Note on enumeration of labeled split graphs

Vladislav Bína, Jiří Přibil

Abstract. The paper brings explicit formula for enumeration of vertex-labeled split graphs with given number of vertices. The authors derive this formula combinatorially using an auxiliary assertion concerning number of split graphs with given clique number. In conclusion authors discuss enumeration of vertex-labeled bipartite graphs, i.e., a graphical class defined in a similar manner to the class of split graphs.

Keywords: graph enumeration; labeled graph; split graph

Classification: Primary 05C30; Secondary 05A15

1. Introduction

The paper concerns counts of vertex-labeled split graphs. Unless explicitly stated otherwise, the text deals with "labeled" graphs as a contrary to the "unlabeled" graphs (or, in other words, we distinguish among the isomorphically equivalent graphs). The choice of labeled graphs is motivated by the first author's efforts in representation of multidimensional probability distributions where the labeled graphs can represent the structure of conditional independence relations among (labeled) variables (see, e.g., Edwards [5] or Bína [4]).

It is clear that for n vertices there are $2^{\binom{n}{2}}$ undirected vertex-labeled graphs. The counts of unlabeled split graphs were derived by Royle [8], but these figures are not of interest for our purpose since we need to count all labeled split graphs (we have labeled variables). Thus Section 3 derives a closed formula for the number of all different labeled split graphs on n vertices.

2. Graph-theoretic preliminaries

Definition 1 (Split graph)**.** An undirected graph $G = (V, E)$ is said to be *split* if there exists a partition of V into two subsets I and K (i.e., $V = I \cup K$) such that I is an independent set and K is a clique.

The edges between the pair of sets K and I are not restricted. For properties see, e.g., Hammer and Simeone [7]. Let us remark that the following rather trivial

DOI 10.14712/1213-7243.2015.112
The research was supported by Grant Agency of the Czech Republic (no. 15-00215S).

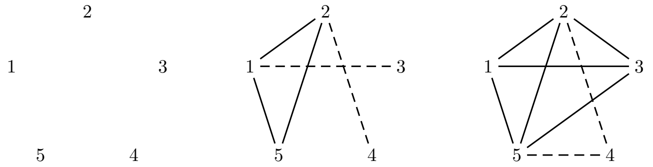

Figure 1. An illustration of simple split graphs on five vertices. Solid lines denote edges in the clique and dotted lines are edges connecting a vertex in the clique with another vertex in the independent set. The split graph on the left has a clique number equal to one, the graph in the middle has a clique number equal to three and the one farthest to the right has a clique number equal to four.

hierarchy of graph classes holds

$$\text{split graphs} \quad \subset \quad \text{chordal graphs} \quad \subset \quad \text{undirected graphs}.$$

Bender et al. [1] showed that a split graph is chordal[1] with chordal complement.

3. Number of labeled split graphs

As we already mentioned, we are interested in counting of labeled split graphs which can be generated on n vertices. But first, let us have a look at some examples of split graphs on five vertices.

We can start the enumerating of split graphs by summation over the size of maximum clique. There is only one split graph with a clique number equal to one (graph without any edge — as shown on the leftmost part of Figure 1). To this single split graph we add a sum of numbers of split graphs with bigger clique numbers. But how many split graphs with clique numbers between two and n are there?

Lemma 2 (Number of split graphs with fixed maximum clique on k vertices)**.** *Let k ($k \leq n$) be the number of vertices that form a fixed maximum clique and let ℓ be the number of vertices that form an independent set. The split graphs on $k+\ell$ vertices are generated by optionally adding edges between vertices of the clique and vertices of the independent set. The number of such split graphs $C_{k,\ell}$ can be computed using the formula*

$$C_{k,\ell} = \left(2^k - 1\right)^{\ell}. \tag{1}$$

Proof: Each vertex in the independent set can be connected with k vertices from the clique. Since its connection with all k vertices produces clique of size

[1] The undirected graph G is *chordal* (or *decomposable* or *triangulated*) if every cycle of a length longer than or equal to four has a chord.

$k+1$, we need to exclude this situation. For each of ℓ vertices from independent set there are $2^k - 1$ possibilities and therefore raising to the power of ℓ gives Formula (1). □

Lemma 3 (Number of split graphs with clique number k)**.** *Let n be the number of vertices, and for any $k \in \{2, \ldots, n\}$ let $N^{\mathrm{S}}_{n,k}$ denote the number of labeled split graphs on n vertices with a clique number of k. Then the following formula holds*

$$N^{\mathrm{S}}_{n,k} = \binom{n}{k} \left(C_{k,n-k} - \sum_{j=1}^{n-k} \frac{j}{j+1} k \binom{n-k}{j} C_{k-1,n-k-j} \right),$$

and the symbols $C_{k,l}$ represent numbers of split graphs with a fixed maximum clique on k vertices and independent set on l vertices, see Formula (1)*.*

Proof: The clique of size k in the graph on n vertices can be chosen $\binom{n}{k}$ ways. If we fix the clique on k vertices, the number of split graphs with a clique number equal to k is $C_{k,n-k}$ (we avoid graphs with a clique number $k+1$).

But in the expression $\binom{n}{k} C_{k,n-k}$ we counted some graphs more than once — the graphs where one or more additional cliques of size k were created. For illustration see Figure 2(a)–(c).

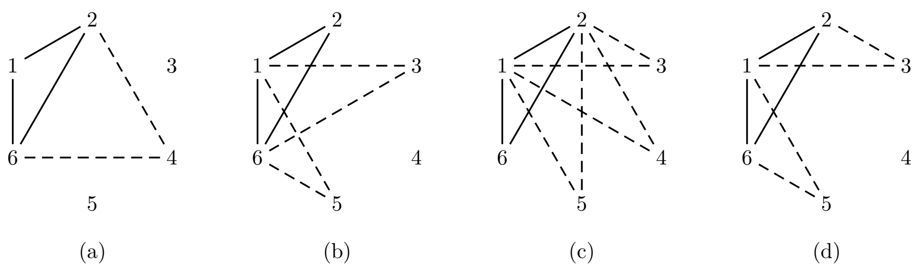

Figure 2. In the left (a) the basic clique of size 3 is generated two times, in (b) the clique of size 3 three times and in (c) we can find it four times. On contrary, in (d) the solid clique of size 3 is generated only once. (Basic clique on vertices $1, 2, 6$ has solid edges; dashed edges connect this clique with the independent set.)

Thus, in the total sum, we need to assign the weight to each situation when a clique of size k appears $(j+1)$-times (by introducing new edges we create j additional cliques of size k). Therefore we set the weight to $\frac{1}{j+1}$. Or, in other words, we need to subtract the repeated situation from resulting sum with a weight of $\frac{j}{j+1}$.

Because these $j+1$ cliques of size k have $k-1$ common vertices, there are $k = \binom{k}{k-1}$ possibilities for choosing $k-1$ vertices from the first primary clique.

And since we choose j vertices out of $n-k$ which are added to $k-1$ common vertices, there are $\binom{n-k}{j}$ such combinations. And there can again be optionally added some more edges. They connect $k-1$ common vertices of all cliques[2] with the remaining $n-k-j$ vertices of the independent set, these are again all split graphs with a fixed clique of size $k-1$ and fixed independent set consisting of $n-k-j$ vertices, their count is $C_{k-1,n-k-j}$. Putting all this together, we need to adjust the total sum by subtracting the term

$$\sum_{j=1}^{n-k} \frac{j}{j+1} k \binom{n-k}{j} C_{k-1,n-k-j}$$

and we can infer that the total number of split graphs with a clique number equal to k is

$$N_{n,k}^{\mathrm{S}} = \binom{n}{k} \left(C_{k,n-k} - \sum_{j=1}^{n-k} \frac{j}{j+1} k \binom{n-k}{j} C_{k-1,n-k-j} \right),$$

where symbol $C_{k,l}$ is defined in Formula (1). □

Theorem 4 (Number of split graphs)**.** *The number of all labeled split graphs on n vertices can be computed using the formula*

$$(2)\quad N_n^{\mathrm{S}} = 1 + \sum_{k=2}^{n} \binom{n}{k} \left(\left(2^k - 1\right)^{n-k} - \sum_{j=1}^{n-k} \frac{j}{j+1} k \binom{n-k}{j} \left(2^{k-1} - 1\right)^{n-k-j} \right).$$

Proof: There is only one split graph with no edges. And for clique number $k \geq 2$ we know from Lemma 3 that there are $N_{n,k}^{\mathrm{S}}$ split graphs with a clique number of k. Summing this up for $2 \leq k \leq n$ we obtain $N_n^{\mathrm{S}} = 1 + \sum_{k=2}^{n} N_{n,k}^{\mathrm{S}}$ and expansion using Lemma 3 and Formula (1) finishes the proof. □

4. Conclusions

The presented paper concerns enumeration of labeled split graphs on a given number of vertices and the main result consists in derivation of explicit formula.

The number of all vertex-labeled split graphs was already published in a form of extended abstract for CTW'11 conference [2] and numerical results together with the formula can also be found in OEIS as a sequence A179534 [3] where the exact counts for up to twenty vertices are given.

Let us mention that the class of bipartite graphs has in a way similar definition. Recall that the vertex set of the split graph can be partitioned into the clique and independent set and vertices from the two sets are optionally connected with edges. Similarly, the vertex set of bipartite graph can be partitioned

[2]We do not set a weight (smaller than 1) when additional edges connect the kth vertex of primary clique (not in the intersection with other cliques) with some of the remaining $n-k-j$ vertices. These graphs are not generated more than once in the entire sum. E.g., Figure 2(d).

into two independent sets such that every edge in bipartite graph connects vertex from one independent set with vertex from the other independent set. To the authors' best knowledge, there does not exist an explicit formula counting vertex-labeled bipartite graphs. But corresponding exponential generating function can be found, e.g., in Wilf's book [10] and recurrence was published by Gebhardt [6], for numerical results see OEIS sequence A047864 [9].

Acknowledgment. The authors thank the anonymous referee for helping them to significantly improve the presentation of the result.

References

[1] Bender E.A., Richmond L.B., Wormald N.C., *Almost all chordal graphs split*, J. Austral. Math. Soc. Ser. A **38** (2) (1985), no. 2, 214–221.

[2] Bína V., *Enumeration of labeled split graphs and counts of important superclasses*, in Proceedings of 10th Cologne-Twente Workshop on Graphs and Combinatorial Optimization (CTW'11), Frascati (2011), pp. 72–75.

[3] Bína V., *Sequence A179534 in The On-Line Encyclopedia of Integer Sequences*, `http://oeis.org/A179534` (2010).

[4] Bína V., *Multidimensional probability distributions: Structure and learning*, Ph.D. Thesis, Faculty of Management in Jindřichův Hradec, Univ. of Economics in Prague, 2011.

[5] Edwards D., Havránek T., *A fast procedure for model search in multidimensional contingency tables*, Biometrika **72** (1985), 339–351.

[6] Gebhardt V., *Computing growth functions of braid monoids and counting vertex-labelled bipartite graphs*, J. Combin. Theory Ser. A **120** (2013), no. 1, 232–244.

[7] Hammer P.L., Simeone B., *The splittance of a graph*, Combinatorica **1** (1981), no. 3, 275–284.

[8] Royle G.F., *Counting set covers and split graphs*, J. Integer Seq. **3** (2000), `https://cs.uwaterloo.ca/journals/JIS/VOL3/ROYLE/royle.html`.

[9] Sloane N.J.A., *Sequence A047864 in The On-Line Encyclopedia of Integer Sequences*, `http://oeis.org/A047864` (1999).

[10] Wilf H.S., *Generatingfunctionology*, Academic Press, San Diego, 1990, p. 80, Equation 3.11.5.

University of Economics in Prague, Faculty of Management in Jindřichův Hradec, Jarošovská 1117/II, 37701 Jindřichův Hradec, Czech Republic

E-mail: bina@fm.vse.cz,
pribil@fm.vse.cz

(*Received* January 21, 2014, *revised* February 4, 2015)