Some general results on enumeration of graphs (a survey)

Martin Klazar
(Charles University, MFF, Praha)

The 44th Czech and Slovak conference GRAPHS 2009
Terchová, Slovakia, May 25–29, 2009

Overview

1. Introduction
2. Hereditary and other classes
3. Regular graphs
4. Ultimate modular periodicity

Survey article: M. K., Overview of some general results in combinatorial enumeration, arXiv:0803.4292. To appear in Proceedings of the Conference on Permutation Patterns, St. Andrews, 2007.

a problem in enumeration of graphs:

$$P = (S_1, S_2, \dots),$$

S_n is a set of (simple, labeled) graphs on $[n] = \{1, 2, \dots, n\}$;
the counting function:

$$f_P(n) = |S_n| = \#G \in S_n.$$

The sets S_n are usually given as sections of an infinite universe of graphs U, by means of size functions and relations on U.

Example. Let U be all simple graphs with finite vertex sets $V \subset \{1, 2, \dots\} = \mathbb{N}$ and $(U, \prec)$ be the induced subgraph relation. We set

$$S_n = \{G \in U \mid V(G) = [n],\ K_{1,2} \not\prec G,\ K_3 \not\prec G\}.$$

Thus $G \in S_n$ iff G is a *partial matching*, a collection of isolated vertices and disjoint edges. How many graphs in S_n?

We have

$$f_P(n) = |S_n| = \sum_{k\geq 0} \binom{n}{2k} \cdot (2k-1)!!$$

where $(2k-1)!! = 1\cdot 3\cdot 5\cdots(2k-1)$, because the numbers m_k of *perfect matchings* on $[2k]$ satisfy recurrence $m_k = (2k-1)m_{k-1}$, $m_1 = 1$.

• The set

$$P = \{G \in U \; | K_{1,2} \not< G,\; K_3 \not< G\}$$

is a *hereditary class*, class of graphs closed to isomorphism and induced subgraphs; $f_P(n) = \#G \in P, V(G) = [n]$.

• The set S_n consists exactly of the graphs G on $[n]$ with $\Delta(G) \leq 1$, i.e., every vertex degree is 0 or 1. Also, we have the recurrence

$f_P(1) = 1$, $f_P(2) = 2$ and, for $n \geq 3$,

$$f_P(n) = f_P(n-1) + (n-1) \cdot f_P(n-2)$$

—again delete the first vertex. Thus

n	1	2	3	4	5	6	7	8	9	...
$f_P(n)$	1	2	4	10	26	76	232	764	2620	...

• Modular behavior: $f_P(n)$, $n = 1, 2, \ldots$, modulo
2 is $1, 0, 0, 0, 0, 0, 0, 0, 0, \ldots = 1, \overline{0}$
3 is $1, 2, 1, 1, 2, 1, 1, 2, 1, \ldots = 1, \overline{2, 1, 1}$
4 is $1, 2, 0, 2, 2, 0, 0, 0, 0, \ldots = 1, 2, 0, 2, 2, \overline{0}$
5 is $1, 2, 4, 0, 1, 1, 2, 4, 0, \ldots = \overline{1, 2, 4, 0, 1}$

Exercise. $f_P(n)$ modulo m is ultimately periodic for any m.

We extend this simple example into the three indicated directions: speed of hereditary (and other) graph classes, counting regular graphs, and ultimate modular periodicity of (the numbers of) combinatorial structures.

Scheme of a general result on enumeration of graphs. The counting function $f_P(n)$ of every problem P in a class $\mathcal{C}$ belongs to the class of functions $\mathcal{F}$:

$$\{f_P \mid P \in \mathcal{C}\} \subset \mathcal{F}.$$

The larger $\mathcal{C}$ is and the more specific the functions in $\mathcal{F}$ are, the stronger the result. *Exact results:* functions in $\mathcal{F}$ explicitly given (polynomials, recurrences, algorithms). *Asymptotic results:* functions in $\mathcal{F}$ given by asymptotic relations. Often mixture of both.

Hereditary and other classes

A class of graphs P is *hereditary* if it is closed to isomorphism and to induced subgraphs. The *speed* or *growth* of P is the function $f_P(n) = \#$ of graphs in P with the vertex set $[n]$.

Theorem. Let P be a hereditary class of graphs. Then

1. $f_P(n) = p_1(n) \cdot 1^n + \cdots + p_k(n) \cdot k^n$ for $n > n_0$ ($p_i \in \mathbb{Q}[x]$) or

2. $f_P(n) = n^{(1-1/k)n+o(n)}$ ($k \in \mathbb{Z}$ and $k \geq 2$) or

3. $n^{n+o(n)} < f_P(n) < 2^{o(n^2)}$ or

4. $f_P(n) = 2^{(1-1/k)n^2/2+o(n^2)}$ ($k \in \mathbb{Z}$ and $k \geq 2$) or

5. $f_P(n) = 2^{n(n-1)/2}$.

1–3 are due to Balogh, Bollobás and Weinreich (2000) and 4–5 to Alekseev (1992) and Bollobás and Thomason (1995). Partial matchings belong to group 2, with speed $n^{n/2+o(n)}$.

Oscillations in region 3 (BBW 2001): for any $c > 1$ and $\varepsilon > 1/c$ there is a hereditary (in fact *monotone*) class P with

$$f_P(n) = n^{cn+o(n)} \quad \text{and} \quad f_P(n) = 2^{(1+o(1))n^{2-\varepsilon}},$$

both for infinitely many n. Lower boundary of region 3:

Theorem (BBW 2005). Let P be a hereditary class. Then

1. $f_P(n) < n^{(1-1/k)n+o(n)}$ $(k \in \mathbb{N})$ or

2. $f_P(n) \geq B_n$ for $n > n_0$, where B_n are the Bell numbers. This lower bound is best possible.

Bell numbers: B_n is the number of partitions of an n-element set,

$\sum_{n\geq 0} B_n x^n = 1 + \frac{x}{1-x} + \frac{x^2}{(1-x)(1-2x)} + \frac{x^3}{(1-x)(1-2x)(1-3x)} + \cdots$

$B_n = n^{n(1-\log\log n/\log n + O(1/\log n))}$
(compare to $n! = n^{n(1-1/\log n + o(1/\log n))}$).

Theorem (Norine, Seymour, Thomas and Wollan 2006). If P is a *minor-closed class* then either $f_P(n) < c^n n!$ or $f_P(n) = 2^{n(n-1)/2}$.

Note that the results on speeds of hereditary and other classes often have form of *jumps in growth* — certain regions of speed are jumped over, i.e., are not realized by any $f_P(n)$.

What about *unlabeled graphs*?

Theorem (Balogh, Bollobás, Saks and Sós 200?). If P is hereditary and $g_P(n)$ counts non-isomorphic graphs in P, then

1. $g_P(n)$ is for $n > n_0$ constantly $0, 1$ or 2, or

2. $g_P(n) = cn^k + O(n^{k-1})$ ($k \in \mathbb{N}$, $c \in \mathbb{Q}$, $c > 0$) or

3. $g_P(n) \geq p_n$ for $n > n_0$, where p_n are the partition numbers. This lower bound is best possible.

The partition numbers: p_n is the number of integer partitions $n = m_1 + m_2 + \cdots + m_k$, $m_1 \geq m_2 \geq \cdots \geq m_k \geq 1$. So

$\sum_{n\ge 0} p_n x^n = \prod_{k\ge 1}(1-x^k)^{-1}$ and $p_n \sim (c_1/n)\cdot \exp(c_2\sqrt{n}\,)$.

What about *ordered graphs*? For $G = (V, E)$, $V \subset \mathbb{N}$, the isomorphism and the (induced) subgraph relation are required to respect the linear order $(V, <)$ inherited from $\mathbb{N}$.

Theorem (Klazar 2000). Let P be a hereditary class of ordered graphs such that in $G \in P$ all components are cliques and the number of components is bounded. Then

$$f_P(n) = p_1(n)\cdot 1^n + \cdots + p_k(n)\cdot k^n,\ n > n_0\ (p_i \in \mathbb{Q}[x]).$$

For example, the number of such graphs with exactly k components and n vertices is

$$S(n,k) = \sum_{i=0}^{k}(-1)^i \frac{(k-i)^n}{i!(k-i)!} \quad \text{(the Stirling number)}$$

(note that $\sum_{k=1}^{n} S(n,k) = B_n$, the Bell number) and the number of all such graphs with $\leq k$ components, which is a hereditary class, is $S(n,1) + S(n,2) + \cdots + S(n,k)$.

If the number of components may be unbounded, we get many other counting functions. For example, for the hereditary class of ordered graphs, with all components cliques,

$$P = \{G \mid \{13, 24\} \not\prec_o G\}$$

(i.e., no two edges in G cross), we get

$$f_P(n) = \frac{1}{n+1}\binom{2n}{n} \text{ (the Catalan number).}$$

Same if components are not restricted to cliques. But for speeds $< 2^n$ one has the following general result.

Theorem (Balogh, Bollobás and Morris 2006). If P is a hereditary class of ordered graphs, then

1. $f_P(n)$ is for $n > n_0$ constant or

2. $f_P(n) = a_0\binom{n}{0} + \cdots + a_k\binom{n}{k}$ for $n > n_0$ ($a_i \in \mathbb{Z}$, $a_k > 0$), and $f_P(n) \geq n$ for every n, or

3. $F_{n,k} \leq f_P(n) \leq n^c F_{n,k}$ ($k \in \mathbb{N}$, $k \geq 2$ and $c > 0$), where $F_{n,k}$ are the generalized Fibonacci numbers, or

4. $f_P(n) \geq 2^{n-1}$.

(This generalizes an analogous result of Kaiser and Klazar (2003) for hereditary classes of permutations.)

$F_{n,k}$, $k \geq 2$: $F_{n,k} = 0$ for $n < 0$, $F_{0,k} = 1$ and

$$F_{n,k} = F_{n-1,k} + F_{n-2,k} + \cdots + F_{n-k,k}, \; n > 0.$$

$F_{n,2}$ are the ordinary Fibonacci numbers; $F_{n,2} \approx 1.618^n$, $F_{n,3} \approx 1.839^n$, $F_{n,4} \approx 1.927^n, \ldots$

Corollary. For hereditary classes of ordered graphs one has the *poly-exp jump:* either $f_P(n) < n^c$ or $f_P(n) \geq F_n \approx 1.618^n$.

Problems. 1. Go above 2^{n-1}. (Accomplished by Vatter for permutations.) 2. Turn case 3 in an exact result, that is, determine the form of functions $f_P(n)$.

Speeds of hereditary classes were investigated for many other structures: permutations, tournaments, posets, words, hypergraphs, relational structures.

Counting regular graphs

G is k-regular: every vertex has degree k, i.e., is incident with k edges. We will be interested in exact numbers of k-regular graphs with the vertex set $[n] = \{1, 2, \ldots, n\}$. (Asymptotics, be it obtained classically or by the theory of random graphs, is another story.)

Already 1-regular graphs are of some interest. From the example on partial matchings we know that their number is

$$m_1(2n) = (2n-1) \cdot m_1(2n-2) = (2n-1)!!, \; m_1(2n-1) = 0.$$

Let $m_k(n) = \#$ k-regular G on $[n]$ and, for $I \subset \mathbb{N}_0 = \{0, 1, 2, \ldots\}$, let $m_I(n) = \#$ G on $[n]$ with every vertex degree in I; so $m_{\{k\}}(n) = m_k(n)$. All graphs considered are simple, loopless and labelled.

A sequence $f(n), n = 1, 2, \ldots$, of numbers is *P-recursive* (or *holonomic*) if, for some $2k$ polynomials $a_i(x)$ and $b_i(x)$, for every $n > n_0$,

$$f(n+k) = \frac{a_0(n)}{b_0(n)} f(n) + \frac{a_1(n)}{b_1(n)} f(n+1) + \cdots + \frac{a_{k-1}(n)}{b_{k-1}(n)} f(n+k-1).$$

Examples. 1. Fibonacci numbers: $F_{n+2} = F_{n+1} + F_n$.
2. Factorials: $n! = n \cdot (n-1)!$.
3. Catalan numbers: $c_n = \frac{1}{n+1}\binom{2n}{n}$ satisfies $c_{n+1} = \frac{4n-2}{n+1} \cdot c_n$.
4. Numbers of partial matchings:

$$m_{\{0,1\}}(n) = m_{\{0,1\}}(n-1) + (n-1) \cdot m_{\{0,1\}}(n-2).$$

The last example is a special case of the following remarkable general result.

Theorem (Gessel, 1990). For any finite set $I \subset \mathbb{N}_0$, the sequence $m_I(n)$, $n = 1, 2, \dots$, of numbers of graphs with vertex set $[n]$ and all vertex degrees in I is P-recursive.

Examples. 1. A non-example: for infinite $I = \mathbb{N}_0$, numbers $m_I(n)$ are not P-recursive, because then $m_I(n) = 2^{n(n-1)/2}$ (all graphs) and these numbers grow too fast. ($f(n)$ P-recursive $\Rightarrow |f(n)| < n^{cn}$.)

2. Let us count 2-regular graphs, $I = \{2\}$. The # of connected 2-reg. graphs on $[n]$ is $(n-1)!/2$, $n \geq 3$. Thus, by the exponential formula for EGF,

$$F(x) = \sum_{n \geq 0} \frac{m_2(n)x^n}{n!} = \exp(G(x)),\ G(x) = \sum_{n \geq 3} \frac{(n-1)! \cdot x^n}{2 \cdot n!} =$$

$= \sum_{n \geq 3} x^n/2n$. So $G' = (\log F)' = F'/F$ and $G' \cdot F = F'$, that is,

$$\sum_{n\geq 2}\frac{x^n}{2}\cdot\sum_{n\geq 0}\frac{m_2(n)x^n}{n!}=\sum_{n\geq 0}\frac{m_2(n+1)x^n}{n!}$$

$$x^2\sum_{n\geq 0}\frac{m_2(n)x^n}{n!}=2(1-x)\sum_{n\geq 0}\frac{m_2(n+1)x^n}{n!}$$

$$m_2(n-2)/(n-2)!=2m_2(n+1)/n!-2m_2(n)/(n-1)!$$

therefore $m_2(1)=m_2(2)=0, m_2(3)=1$ and

$$m_2(n+3)=\frac{(n+2)(n+1)}{2}\cdot m_2(n)+(n+2)\cdot m_2(n+2),\ n\geq 1.$$

So $m_2(4)=3$, $m_2(5)=12$, $m_2(6)=70$, $m_2(7)=465,\ldots$. Alternatively, considering the component containing 1, we get directly for $m_2(n)$ the recurrent expression

$$m_2(n) = \sum_{k=2}^{n-1} \binom{n-1}{k} \frac{k!}{2} \cdot m_2(n-k-1)$$

which can be inductively verified to satisfy the P-recurrence of order 3.

Gessel proved his theorem by means of symmetric functions and the generating function in infinitely many variables

$$F(x_1, x_2, \dots) = \prod_{1 \le i < j} (1 + x_i x_j) = \sum_{d_1, \dots, d_n \ge 0} m(d_1, \dots, d_n) x_1^{d_1} \dots x_n^{d_n}$$

where the coefficient $m(d_1, \dots, d_n)$ is the # of graphs G on $[n]$ such that $\deg_G(i) = d_i$ for $i = 1, 2, \dots, n$.

Ultimate modular periodicity

Fibonacci numbers $(F_n)_{n\geq 1} = (1,2,3,5,8,13,\dots)$, where $F_{n+2} = F_{n+1} + F_n$, are ultimately periodic modulo m for any m. By the pigeonhole, $(F_k, F_{k+1}) \equiv (F_l, F_{l+1})$ mod m for some $k < l$, which implies that for $0 \leq i < l - k$,

$$F_{k+i} \equiv F_{k+(l-k)+i} \equiv F_{k+2(l-k)+i} \equiv \dots \mod m.$$

Running the recurrence backwards, we see that F_n mod m are in fact fully periodic, with period $l - k$.

Recall that the numbers $f(n)$ of partial matchings on $[n]$ satisfy the P-recurrence $f(n+2) = (n+1)f(n) + f(n+1)$. Again, $f(n)$ mod m are ultimately periodic for any m: $(k+1, f(k), f(k+1)) \equiv (l+1, f(l), f(l+1))$ mod m for some $k < l$ and so on. However, there are preperiods and $f(n)$ mod m is in general not fully periodic, as its P-recurrence cannot be run backwards.

And this is the reason why P-recursive sequences in general are not ultimately periodic mod m. Let us look at the Catalan numbers $c_n = \frac{1}{n+1}\binom{2n}{n}$, satisfying the P-recurrence

$$c_{n+1} = \frac{4n-2}{n+1} \cdot c_n.$$

Modulo m we divide by 0 infinitely often. In fact, it is not hard to show that $c_n \equiv 1$ mod 2 iff $n = 2^m - 1$, which is not ultimately periodic.

What about the Bell numbers?

$\sum_{n\geq 0} B_n x^n = 1 + \frac{x}{1-x} + \frac{x^2}{(1-x)(1-2x)} + \frac{x^3}{(1-x)(1-2x)(1-3x)} + \cdots$

$\equiv 1 + \frac{x}{1-x} + \frac{x^2}{1-x} + \frac{x^3}{(1-x)^2} + \frac{x^4}{(1-x)^2} + \frac{x^5}{(1-x)^3} + \cdots \mod 2$

$$= 1 + (1 + x^{-1}) \sum_{n\geq 1} \frac{x^{2n}}{(1-x)^n} = \frac{1}{1-x-x^2}.$$

So $(1-x-x^2)\sum_{n\geq 0} B_n x^n \equiv 1 \mod 2$ and

$$B_{n+2} \equiv B_{n+1} + B_n \mod 2.$$

Thus B_n are periodic mod 2 and

$$B_n = 1, 2, 5, 15, 52, \ldots \mod 2 \ \equiv \overline{1, 0, 1}.$$

Similarly, B_n is periodic mod m for any m.

Theorem (Blatter and Specker, 1981). Let ϕ be a closed formula in the MSOL, using only unary and binary predicates (i.e., symbols for relations), and $f_\phi(n)$ be the number of models of ϕ on $[n]$ (i.e., relational systems on $[n]$ in which ϕ holds). Then the sequence

$$f_\phi(n) \bmod m,\ n = 1, 2, \ldots,$$

is ultimately periodic for any m.

As an example, with variables a, b, c and one binary predicate $\sim$, we may take this ϕ:

$$\forall a, b, c:\ (a \sim a)\ \&\ (a \sim b \Rightarrow b \sim a)\ \&\ ((a \sim b\ \&\ b \sim c) \Rightarrow a \sim c).$$

The models of ϕ on $[n]$ are exactly the equivalence relations, thus $f_\phi(n) = B_n$. This ϕ is in fact even FOL formula.

MSOL = monadic second order logic: the language of FOL (predicates, no functions) + variables S for sets of elements, which can be quantified by $\forall, \exists$; atomic formulas of the type $x \in S$.

Examples. Gives ultimate modular periodicity for sequences of numbers of many classes of labelled graphs on $[n]$, for example

• triangle-free graphs (FOL definable)
• graphs avoiding finitely many forbidden induced subgraphs (FOL definable)
• k-regular graphs (FOL definable)
• k-colorable graphs (MSOL definable)
• planar graphs (MSOL definable, via Kuratowski's theorem)

Fischer (2003) showed that the theorem does not hold for quaternary relations.

Problem. Does it hold for ternary relations?

Thank you for your attention!