Enumeration of P_4-free chordal graphs

Robert Castelo*
Biomedical Informatics Group (GRIB)
Pompeu Fabra University
Dr. Aiguader 80, 08003 Barcelona
Spain
email: rcastelo@imim.es

Nick Wormald†
Department of Mathematics and Statistics
University of Melbourne, VIC 3010
Australia
email: nick@ms.unimelb.edu.au

Abstract

We count labelled chordal graphs with no induced path of length 3, both exactly and asymptotically. These graphs correspond to rooted trees in which no vertex has exactly one child, and each vertex has been expanded to a clique. Some properties of random graphs of this type are also derived. The corresponding unlabelled graphs are in 1-1 correspondence with unlabelled rooted trees on the same number of vertices.

1 Introduction

A graph is *chordal*, also known as *triangulated*, if it does not contain a chordless cycle on more than three vertices as an induced subgraph. Equations which effectively gave recurrence relations for counting labelled chordal graphs were derived in [11]. A graph is a *comparability* graph if its set of edges admits a transitive orientation. The class of graphs we enumerate in this paper corresponds to a subclass of chordal comparability graphs, which was described by Golumbic [4] in terms of forbidden subgraphs as follows.

Let P_4 denote a graph formed by a path on four vertices. A graph is P_4-free if and only if it contains no induced subgraph isomorphic to P_4. It is also shown in [4] that a trivially perfect graph is characterized by being a P_4-free chordal

*Research carried out while this author was working at CWI and Utrecht University, Netherlands.

†Research supported by the Australian Research Council.

graph, since its stability number (i.e. cardinality of the largest independent set) equals the number of maximal cliques. In [3] the P_4-free chordal graphs were considered in relation to graphical Markov models.

In the next section we give the theorems we need on the structure of P_4-free chordal graphs. In Section 3 we find a generating function equation and recursive formulae for the numbers of labelled connected graphs, counted by vertices or by vertices and edges. We also obtain an asymptotic formula for the numbers counted by vertices, and some properties relating to the number of vertices of degree $n-1$ in the graphs on n vertices. These vertices play a central role in the structural results. The recurrences are all easy to compute, so we give just a few small numbers in tables along the way. The last result in that section shows that the unlabelled P_4-free chordal graphs correspond to unlabelled rooted trees, so that the numbers are the same.

2 Structure of P_4-free chordal graphs

The first characterization of P_4-free chordal graphs was by Wolk [9], who was investigating necessary and sufficient conditions on a graph to admit a transitive orientation. In this paper it is shown that a graph is P_4-free and chordal if and only if it is the comparability graph of a tree poset. A tree poset is one in which x and y are comparable whenever $x < z$ and $y < z$ for some z, and the comparability graph joins any two points of the poset which are comparable. This characterization can be used to obtain the result we need, but we derive it from the beginning for completeness and since the argument is almost as short.

For any graph G, define $D(G)$ to be the set of vertices of degree $|V(G)|-1$.

Proposition 1 *Let G be a connected graph. Then G is P_4-free and chordal if and only if it is complete or $G - D(G)$ is a disconnected P_4-free chordal graph.*

Proof We proceed by induction on $n = |V(G)|$. For $n = 1$ it is immediate. The complete case is trivial, so assume that G is a connected P_4-free chordal non-complete graph. As was shown by Wolk [10], G must have at least one vertex of degree $|V(G)|-1$. (The argument goes like this: if not, let u be a vertex of maximum degree, and let v be a neighbour of u and with a neighbour w not adjacent to u. Then for any other neighbour x of u the fact that $wvux$ is not an induced P_4 or 4-cycle implies that $vx \in E(G)$. This implies that the degree of v exceeds the degree of u, a contradiction.) Since $G' = G - v$ is an induced subgraph of a P_4-free chordal graph, it too is P_4-free and chordal.

If G' is disconnected, we are done. If G' is connected, then by induction $G' - D(G')$ is a disconnected P_4-free chordal graph. But clearly $G - D(G) = G' - D(G')$. The proposition follows. ∎

3 Enumeration

We use Proposition 1. Exact counting is considered first. Then an asymptotic formula is given in Theorem 1, which also gives asymptotics of the expected size of $D(G)$ and number of components of $G - D(G)$.

Let a_n be the number of labelled connected P_4-free chordal graphs on n vertices, and let $f = f(x) = \sum_{n=1}^{\infty} a_n x^n/n! = x + \frac{1}{2}x^2 + \cdots$ be the corresponding exponential generating function. Then by standard arguments as in Harary and Palmer [5] or Wilf [8], the graph $G - D(G)$ is counted by $e^f - f$, since e^f counts all P_4-free chordal graphs including the empty graph and the connected ones. (In this case we need to leave the empty graph as a possibility for $G - D(G)$, in case G is complete.) Thus

$$f = (e^x - 1)(e^f - f) \tag{1}$$

where $e^x - 1$ takes account of the deleted vertices $D(G)$. By simple algebra, this can be rewritten as

$$f = (1 - e^{-x})e^f. \tag{2}$$

We can observe that that the same deletion operation recursively gives a correspondence between these graphs and rooted labelled trees in which no vertex has exactly one child — or, equivalently, rooted trees with no non-root vertex of degree 2 and with root vertex of degree at least 2 — and in which each vertex has been expanded to a clique. The root vertex expands to the clique $D(G)$, and the other expansions are defined recursively. If edges are added from each vertex in a clique to all vertices "above" it in the tree, we recover the graph G. Thus, we can count these graphs alternatively by counting the corresponding trees: from the correspondence with these trees we have $f(x) = T(e^x - 1)$ where $T(x)$ is the exponential generating function for such trees. These are strongly related to homeomorphically irreducible labelled trees. Using the standard technique for counting trees of various types, it is easy to establish that T satisfies the equation $T = x(e^T - T)$. The resulting equation is (1).

3.1 Exact numbers of labelled graphs

We can easily use the well known "$x\frac{d}{dx}\log$" trick (see [8, Section 1.6] for example) to get a recurrence relation for the coefficients of f from (2) say, in terms of the coefficients of $1/(1 - e^{-x})$, which can of course be pre-computed.

If A_n is the total number of (not necessarily connected) labelled P_4-free and chordal graphs on n vertices then the corresponding exponential generating function is given by $e^{f(x)}$ using the standard exponential relationship expoused in [5] or [8] (with $A_0 = 1$ by convention). As in [5, p.9] this gives the recurrence $A_n = a_n + \frac{1}{n}\left(\sum_{k=1}^{n-1} k\binom{n}{k} a_k A_{n-k}\right)$. One can alternatively argue, by directly counting the results of deleting $D(G)$, that

$$a_n = 1 + \sum_{k=1}^{n-2} \binom{n}{k} (A_{n-k} - a_{n-k})$$

which can be combined with the previous recurrence to compute a_n and A_n recursively and simultaneously. The initial values of the recurrences are $a_1 =$

$a_2 = A_1 = 1$ and $A_2 = 2$. There are also simple ways to compute the coefficients using recursive expansions with an algebraic manupulation package such as Maple. Some numbers resulting from these recursions and computations are given in Table 1.

Table 1: Numbers of labelled connected (a_n) and all (A_n) P_4-free chordal graphs with n vertices

a_n	A_n	n
1	1	1
1	2	2
4	8	3
23	49	4
181	402	5
1812	4144	6
22037	51515	7
315569	750348	8
5201602	12537204	9
97009833	236424087	10
2019669961	4967735896	11
46432870222	115102258660	12

We turn to computing the number $a_{n,q}$ of labelled connected P_4-free chordal graphs on n vertices and q edges. A recurrence may be obtained by summing over all different possibilities for the set $D(G)$. For each of these sets D with k vertices, one has for $G - D(G)$ all the disconnected P_4-free chordal graphs on $n-k$ vertices. If G has q edges, $D(G)$ needs $k(k-1)/2$, since it is a clique. Also there are edges from D to each of the $n-q$ vertices of $G \setminus D$. Therefore, the recurrence for $a_{n,q}$ is

$$a_{n,q} = \sum_{k=1}^{n-2} \binom{n}{k} \left(A_{n-k,q-k(k-1)/2-k(n-k)} - a_{n-k,q-k(k-1)/2-k(n-k)}\right) \qquad (3)$$

where $A_{n,q}$ is the total number of labelled P_4-free chordal graphs on n vertices and q edges, with $A_{0,0} = 1$ by convention. Again by the exponential relationship, $\sum_{n\geq 0, q\geq 0} A_{n,q} x^n y^q / n! = \exp(\sum_{n\geq 1, q\geq 0} a_{n,q} x^n y^q / n!)$. This leads to

$$A_{n,q} = a_{n,q} + \sum_{l=0}^{q} \left(\frac{1}{n} \left(\sum_{k=1}^{n-1} k \binom{n}{k} a_{k,l} A_{n-k,q-l} \right) \right). \qquad (4)$$

Together, (3) and (4) determine the numbers $a_{n,q}$ recursively, beginning with $a_{2,1} = 1$. Table 2 gives the resulting values of $a_{n,q}$ for small n.

3.2 Asymptotics for labelled graphs

Here we find asymptotic expressions for a_n and A_n. Building on this, it is routine to obtain asymptotic properties of a random labelled P_4-free chordal graph G on

Table 2: Numbers of labelled connected P_4-free chordal graphs with n vertices and q edges

n = 2	3	4	5	6	7	8	q
1	0	0	0	0	0	0	1
	3	0	0	0	0	0	2
	1	4	0	0	0	0	3
		12	5	0	0	0	4
		6	30	6	0	0	5
		1	75	60	7	0	6
			30	270	105	8	7
			30	360	735	168	8
			10	435	1 925	1 680	9
			1	270	2 940	7 280	10
				255	3 591	16 800	11
				80	4 165	25 536	12
				60	2 310	38 108	13
				15	2 520	42 420	14
				1	1 925	35 700	15
					882	39 060	16
					630	28 728	17
					175	28 784	18
					105	20 860	19
					21	11 340	20
					1	9 240	21
						5 726	22
						2 268	23
						1 330	24
						336	25
						168	26
						28	27
						1	28

n vertices by way of the methods comonly used for tree enumeration problems. As examples of this, we include some properties relating to $D(G)$: its expected size (i.e. number of vertices of degree $n-1$) and also the number of components of $G - D(G)$ (i.e. number of branches at the root vertex of the corresponding rooted tree).

Theorem 1 *(a) As $n \to \infty$*

$$a_n \sim \sqrt{r(e-1)}n^{-1}\left(\frac{n}{er}\right)^n$$

where $r = 1 - \ln(e-1) \approx .4587$, and $A_n \sim ea_n$.

(b) Let G be a random labelled connected P_4-free chordal graph with n vertices and let D be the set of vertices of degree $n-1$ in G. Then as $n \to \infty$ the expected cardinality of D tends towards $e(1 - \ln(e-1)) \approx 1.2468$ and the expected number of components in $G - D$ tends towards $\frac{2e-1}{e-1} \approx 2.5820$.

Proof For asymptotics, we can use (1) which determines f implicitly as a function of x. We can more or less apply the theorem stated by Bender [1, Theorem 5], though care has to be taken, as noticed by Canfield [2] due to the possible multiple definition of f. This problem is that the correct singularity has to be identified. This has been remedied for special cases such as in [2], and in fact Meir and Moon [6, Theorem 1] give a result which applies immediately to (1) (see also [7]). Writing (2) as $f = F(x, f)$, this theorem guarantees that there is a singularity of $f(x)$ at the unique solution $z = r$, for positive real z, of the equations $w = F(z, w)$ and $1 = F_w(z, w)$, that is,

$$w = (1 - e^{-z})e^w = 1, \tag{5}$$

and that there are no other singularities of $f(z)$ for complex z with $|z| \le r$. Thus Bender [1, Theorem 5] is valid, and gives the asymptotics, as follows.

Solving (5), we observe $w = 1$, and the solution for z is then

$$r = 1 - \ln(e-1). \tag{6}$$

For [1, Theorem 5] we need to check

$$F_z(r, 1) = e - 1, \qquad F_{ww}(r, 1) = (1 - e^{-r})e = 1$$

and the result, after multiplying the coefficient of x^n in $f(x)$ by $n!$ and applying Stirling's formula, is the expression for a_n in (a). We postpone A_n until considering (b).

For the first part of (b) we have to find $\hat{a}_n/a_n$, where $\hat{a}_n$ is the number of P_4-free chordal graphs weighted according to the number of vertices in the set D. This can be deduced from [1, Theorem 2], but we prefer to give a standard singularity argument which in some gives the same approach to both parts of (b). Since D has exponential generating function $e^x - 1$, for such a weighted D we use $x\frac{d}{dx}(e^x - 1) = xe^x$. So letting $\hat{f} = \sum_{n=1}^{\infty} \hat{a}_n x^n/n!$, we have

$$\hat{f} = xe^x(e^f - f). \tag{7}$$

Together with (1), this gives

$$\hat{f} = \frac{xe^x f}{e^x - 1}. \tag{8}$$

From (7) it is clear that all singularities of $\hat{f}$ are also singularities of f. The coefficients of $\hat{f}$ are by definition greater than the corresponding coefficients of f, so the radius of convergence of $\hat{f}$ is at most that of f, i.e. r as given by (6). Thus (by Pringsheim's theorem), $\hat{f}$ has a unique singularity on its radius of convergence, at r. From the proof of [1, Theorem 5], we know that

$$f(z) = h(z) + c(z-r)^{1/2} + O((z-r)^{3/2}) \tag{9}$$

as $z \to r$ for a function h analytic at r. From (8), a similar statement is true of $\hat{f}$, with $h(z)$ replaced by $\hat{h}(z) = \frac{ze^z h(z)}{e^z-1}$. Hence by Darboux's theorem (see [1, Theorem 4]) $\hat{a}_n/a_n \sim \frac{re^r}{e^r-1}$, which is re by (6).

For the second part of (b), we require $\bar{a}_n/a_n$, where $\bar{a}_n$ is the number of P_4-free chordal graphs weighted according to the number of components when D is removed. These components are counted in (1) by $e^f - f$, so to give the required weighting we replace this factor by $f\frac{d}{df}(e^f - f) = (e^f - 1)f$. Thus, with $\bar{f} = \sum_{n=1}^{\infty} \bar{a}_n x^n/n!$,

$$\bar{f} = (e^x - 1)(e^f - 1)f = e^x f^2 - (e^x - 1)f \tag{10}$$

after a little manipulation using (1). From the form of this equation, $\bar{f}$ can have no singularity other than a singularity of f, and so no positive real singularity other than r in (6). The solution to (5) found in the proof of (a) has $w = 1$, and so $f(r) = 1$. Near r, the function f behaves as given in (9), and so we deduce $h(r) = 1$ and $f^2 = h^2(z) + 2ch(r)(z-r)^{1/2} + O((z-r)^{3/2})$. Thus from (10),

$$\bar{f}(z) = e^z h^2(z) - (e^z - 1)h(z) + (e^r + 1)c(z-r)^{1/2} + O((z-r)^{3/2})$$

as $z \to r$, and so, by Darboux's theorem and (9), $\bar{a}_n = 2e^r a_n - (e^r - 1)a_n + o(a_n)$. The rest of part (b) follows, since $e^r + 1 = \frac{2e-1}{e-1}$ by (6).

Finally it is easy to apply the same method as in (b) to obtain $A_n \sim ea_n$. One can note as above that the exponential generating function for A_n is $e^{f(x)}$ and then use $e^{f(z)} = e^{h(z)}(1 + c(z-r)^{1/2} + c^2(z-r)/2 + O((z-r)^{3/2})$ near $z = r$. ∎

3.3 Unlabelled enumeration

The characterization of P_4-free chordal graphs given in [10] suffices to show that these unlabelled connected graphs correspond to unlabelled rooted trees. We can also see this easily from Proposition 1: a connected P_4-free chordal graph G corresponds to a rooted tree T in which the length of the path P from the root vertex to the nearest vertex of degree at least 3 is $|D(G)|-1$, and the components of $T - P$ are the rooted trees corresponding to $G - D(G)$ (recursively defined). The exact and asymptotic numbers of unlabelled rooted trees with n vertices are given in [5].

Acknowledgement We wish to thank the anonymous referee who made a number of corrections and suggestions for improvement of this paper.

References

[1] E. A. Bender, Asymptotic methods in enumeration, *SIAM Rev.*, **16** (1974), 485–515.

[2] E. R. Canfield, Remarks on an asymptotic method in combinatorics, *J. Combin. Theory, Ser. A*, **37** (1984), 448–352.

[3] R. Castelo and A. Siebes, A characterization of moral transitive acyclic directed graph Markov models as labeled trees, *Journal of Statistical Planning and Inference*, *in press* (2002).

[4] M. C. Golumbic, Trivially perfect graphs, *Discrete Mathematics*, **24** (1978), 105–107.

[5] F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, New York, 1973.

[6] A. Meir and J. W. Moon, On an asymptotic method in enumeration, *J. Combin. Theory, Ser. A*, **51** (1989), 77–89.

[7] A. Meir and J. W. Moon, Erratum: “On an asymptotic method in enumeration”, *J. Combin. Theory, Ser. A*, **52** (1989), 163.

[8] H. S. Wilf, *generatingfunctionology*, Academic Press, 1990.

[9] E. S. Wolk, The comparability graph of a tree, *Proc. Am. Math. Soc.*, **13** (1962), 789–795.

[10] E. S. Wolk, A note on “The comparability graph of a tree”, *Proc. Am. Math. Soc.*, **16** (1965), 17–20.

[11] N. C. Wormald, Counting labelled chordal graphs, *Graphs and Combinatorics* **1** (1985), 193–200.