes	yes	yes	yes	yes	no	no	yes
#temp.	1	1	1	1	1	1	40	68950	172*
time	$<$ 1ms	78ms	95ms	1.45s	0.52*s*	4.1d	3.6*h*	1.3*y*	2.7*d**

**For $k = 7$, $n = 18$, the complete $F(8)$ is not computed. The time listed here is for showing $f(7) \leq 18$ is the number of cubes times the shorted time in which the solver solves the cube with a positive answer. The number of templates is also reduced under the assumption of a pair of disjoint K_k and I_k.

Time Distribution for $k = 7$



Time Distribution for $k = 7$ (Cont'd)



References

Ansótegui, C. and Manya, F. (2004).
Mapping problems with finite-domain variables to problems with boolean variables.
In *International conference on theory and applications of satisfiability testing*, pages 1–15. Springer.

Gent, I. P. and Nightingale, P. (2004).
A new encoding of alldifferent into sat.
In *International Workshop on Modelling and Reformulating Constraint Satisfaction*, volume 3, pages 95–110.

Kirchweger, M. and Szeider, S. (2021).
SAT modulo symmetries for graph generation.
In Michel, L. D., editor, *27th International Conference on Principles and Practice of Constraint Programming (CP 2021)*, Leibniz International Proceedings in Informatics (LIPIcs), page 39:1–39:17. Dagstuhl.

OEIS Foundation Inc.
The On-Line Encyclopedia of Integer Sequences.
Published electronically at `http://oeis.org`.

Rado, R. (1964).
Universal graphs and universal functions.
Acta Arithmetica, 9(4):331–340.