Kolmogorov Complexity of Graph Classes

Ethan Mader   Zachary Lee

Purdue University

Spring 2025

Introduction

- For a graph G, we encode its adjacency matrix as binary string x_G; its Kolmogorov complexity is $K(G) := K(x_G)$.
- Main question: how does graph structure impact compressibility?
- We aim to provide upper bounds for the asymptotic Kolmogorov complexity of graph classes in terms of graph-theoretic properties
- Few previous results use conditional complexity - we define our own model

Preliminaries

Graph Class

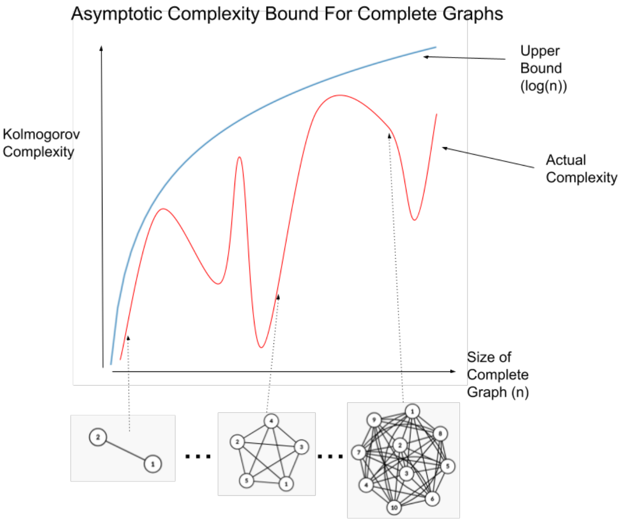
Let G^* be all simple undirected graphs. $X \subseteq G^*$ is a *graph class* if for all n, there exists $G \in X$ such that $|V(G)| \geq n$.

Complexity of Graph Classes

$X \in K\big(f(n)\big)$ if there exists $c > 0$, $n_0 \in \mathbb{N}$ such that $\forall G \in X$ where $|V(G)| > n_0$ we have $K(G) \leq cf(n)$.

Labeled vs. unlabeled setting: $K(G) = \min_\pi K(x_{\pi(G)})$ in the unlabeled case which is the setting we consider.

Complexity of a Graph Class



Trivial Facts

Every graph class X satisfies $X \in K(n^2)$ since $|x_G| = n(n-1)/2$.

Edge Bound

Let X be a graph class such that for all $G \in X$ where $|V(G)| = n$, we have $|E(G)| = O(f(n))$ for some $f : \mathbb{N} \to \mathbb{N}$. Then $X \in K(f(n)\log(n))$.

The $\log(n)$ factor comes from the indices needed to denote each edge.

Corollary (Bounded Degree)

If X is a graph class and there exists d such that for all $G \in X$ we have $\Delta(G) \leq d$, then $X \in K(n\log(n))$.

Basic Graph Classes

Definition

A graph class X is *basic* if $X \in K(\log(n))$.

Theorem (Compression by Enumeration)

X is basic if the following hold:
(1) There is a recognizer R for $L_X = \{x_G : G \in X\}$.
(2) $|\{G \in X : |V(G)| = n\}| = O(n^k)$ for some k

Examples: complete graphs K_n, cycles C_n, paths, wheels, stars, k-edge graphs, $\sqrt{n} \times \sqrt{n}$ grids, complete binary trees, 0101 graphs.

Set Operations on Graph Classes

Let X_1, X_2 be graph classes with $X_1 \in K(f_1(n))$, $X_2 \in K(f_2(n))$.

- **Subset:** If $Y \subseteq X_1$ then $Y \in K(f_1(n))$.
- **Union:** $X_1 \cup X_2 \in K(\max f_1(n), f_2(n))$.
- **Intersection:** $X_1 \cap X_2 \in K(\min f_1(n), f_2(n))$.
- **Complement:** If $Y = \{G : \bar{G} \in X_1\}$, then $Y \in K(f_1(n))$.

These follow from basic definitions.

Encoding Trees in $O(n)$ via DFS Traversal

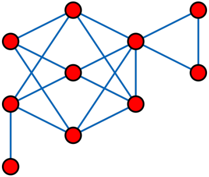
Encoding Scheme

We represent a rooted tree using a binary string of length $2(n-1)$, where:

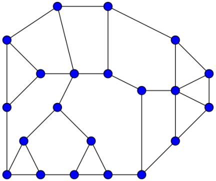
- 1 means "create a new child and move to it"
- 0 means "return to the parent"

This encodes a preorder traversal that visits each edge twice (once forward, once back).

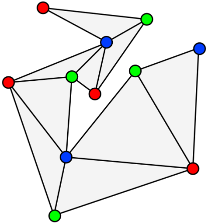
Bounded Clique-Width



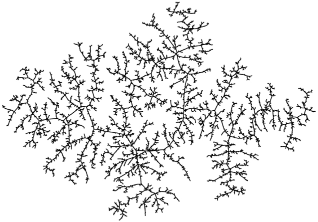
(a) Distance hereditary graph



(b) Halin graph



(c) Outer-planar graph



(d) Tree

Bounded Clique-Width

Definition

Define the following operations over labels $[1, c]$:

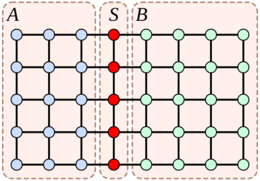
1. $i(v)$ — create vertex v labeled i.
2. $G_1 \oplus G_2$ — disjoint union.
3. $\eta_{i,j}$ — connect every i-labeled vertex to every j-labeled vertex.
4. $\rho_{i \to j}$ — relabel i to j.

The clique-width of G, $cwd(G)$, is the smallest c where we can produce G.

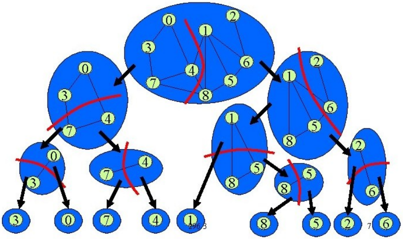
Theorem

If there exists constant c such that $cwd(G) \leq c$ for all $G \in X$, then $X \in K(n)$.

Separation Number



(a) Separator of grid



(b) Recursive application of separator

Separation Number

Definition (Separator)

$S \subseteq V$ is an (s, α)-separator if $V = S \sqcup A \sqcup B$, $|S| \leq s$, $|A|, |B| \leq \alpha n$, $E(A, B) = \emptyset$.

Separation number $s(G)$: minimum s such that every induced subgraph of G admits an $(s, \frac{2}{3})$-separator.

Theorem

If there exists $s : \mathbb{N} \to \mathbb{N}$ such that for all $G \in X$ where $|V(G)| = n$, we have $s(G) = O(s(n))$, then $X \in K(s(n) n \log^2(n))$.

Many corollaries, but we may lose tightness of the edge bound.

Graph Products

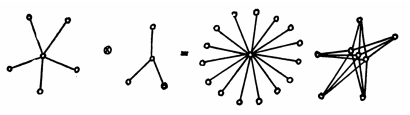
Definition

The *graph product* $G = G_1 \circ_e G_2$ has vertex set $V_1 \times V_2$, with an edge between (a, b) and (a', b') if a computable function e returns 1 on that pair.

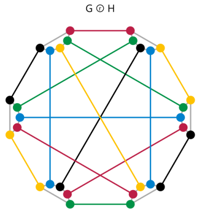
Theorem

Let $X_1 \in K(f_1(n))$ and $X_2 \in K(f_2(n))$ be graph classes, and $\circ_e$ a graph product computed by a Turing machine with $O(1)$ description length with access to both adjacency matrices. Then

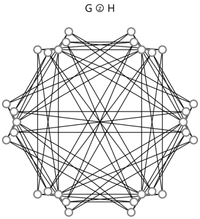
$$X_1 \circ_e X_2 \in K(f_1(n) + f_2(n)).$$



(a) Kronecker product



(b) Replacement product



(c) Zig-zag product

Graph Sums

The graph sum produces a copy of each graph and specifies edges between them via a computable function $e : (V_1 \sqcup V_2)^2 \to \{0, 1\}$.

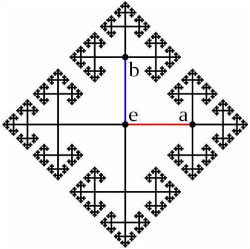
Theorem

Let $X_1, X_2 \in K(f_1(n)), K(f_2(n))$ respectively, and $+_e$ a graph sum computed by a Turing machine with $O(1)$ description length. Then

$$X_1 +_e X_2 \in K(f_1(n) + f_2(n)).$$

Conclusion

Cayley Graph



- Obtained reasonable bounds for sparse and bounded clique-width graphs
- Developed tools for proving more upper bounds
- Include constants instead of $O(f(n))$?
- Fully classify basic graphs?
- Directed graphs or hypergraphs?
- Other graph parameters?
- Complexity via spectral graph theory (structured eigenvectors)?

References

- *Kolmogorov Basic Graphs and Their Application in Network Complexity Analysis* (Farzaneh, Coon, Badiu, 2021)
- *Kolmogorov Complexity of Graphs* (John Hearn, 2006)
- *Kolmogorov Random Graphs and the Incompressibility Method* (Buhrman, Li, Tromp, Vitányi, 1999)
- *Critical Properties of Graphs of Bounded Clique-width* (Lozin, Milanic, 2013)
- *Bandwidth, Expansion, Treewidth, Separators and Universality for Bounded-degree Graphs* (Böttcher, Pruessmann, Taraz, Würfl, 2010)
- *Information System on Graph Classes and their Inclusions* (H.N. de Ridder et al., 2001-2014)