Advanced Topics in Computer Science 236605
(Winter 2009/10)
Logical Methods in Combinatorics

Lecturer: Prof. János Makowsky

e-mail: janos@cs.technion.ac.il,

Homepage: `http://www.cs.technion.ac.il/ janos/COURSES/236605-09`

Office: Taub 628

Reception hours: Monday, 14-16, or by appointment via e-mail.

Outline of Lecture 1

Density, speed and probalities

- Organisational matters
- Purpose of the course
- Counting finite structures: Density, speed and probabilities
- 0-1 Law for First Order Logic (outline)

Course prerequisites and requirements

Lectures:

Weekly two hour lectures.
Topic of the week.

Homework:

Weekly homework assignments.
Complementing material of the course.
No hand-in required.
Connect passive and active knowledge.
Measure your understanding.
Control it yourself or with a partner.

Purpose of the course

We want to explore theorems in

combinatorial and algorithmic graph theory

the proofs of which use methods from **logic**.

- 0-1 laws for graph properties.
- Linear recurrence relations for counting graphs on n vertices.
- Linear recurrences for graph polynomials.
- Parametrized complexity of graph polynomials.

Counting finite topologies

Let T_n be the number of topologies on the set $\{1, \ldots, n\}$.

$T_1 = 1$, as the underlying set is always open.

$T_2 = 4$, for each singleton, we can decide whether it is open or not.

T_n is bounded by 2^{2^n}, hence $T_5 \leq 2^{32}$.

Two papers

$T_5 = 7181$

A. Shaafat, *On the number of topologies definable for a finite*, J. Australian Mat. Soc., vol 8 (1968), 194-198.

$T_5 = 6942$

J. Evans, F. Harary and M.S. Lynn, *On the computer enumeration of finite toplogies*, Communications of the ACM, 10 (1967), 295-297.

In the course we shall prove that $T_5 = 2 \pmod 5$.

$$7181 \neq 2 \mod 5$$

$$6942 = 2 \mod 5$$

This will allow us to conclude that $T_5 = 7181$ is not possible.

Logic and Combinatorics

- The class of finite topologies is not definable in First Order or even Second Order Logic.
- But the number of topologies on n points is the same as the number of reflexive transitive relations on n points.
- The class of reflexive transitive relations on n points is First order definable.
- Counting First Order definable relations is amenable using techniques from logic.

Toplogies and preorders

A *preorder* R is a
transitive and reflexive binary relation.

If additionally $R(a, b) \wedge R(b, a) \Rightarrow a = b$, R is a *partial order*.

$PreO_n$ is the number of preorders on the set $\{1, \ldots, n\}$.

Theorem:
The number of preorder on a finite set equals the number of topolgies, i.e., $PreO_n = T_n$.

Proof: Homework 1

Graph properties

A simple graph

$$G =< V, E >=< V(G), E(G) >$$

has a set V of vertices and $E \subseteq V^2$ of edges. Simple graphs have no multiple edges. G is finite if V is finite.

A **graph property** $\mathcal{P}$ is a class of (finite) graphs
closed under graph isomorphisms.

A graph property $\mathcal{P}$ is **monotone**, if it is closed under subgraphs, and **hereditary**, if it is closed under induced subgraphs.

A graph property $\mathcal{P}$ is $FOL(\tau)$-**definable** if it consists of all (finite) graphs satisfying an $FOL(\tau)$-formula ϕ. (Similarly for $SOL, MSOL$)

Examples and Exercise:

Check for monotonicity, hereditarity and definability:

Planar graphs, regular graphs, Eulerian graphs, Hamiltonian graphs,

connected graphs, 3-colorable graphs, bipartite graphs

Density and asymptotic probabilities

Counting graphs

We want to count the number of finite graphs having a given graph property.

There are various ways of doing this.

- Density functions of a graph property
- Asymptotic probabilities of a graph property
- Given a graph, count the number of induced subgraphs with a given property
- Turn any of the above into a generating function.

Counting graphs: The density function.

- Graphs with n vertices will have $V = \{0, 1, \ldots, n-1\} = [n]$.
- There are $gr(n) = 2^{\binom{n}{2}} = 2^{\frac{n(n-1)}{2}}$ many graphs with n vertices.
- For a property $\mathcal{P}$ denote by $\mathcal{P}^n$ the graphs with n vertices in $\mathcal{P}$, and by $f_{\mathcal{P}}(n) = \mid \mathcal{P}^n \mid$, the number of graphs G with $V(G) = [n]$ which are in $\mathcal{P}$. $f_{\mathcal{P}}(n)$ is called the **density function** of $\mathcal{P}$.

  If $\mathcal{P}$ is hereditary, the density function of $\mathcal{P}$ is also called the **speed** $\mathcal{P}$, since it is an **ultimately monotone increasing function**.
- The **spectrum** $sp(\mathcal{P})$ of $\mathcal{P}$ is the set $\{n \in \mathbb{N} : f_{\mathcal{P}}(n) \neq 0\}$.

What can we say about $f_{\mathcal{P}}(n)$ and $sp(\mathcal{P})$?

Density functions and generating functions

- The power series

$$F_{\mathcal{P}}(n) = \sum_{n}^{\infty} f_{\mathcal{P}}(n) \cdot X^n$$

is called the **generating function** of $\mathcal{P}$.

- The power series

$$\mathbb{F}_{\mathcal{P}}(n) = \sum_{n}^{\infty} f_{\mathcal{P}}(n) \cdot \frac{X^n}{n!}$$

is called the **exponential generating function** of $\mathcal{P}$.

What can we say about $F_{\mathcal{P}}(n)$, and $\mathbb{F}_{\mathcal{P}}(n)$?

Density functions: References

There is rich literature on density functions.

[Rio68] J. Riordan,
Combinatorial identities, Robert E. Krieger, New York, 1968

[HP73] F. Harary and E. Palmer,
Graphical enumeration, Academic Press, 1973.

[Wil90] H.S. Wilf,
generatingfunctionology, Academic Press, 1990 (2nd ed. 1994).
Also: `www.math.upenn.edu/~wilf/DownldGF.html`

[BBW02] J. Balogh, B. Bollobás and David Weinreich,
Measures on monotone properties of graphs,
Discrete Applied Mathematics 116 (2002), 17-36.

We shall see examples in a moment.

Counting graphs: Asymptotic probability

With the density function we can also define probabilities:

- We define $\mu_n(\mathcal{P}) = \frac{|\mathcal{P}^n|}{gr(n)}$, the **fraction** of graphs in $\mathcal{P}$.

- $\mu(\mathcal{P}) = \lim_{n\to\infty} \mu_n(\mathcal{P})$ is called the **asymptotic probability** of $\mathcal{P}$.

- Let $G(n, p(n))$ be the **random graph** on n labeled vertices where each edge is chosen with probability $p(n)$, and $\mu_n^p(\mathcal{P})$ and $\mu^p(\mathcal{P})$ the probability, respectively asymptotic probability that $G(n, p(n))$ in in $\mathcal{P}$.

  For $p(n) = \frac{1}{2}$ constant we get $\mu_n(\mathcal{P}) = \mu_n^{\frac{1}{2}}(\mathcal{P})$

What can we say about $\mu_n(\mathcal{P})$ and $\mu(\mathcal{P})$?

Random graphs: References

There is rich literature on random graphs.

[Bol85] B. Bollobás,
Random Graphs, Academic Press, 1985

A new book on the same topic by the same author is in preparation.

[ASE92] N. Alon, J. Spencer and P. Erdös,
The Probabilistic Method, John Wiley, 1992

[Spe01] J. Spencer,
The Strange Logic of Random Graphs, Springer, 2001

Some motivating (and confusing) examples

- Trees
- Cliques, stars and disjoint unions of cliques
- Connected graphs
- Regular graphs
- Forbidden subgraphs (homework)

Trees

Let $\mathcal{P}$ be the class of labeled trees.

- $\mathcal{P}$ not hereditary hence not monotone.
  $\mathcal{P}$ is not FOL-definable, but $MSOL$-definable.

- $f_{\mathcal{P}}(n) = n^{n-2}$ (Caley's Theorem, 1889).

- $sp(\mathcal{P}) = \mathbb{N} - \{0\}$.

- $F_{\mathcal{P}}(n) = \sum_n n^{n-2} X^n$ and $\mathbb{F}_{\mathcal{P}}(n) = \sum_n \frac{n^{n-2} \cdot X^n}{n!}$.

- $\mu_n(\mathcal{P}) = \frac{n^{n-2}}{2^{\binom{n}{2}}}$ and $\mu(\mathcal{P}) = 0$

Exercise: Compute n^{n-2} modulo m. Show that for fixed m, this gives an ultimately periodic sequence.

Hint: Use Little Fermat: $n^{p-1} = 1 \pmod{p}$ if p does not divide n.

Counting graphs: Cliques

Let $\mathcal{P} = \{K_n : n \in \mathbb{N}\}$ be the cliques (complete graphs).

- $\mathcal{P}$ is hereditary but not monotone.
  $\mathcal{P}$ is FOL-definable.
- $f_{\mathcal{P}}(n) = 1$ and trivially periodic modulo any m.
- $sp(\mathcal{P}) = \mathbb{N} - \{0\}$.
- $F_{\mathcal{P}}(n) = \sum_n X^n$ and $\mathbb{F}_{\mathcal{P}}(n) = \sum_n \frac{X^n}{n!}$.
- $\mu_n(\mathcal{P}) = \frac{1}{2^{\binom{n}{2}}}$ and $\mu(\mathcal{P}) = 0$

Counting graphs: Stars

For $\mathcal{P} = \{K_{1,n} : n \in \mathbb{N}\}$, the stars
(complete bipartite graphs with 1 and n vertices),

- $\mathcal{P}$ is not hereditary.
  $\mathcal{P}$ is FOL-definable.
- $f_{\mathcal{P}}(n) = n$ and trivially periodic modulo any m.
- $sp(\mathcal{P}) = \mathbb{N} - \{0, 1\}$.
- $F_{\mathcal{P}}(n) = \sum_n n \cdot X^n$ and $\mathbb{F}_{\mathcal{P}}(n) = \sum_n n \cdot \frac{X^n}{n!}$.
- $\mu_n(\mathcal{P}) = \frac{n}{2^{\binom{n}{2}}}$ and $\mu(\mathcal{P}) = 0$

If $\mathcal{Q}$ is the closure of $\mathcal{P}$ under induced substructures, we get either stars or sets of isolated points, and $f_{\mathcal{Q}}(n) = n + 1$.

The density function of connected graphs

The class CONN of connected labeled graphs

- is not hereditary,
- is not FOL(R)-definable,
- but it is MSOL(R)-definable using a universal quantifier over set variables.
- It is also definable in Fixed Point Logic (FPL).

Counting labeled connected graphs is treated in [HP74] chapters 1 and 7 and in [Wil90] chapter 3. [HP74] chapter 1, page 7 gives:

$$f_{\mathsf{CONN}}(n) = 2^{\binom{n}{2}} - \frac{1}{n}\sum_{k=1}^{n-1} k\binom{n}{k} 2^{\binom{n-k}{2}} f_{\mathsf{CONN}}(k).$$

This does not look vey useful, but we get $\mu(\mathsf{CONN}) = 1$.

The exponential formula for connected graphs

Corollary 1.2.8 in [HP74] and Corollary 3.4.1 in [Wil90].

Let $\mathcal{C}$ be a class of graphs and $\mathcal{D}$ the set of connected graphs in $\mathcal{C}$.

We note that if $\mathcal{C}$ is $MSOL$-definable so is $\mathcal{D}$.

There is a remarkable theorem due to R.J. Riddell,
which relates the exponential generating functions $\mathbb{F}_{\mathcal{C}}$ and $\mathbb{F}_{\mathcal{D}}$.

With $\sum_n \frac{g(\bar{x})^n}{n!} = e^{g(\bar{x})}$ we have

$$\mathbb{F}_{\mathcal{C}} = e^{\mathbb{F}_{\mathcal{D}}}$$

For the class of all graphs and connected graphs we get:

$$\mathbb{F}_{graphs} = \sum_n 2^{\binom{n}{2}} \cdot \frac{X^n}{n!} = e^{\mathbb{F}_{\text{CONN}}}$$

Counting graphs: Two disjoint cliques

Let $\mathcal{P}$ consist of graphs which are a union of two disjoint cliques.

- $\mathcal{P}$ is not hereditary hence not monotone.
  $\mathcal{P}$ is FOL-definable. **Homework:** Find the formula!

- $f_{\mathcal{P}}(n) = 2^{n-1}$ and and ultimately periodic modulo any m (using Little Fermat again).

- $sp(\mathcal{P}) = \mathbb{N} - \{0, 1\}$.

- $F_{\mathcal{P}}(n) = \sum_n 2^{n-1} \cdot X^n$ and $\mathbb{F}_{\mathcal{P}}(n) = \sum_n 2^{n-1} \cdot \frac{X^n}{n!}$.

- $\mu_n(\mathcal{P}) = \frac{2^{n-1}}{2^{\binom{n}{2}}}$ and $\mu(\mathcal{P}) = 0$

For arbitrary disjoint unions of cliques (DUCliques) we use Riddell's formula and get for the exponential generating function

$$\mathbb{F}_{DUCliques} = e^{\mathbb{F}_{Cliques}} = e^{e^X}$$

Counting graphs: 1-regular graphs

Let $\mathcal{P}$ consist of graphs which are 1-regular
(disjoint union of non-connected edges, perfect matchings).

- $\mathcal{P}$ is not hereditary.
- $\mathcal{P}$ is FOL-definable.
- **Exercise:** Compute $f_{\mathcal{P}}(2m)$ and $sp(\mathcal{P})$.
- **Exercise:** Compute $\mu_m(\mathcal{P})$ and $\mu(\mathcal{P})$.

2-regular graphs

2-regular graphs are disjoint unions of circles.

$f_{Circle} = \frac{n!}{n}$ hence the exponential generating function is

$$\mathbb{F}_{Circle} = \sum_{n} \frac{(\frac{n!}{n})^n \cdot X^n}{n!}$$

We have using Riddel's formula

$$\mathbb{F}_{\mathcal{R}EG_2} = e^{\mathbb{F}_{Circle}} = e^{\left(\sum_n \frac{(\frac{n!}{n})^n \cdot X^n}{n!}\right)}$$

For 2-regular graphs [Wil90] determines the exponential generating function:

$$\mathbb{F}_{\mathcal{R}EG_2} = \sum_{n=0} f_{\mathcal{R}EG_2}(n) \frac{x_n}{n!} = \frac{e^{-\frac{1}{2}x - -\frac{1}{4}x^2}}{\sqrt{1-x}}$$

The density function for regular graphs

The class $\mathrm{REG_r}$ of simple regular graphs where every vertex has degree r is FOL-definable (for fixed r).

The formulas says that every vertex has exactly r different neighbors. The formula grows with r. Regularity without specifying the degree is not FOL-definable, actually not even $CMSOL$-definable.

Counting the number of labeled regular graphs is treated completely in Chapter 7 of [HP74], where an explicit formula is given, essentially due to J.H. Redfield (1927) and rediscovered by R.C. Read (1959).

However, the formula is very complicated.

For cubic graphs, the function is explicitly given: $f_{\mathcal{R}_3}(2n+1) = 0$ and

$$f_{\mathcal{R}_3}(2n) = \frac{(2n)!}{6^n} \sum_{j,k} \frac{(-1)^j (6k-2j)! 6^j}{(3k-j)!(2k-j)!(n-k)!} 48^k \sum_i \frac{(-1)^i j!}{(j-2i)! i!}$$

Where does logic enter?

0-1 Laws for asymptotic probabilities

Theorem A: (Kolaitis and Vardi, 1992)
(generalizing a long sequence of earlier papers since 1964)

For $\mathcal{P}$ definable in infinitary logic with finitely many variables, $\mathcal{L}^{\omega}_{\infty,\omega}$, either $\mu_{\mathcal{P}} = 0$ or $\mu_{\mathcal{P}} = 1$.

This works also for any constant probability p and $\mu^{p}_{\mathcal{P}}$.

Theorem B: (Shelah and Spencer, 1988)

For $\alpha \in [0, 1]$ irrational, $\mathcal{P}$ definable in FOL, either $\mu^{n^{-\alpha}}_{\mathcal{P}} = 0$ or $\mu^{n^{-\alpha}}_{\mathcal{P}} = 1$.

For all rational α there are counterexamples.

Where does logic enter?

Density functions modula m

Theorem C: (Blatter and Specker, 1981)

For $\mathcal{P}$ definable in $MSOL(\tau)$,
where τ is relational and has relation symbols of arity at most 2,
and for every $m \in \mathbb{N}$,
the function $f_{\mathcal{P}}(n) \pmod{m}$ on $\mathbb{Z}_m$
satisfies a linear recurrence relations, i.e
there are $q, n_0, a_1(m), \ldots, a_q(m) \in \mathbb{N}$ such that for $n \geq n_0$ we have

$$f_{\mathcal{P}}(n+q) = \sum_{k=0}^{q-1} a_k(m) \cdot f_{\mathcal{P}}(n+k) \pmod{m}$$

and hence is ultimately periodic.

For relations of arity ≥ 4, E. Fischer (2002) found counterexamples.

When does logic enter? (References)

1964 H. Gaifman,
Concerning measures of first order calculi,
Israel Journal of Mathematics, 2 (1964), pp. 1-17.

1969 Y.V. Glebskii, D.I. Kogan, M.I. Liagonkii and V.A. Talanov,
Range and degree of realizability of formulas in the restricted predicate calculus,
Cybernetics 5 pp. 142-154 (Russian original: Kybernetika 5 (1969), pp. 17-27.

1976 R. Fagin,
Probabilities on finite models, Journal of Symbolic Logic, 41 (1976), pp. 50-58.

1981 C. Blatter and E. Specker,
Le nombre de structures finies d'une th'eorie à charactère fin, Sciences Mathématiques, Fonds Nationale de la recherche Scientifique, Bruxelles, 1981, pp. 41-44.

1984 C. Blatter and E. Specker,
Recurrence relations for the number of labeled structures on a finite set, In *Logic and Machines: Decision Problems and Complexity*, E. Börger and G. Hasenjaeger and D. Rödding, Springer Lecture Notes in Computer Science, 171 (1984), pp. 43-61.

1988 S. Shelah and J. Spencer,
Zero-One Laws for Sparse Random Graphs,
Journal of the AMS 1 (1988), pp. 97-115.