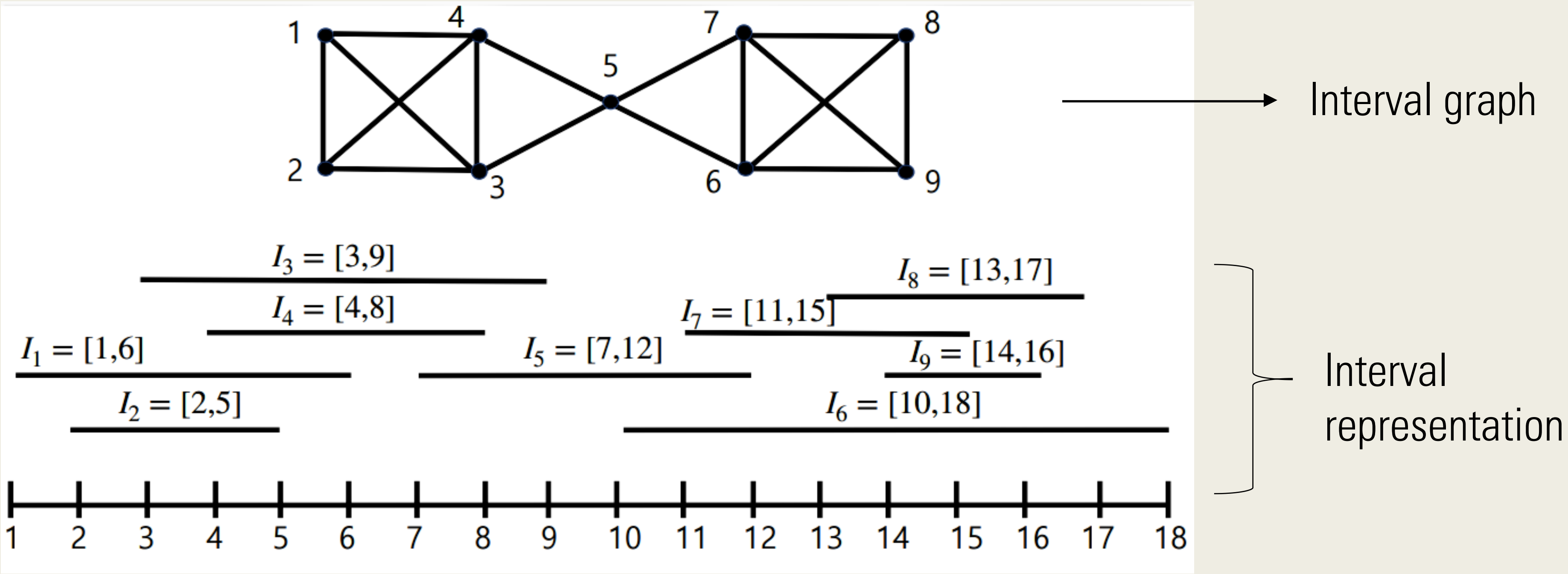
COMPACT REPRESENTATION OF INTERVAL GRAPHS OF BOUNDED DEGREE AND CHROMATIC NUMBER

Sankardeep Chakraborty
The University of Tokyo
Japan

Seungbum Jo
Chungnam National University
South Korea

Set up

A simple undirected graph G is called an Interval graph if (a) with every vertex we can associate a closed interval on the real line, and (b) two vertices share an edge if and only if the corresponding intervals are not disjoint.



Set up

Given a set T consisting of combinatorial objects with certain property, a data structure Z is called

- Succinct if Z can store any arbitrary member x from T using $\log(|T|) + o(\log(|T|))$ bits, OR
- Compact if Z can store any arbitrary member x from T using $O(\log(|T|))$ bits,

along with fast query support.

There already exist succinct/compact data structures for various combinatorial objects like arbitrary graphs, planar graphs, trees, deterministic finite automata, permutations, equivalence classes and many more.

Prior Work

- Acan et al. (Algorithmica 2021) proposed a succinct data structure for storing and navigating interval graphs.
- More specifically, given an unlabeled interval graph G with n vertices, they first show that at least n log n - 2n log log n - O(n) bits are necessary to represent G.
- This is followed by a matching data structure consuming n log n + O(n) bits of space with constant time queries i.e., degree, adjacency and neighborhood.

Can we design efficient data structures whose space consumption beats the information-theoretic lower bound under some bounded parameter condition?

Our Main Results

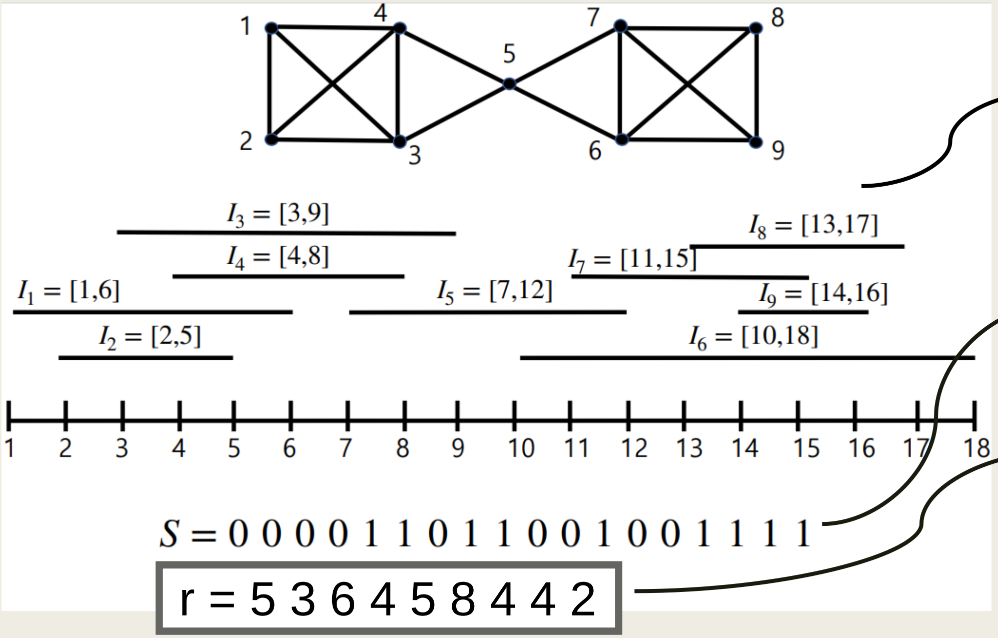
- When the maximum degree of any interval graph is bounded by k, we show that there exists an $(n \log k + O(n))$-bit data structure. Thus, our data structure surpasses the information-theoretic lower bound when $k = O(n^\varepsilon)$ where $0 < \varepsilon < 1$.

- We augment the upper bound result by giving an explicit $((1/6)\, n \log k - O(n))$-bit enumerative lower bound. Our result provides counting lower bound by taking into consideration maximum degree as a parameter for interval graphs for the first time in the related literature.

- Finally, we consider interval graphs with bounded chromatic number p, and here, we design a $(p-1)n + o(pn)$ bit data structure with efficient navigational query support. Thus, our data structure surpasses the information-theoretic lower bound when $p = o(\log n)$.

Upper Bound

This upper bound follows from the result of Acan et al.'s (Algorithmica 2021) result in a straightforward manner.



Every interval has distinct start and end point. Overall, for n intervals, all the endpoints make up 2n distinct integers from 1 to 2n without loss of generality.

S is a bit string of size 2n bits with 0s in starting locations and 1 at the ending locations.

r stores difference between end point and start point of the intervals starting from left to right.

With additional o(n) bit structures (Rank and Select), it is possible to store r and S using n log n + O(n) bits such that degree/adjacency/neighborhood queries can be supported.

As the maximum degree of our input interval graph is bounded by k, in total the data structure consumes (n log k + O(n))-bit along with supporting fast queries.

Is this optimal?

Lower Bound

For any interval graph G with n vertices, if the maximum degree of G is k, at least ((1/6) n log k - O(n))-bits are necessary to represent G.

Proof Idea:

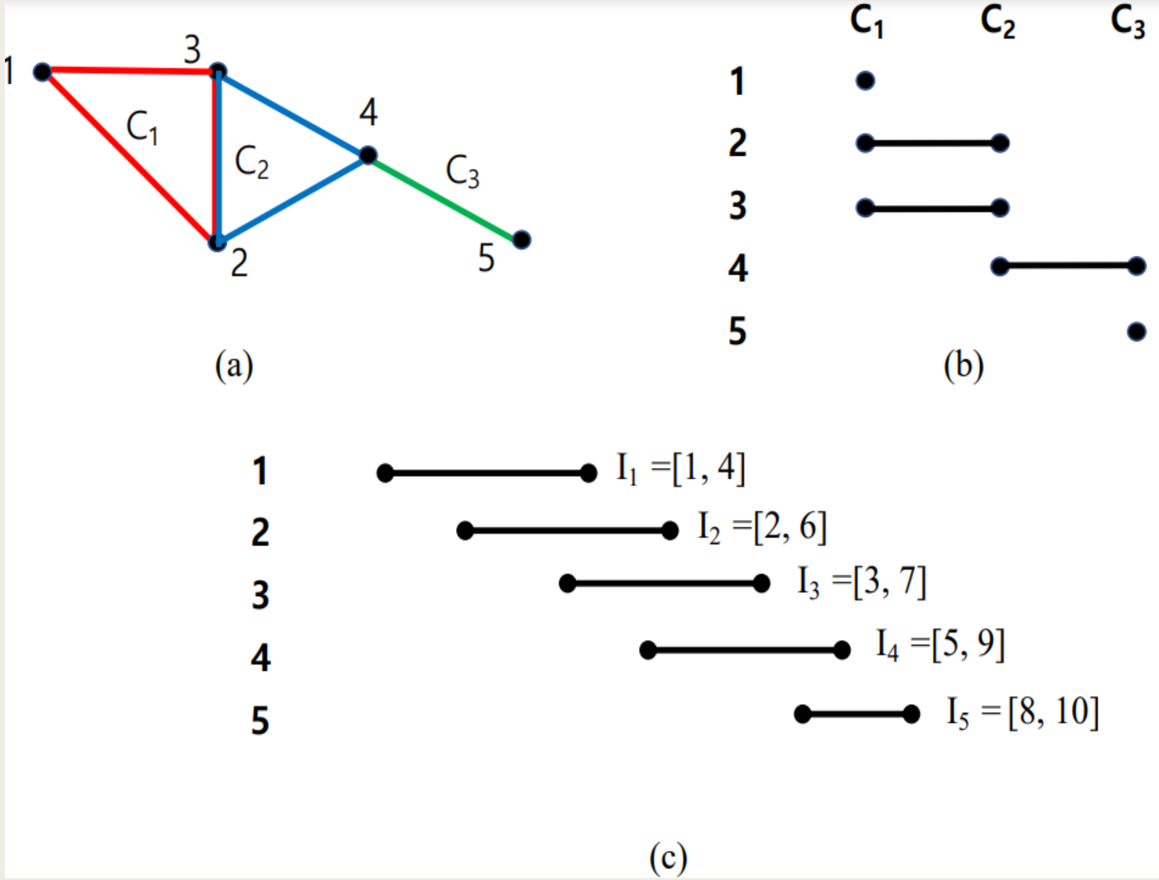
- Let T be a set of all non-isomorphic interval graphs with n vertices where for each graph G in T, its interval representation satisfies (i) all the starting and endpoints of the intervals are distinct, and (ii) the maximum degree of G is at most k. Then |T| gives our desired lower bound.

- We first obtain an interval representation from the interval graph using *bundle hypergraph.*

- For an interval graph G, let C be a set of all the inclusion-maximal cliques in G. Also, let B_v be the *bundle* at vertex v, which is a set of maxcliques in C containing v. Then the bundle hypergraph $\Delta = (C, E(\Delta))$ is a hypergraph where its hyperedges are the bundles of G. i.e., $E(\Delta) = \{B_v \mid v \text{ in } V\}$.

Proof Idea

- It is known that if G is an interval graph, one can define an ordering among the maxcliques in C to satisfy the property that every hyperedge of Δ consists of consecutive maxcliques in C. Thus, by denoting the i-th maxclique in C as C_i, one can define the interval representation of G as for each v as $I_v = [i, j]$ if $B_v = \{C_i, C_{i+1}, \ldots, C_j\}$.



- Note that multiple intervals can share same end points, but we can easily make the endpoints distinct by changing the shared endpoints as consecutive integers.

- We say B_u and B_v *overlap* if $B_u \cap B_v \neq \{B_u, B_v, \Phi\}$. Similarly, Δ is called *overlap connected* if for any two hyperedges B_u and B_v, there exists a sequence S of hyperedges from B_u to B_v where any two consecutive hyperedges in S are overlapped.

Proof Idea

(Kobler et al. SICOMP 2011) showed the following

- there exists a bijection between the set of all non-isomorphic interval graphs and the set of all non-isomorphic bundle hypergraphs, and
- if Δ is overlap connected, there exist at most two minimal interval representations of Δ, C and D, where D is a mirror image of C.

- Thus, by counting the number of distinct minimal interval representation whose corresponding bundle hypergraph Δ is overlap connected, we can obtain the lower bound of the number of non-isomorphic interval graphs.

Constructive proof

((1/6) n log k - O(n)) bits are needed when maximum degree is bounded by k.

In the extended version of our paper, we show similar results can be obtained for circular-arc graphs when parameterized by maximum degree and chromatic number.

Conclusion

- All our data structures are compact. Can we make them succinct?
- Which parameter would give better compression than the ones we considered here?
- Systematic study of parameterized data structures for combinatorial objects.

Thank you for your attention