Connection Matrices
and
Definability of Graph Invariants

Johann A. Makowsky

Faculty of Computer Science,
Technion - Israel Institute of Technology,
Haifa, Israel

http://www.cs.technion.ac.il/~janos
e-mail: janos@cs.technion.ac.il

****

Reporting also recent work by

M. Freedman, L. Lovász, A. Schrijver and B. Szegedy

Based on joint work with B. Godlin and T. Kotek

File:new-title

Three Lectures on Connection Matrices

- Lecture 1: Definability of graph properties and graph parameters

- Lecture 2: Characterizing partition functions

- Lecture 3: Definability of graph polynomials

Lecture 1: Overview

- Tame Logics

- Numeric graph invariants: Properties and guiding examples

- Non-definability via Complexity Theory

- Typical properties of graph parameters

- Connection matrices

- More connection matrices Parametrized numeric graph invariants and graph polynomials

- MSOL-definable graph polynomials

- Definability and non-definability in MSOL of graph parameters

Logic

Logics

In this talk a logic $\mathcal{L}$ is a fragment of Second Order Logic **SOL**.

Let $\mathcal{L}$ be a subset of **SOL**. $\mathcal{L}$ is a fragment of **SOL** if the following conditions hold.

(i) For every finite relational vocabulary τ the set of $\mathcal{L}(\tau)$ formulas contains all the atomic τ-formulas and is closed under boolean operations and renaming of relation and constant symbols.

(ii) $\mathcal{L}$ is equipped with a notion of quantifier rank and we denote by $\mathcal{L}_q(\tau)$ the set of formulas of quantifier rank at most q. The quantifier rank is subadditive under substitution of subformulas,

(iii) The set of formulas of $\mathcal{L}_q(\tau)$ with a fixed set of free variables is, up to logical equivalence, finite.

(iv) Furthermore, if $\phi(x)$ is a formula of $\mathcal{L}_q(\tau)$ with x a free variable of $\mathcal{L}$, then there is a formula ψ logically equivalent to $\exists x \phi(x)$ in $\mathcal{L}_{q'}(\tau)$ with $q' \geq q + 1$.

(v) A fragment of **SOL** is called **tame** if it is closed under scalar transductions.

Typical fragments

- First Order Logic **FOL**.

- Monadic Second Order Logic **MSOL**.

- Logics augmented by modular counting quantifiers: $D_{m,i}x\phi(x)$ which says that the numbers of elements satisfying ϕ equals i modulo m.

- **CFOL**, **CMSOL** denote the logics **FOL**, resp. **MSOL**, augmented by all the modular counting quantifiers.

- Logics augmented by Lindström quantifiers.

- Logics restricted a fixed finite set of bound or free variables.

Boolean and Numeric graph invariants

aka

Graph properties and Graph parameters

Graph properties (boolean graph invariants)

We denote by $G = (V(G), E(G))$ a graph,
and by $\mathcal{G}$ and $\mathcal{G}_{simple}$ the class of finite (simple) graphs, respectively.

A **graph property** or **boolean graph invariant** is a function

$$f : \mathcal{G} \to \mathbb{Z}_2$$

which is invariant under graph isomorphism.

More traditionally, a graph property $P = P_f$ is a family of graphs closed under isomorphisms given by $P_f = \{G : f(G) = 1\}$.

  (i) P is hereditary, if it is closed under induced subgraphs.

 (ii) P is monotone, if it is closed under (not necessarily induced) subgraphs.

(iii) P is definable in some logic $\mathcal{L}$ if there is a formula $\phi \in \mathcal{L}$ such that $P = \{G : G \models \phi\}$.

(iv) Regular graphs of fixed degree d are definable in First order Logic **FOL**.

 (v) Connectivity and planarity are definable in Monadic Second Order Logic **MSOL**.

File:gpar

Numeric graph invariants (graph parameters)

We denote by $G = (V(G), E(G))$ a graph,
and by $\mathcal{G}$ and $\mathcal{G}_{simple}$ the class of finite (simple) graphs, respectively.

A **numeric graph invariant** or **graph parameter** is a function

$$f : \mathcal{G} \to \mathbb{R}$$

which is invariant under graph isomorphism.

(i) Cardinalities: $|V(G)|$, $|E(G)|$

(ii) Counting configurations:

$k(G)$ the number of connected components,
$m_k(G)$ the number of k-matchings

(iii) Size of configurations:

$\omega(G)$ the clique number
$\chi(G)$ the chromatic number

(iv) Evaluations of graph polynomials:

$\chi(G, \lambda)$, the chromatic polynomial, at $\lambda = r$ for any $r \in \mathbb{R}$.
$T(G, X, Y)$, the Tutte polynomial, at $X = x$ and $Y = y$ with $(x, y) \in \mathbb{R}^2$.

Definability of numeric graph parameters, I

We first give examples where we use small, i.e., polynomial sized sums and products:

(i) The cardinality of V is **FOL**-definable by

$$\sum_{v \in V} 1$$

(ii) The number of connected components of a graph G, $k(G)$ is **MSOL**-definable by

$$\sum_{C \subseteq V : \text{component}(C)} 1$$

where component(C) says that C is a connected component.

(iii) The graph polynomial $X^{k(G)}$ is **MSOL**-definable by

$$\prod_{c \in V : \text{first}-\text{in}-\text{comp}(c)} X$$

if we have a linear order in the vertices and first $-$ in $-$ comp(c) says that c is a first element in a connected component.

Definability of numeric graph parameters, II

Now we give examples with possibly large, i.e., exponential sized sums:

(iv) The number of cliques in a graph is **MSOL**-definable by

$$\sum_{C \subseteq V : \text{clique}(C)} 1$$

where clique(C) says that C induces a complete graph.

(v) Similarly "the number of maximal cliques" is **MSOL**-definable by

$$\sum_{C \subseteq V : \text{maxclique}(C)} 1$$

where maxclique(C) says that C induces a maximal complete graph.

(vi) The clique number of G, $\omega(G)$ is is **SOL**-definable by

$$\sum_{C \subseteq V : \text{largest}-\text{clique}(C)} 1$$

where largest $-$ clique(C) says that C induces a maximal complete graph of largest size.

File:gpar

Definability of numeric graph parameters, III

A numeric graph parameter is $\mathcal{L}$-definable if it can be defined by similar expressions using large and small sums and only small products.

Usually, summation is allowed over second order variables, whereas products are over first order variables.

How can we prove definability and non-definability of graph parameters in some logic $\mathcal{L}$? In particular:

- How to prove that $k(G)$ is not **CFOL**-definable?

- How to prove that $\omega(G)$ is not **CMSOL**-definable?

- How to prove that the chromatic number $\chi(G)$ or the chromatic polynomial $\chi(G, X)$ is not **CMSOL**-definable?

File:gpar

Non-definability via complexity assumptions:

Harmonious colorings

A vertex coloring of a graph G with k colors is harmonious if it is proper and each pair of colors appears at most once along an edge.

The harmonious index of a graph G is the smallest k such that there is a harmonious coloring with k colors.

- J.E. Hopcroft and M.S. Krishnamoorthy studied harmonious colorings in 1983.

- B. Courcelle has shown that graph parameters (polynomials) definable in **CMSOL** can be
  computed in polynomial time for graphs of tree-width at most k.

- K. Edwards and C. McDiarmid showed that computing the harmonious index is **NP**-hard even on trees.

- So assuming $\mathbf{P} \neq \mathbf{NP}$, the harmonious index is not **CMSOL**-definable, because trees have tree-width 1.

File:gpar

Non-definability via complexity assumptions: Chromaticity

- B. Courcelle, J.A.M. and U. Rotics proved that
  graph parameters (polynomials) definable in **CMSOL** in the
  language of graphs can be computed in polynomial time
  for graphs of clique-width at most k.

- The Exponential Time Hypothesis (**ETH**) says that $3 - SAT$ cannot be
  solved in time $2^{o(n)}$. It was first formulated by R. Impagliazzo, R. Paturi
  and F. Zane in 2001.

- F. Fomin, P. Golovach, D. Lokshtanov and S. Saurabh proved that,
  assuming that **ETH** holds,
  the chromatic number $\chi(G)$ cannot be computed in polynomial time.

- Therefore, assuming **ETH**, the chromatic number and the chromatic
  polynomial are not **CMSOL**-definable.

There are many other non-definability results which can obtained like this,
for example graph paremeters derived from dominating sets or the size of a
maximal cut.

File:gpar

Our goal is to prove non-definability

without complexity theoretic assumptions.

Typical properties of graph parameters

Multiplicative graph parameters

with respect to ⊔

Let $G_1 \sqcup G_2$ denote the disjoint union of two graphs.

f is **multiplicative** if $f(G_1 \sqcup G_2) = f(G_1) \cdot f(G_2)$.

  (i) $|V(G)|, |E(G)|, k(G)$ are not multiplicative

 (ii) $\chi(G)$ and $\omega(G)$ are not multiplicative

(iii) The number of perfect matchings $pm(G)$ is multiplicative and so is the generating matching polynomial $\sum_k m_k(G)X^k$.
Note that $m_k(G)$ is not multiplicative.

(iv) The graph polynomials $\chi(G, \lambda)$ and $T(G, X, Y)$ are multiplicative.

Additive graph parameters

with respect to ⊔

Let $G_1 \sqcup G_2$ denote the disjoint union of two graphs.

f is **additive** if $f(G_1 \sqcup G_2) = f(G_1) + f(G_2)$.

(i) $|V(G)|, |E(G)|$ are additive.

(ii) $k(G)$ are additive

   Let $b(G)$ be the number of 2-connected components of G.
   $b(G)$ is additive.

(iii) $\chi(G)$ and $\omega(G)$ are not additive

(iv) If f is additive and $r \in \mathbb{R}$, then r^f is multiplicative.

Maximizing and minimizing graph parameters

with respect to $\sqcup$

Let $G_1 \sqcup G_2$ denote the disjoint union of two graphs.

f is **maximizing** if $f(G_1 \sqcup G_2) = \max\{f(G_1), f(G_2)\}$.

f is **minimizing** if $f(G_1 \sqcup G_2) = \min\{f(G_1), f(G_2)\}$.

 (i) The various chromatic numbers $\chi(G)$, $\chi_e(G)$, $\chi_t(G)$ are maximizing.

 (ii) The size of the maximal clique $\omega(G)$ and the maximal degree $\Delta(G)$ are maximizing.

(iii) The tree-width $tw(G)$ and the clique-width $cw(G)$ of a graph are maximizing.

(iv) The minimum degree $\delta(G)$, the girth $g(G)$ are minimizing.

The girth is the minimum length of a cycle in G.

The connection matrix of a graph parameter

with respect to the disjoint union ⊔

Connection matrix $M(f, \sqcup)$.

Let G_i be an enumeration of all finite graphs (up to isomorphism).

The (full) **connection matrix** $M(f, \sqcup) = m_{i,j}(f, \sqcup)$ is defined by

$$m_{i,j}(f, \sqcup) = f(G_i \sqcup G_j)$$

The **rank** of $M(f, \sqcup)$ is denoted by $r(f, \sqcup)$.

We shall often look at various infinite submatrices of the full connection matrix.

****

Examples: Check with $|V(G)|$ and $2^{|V(G)|}$.

Computing $r(f, \sqcup)$

Proposition:

(i) If f is multiplicative, $r(f, \sqcup) = 1$.

(ii) If f is additive, $r(f, \sqcup) = 2$.

(iii) If f is maximizing or minimizing, $r(f, \sqcup)$ is infinite.

(iv) For the average degree $d(G)$ of a graph, $r(d, \sqcup)$ is infinite.

Proof: The first three statements are easy.

For $f = d(G)$ we have

$$M(d, \sqcup) = 2\frac{|E_1| + |E_2|}{|V_1| + |V_2|}.$$

This contains, for graphs with a fixed number e of edges, the Cauchy matrix $(\frac{2e}{i+j})$, hence $r(d, \sqcup)$ is infinite. $\square$.

Characterizing multiplicative graph parameters

M. Freedman, L. Lovász and A. Schrijver, 2007

Theorem: ([FLS] Proposition 2.1.)

Assume $f(G) \neq 0$ for some graph G.
f is multiplicative iff $M(f, \sqcup)$ has rank 1 and is positive semi-definite.

Recall: A finite square matrix M over an ordered field is **positive semi-definite** if for all vectors $\bar{x}$ we have $\bar{x} M \bar{x}^{tr} \geq 0$. An infinite matrix is positive semi-definite, if every finite principal submatrix is positive semi-definite.

General Connection Matrices

General Connection Matrices: I

Let $\mathcal{C}$ be a class possibly labeled graphs, hyper-graphs or τ-structures.

Let $\square$ be a binary operation define on $\mathcal{C}$.

Let G_i be an enumeration of all (labeled) finite graphs

Let f be graph parameter.

The (full) connection matrix $M(f, \square)$ is defined by

$$M(f, \square)_{i,j} = f(G_i \square G_j)$$

and is called the **Full Connection Matrix of f for $\square$ on $\mathcal{C}$**,

or just a **connection matrix**.

We shall often look at infinite submatrices of $M(f, \square)$.

$\mathcal{L}$-smooth operations.

Let $\mathcal{L}$ be a logic.

We say that two graphs G, H are $(\mathcal{L},)q$-equivalent, and write $G \sim_{\mathcal{L}}^q H$, if G and H satisfy the same $\mathcal{L}$-sentences of quantifier rank q.

We say that $\square$ is $\mathcal{L}$-smooth, if wwhenever we have

$$G_i \sim_{\mathcal{L}}^q H_i, i = 0, 1$$

then

$$G_0 \square G_1 \sim_{\mathcal{L}}^q H_0 \square H_1$$

This definition can be adapted to k-ary operations for $k \geq 1$.

Proving that an operation $\square$ is $\mathcal{L}$-smooth may be difficult.

For **FOL** this can be achieved using

Ehrenfeucht-Fraïssé games also know as pebble games.

Anther way of establishing smoothness is via the Feferman-Vaught theorem.

File:cmatrices

Examples of $\mathcal{L}$-smooth operations.

(i) Quantifier-free scalar transductions are both **FOL** and **MSOL**-smooth.

(ii) Quantifier-free vectorized transductions are **FOL** but not **MSOL**-smooth.

(iii) The cartesian product is **FOL**-smooth but not **MSOL**-smooth.

   This was shown by A. Mostowski in 1952.

(iv) The (rich) disjoint union is both **FOL** and **MSOL**-smooth.

   The rich disjoint union has two additional unary predicates to distinguish the universes.

   For **FOL** this was shown by E. Beth in 1952.
   For **MSOL** this is due to H. Läuchli, 1966, using Ehrenfeucht-Fraïssé games

(v) Adding modular counting qunatifiers preserves smoothness.

   For **CMSOL** and the disjoint union this is due to B. Courcelle, 1990.
   For $CFOL$ and the red product this is due to T. Kotek and J.A.M., 2012.

The Finite Rank Theorem

THEOREM (Godlin, Kotek, Makowsky 2008):

Let f be a numeric parameter or polynomial for τ-structures definable in $\mathcal{L}$ and taking values in an integral domain $\mathcal{R}$.

Let $\square$ be an $\mathcal{L}$-smooth operation.

Then the connection matrix $M(f, \square)$ has finite rank over $\mathcal{R}$.

$$*********************$$

The **Proof** uses a Feferman-Vaught-type theorem for graph polynomials, due to B. Courcelle, J.A.M. and U. Rotics, 2000.

Applications of the Finite Rank Theorem, I

Disjoint unions

The following graph parameters or not **CMSOL**-definable

because they are maximizing (minimizing) for the disjoint union.

- $\omega(G)$, the clique number and $\alpha(G)$, the independence number of G.

- The chromatic number $\chi(G)$ and the chromatic index $\chi_e(G)$.

- The degrees $\delta(G)$ (minimal), $\Delta(G)$ (maximal)

The same holds for the average degree $d(G)$, but here we use the fact that the Cauchy matrix growing rank.

Applications of the Finite Rank Theorem, I

Products combined with translation schemes
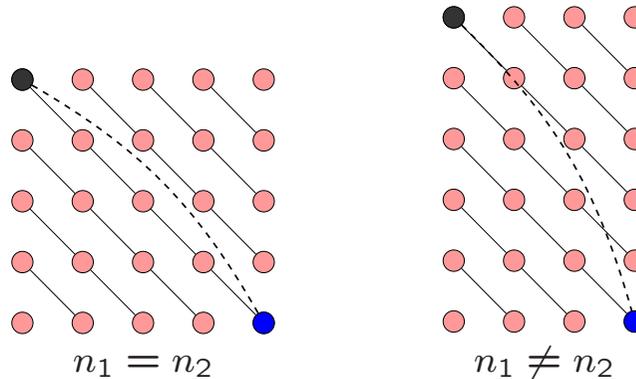

The transduction

$$\Phi_F\left((v_1, v_2), (u_1, u_2)\right) \quad = \quad (E_1(v_1, u_1) \wedge E_2(v_2, u_2)) \vee$$
$$((v_1, v_2), (u_1, u_2)) = ((start_1, start_2), (end_1, end_2))$$

transforms the caresian product of two directed paths $P_{n_i}^i = (V_1, E_1, start_i, end_i)$ of length n_i with the two constants $start_i$ and end_i, $i = 1, 2$ into an undirected graph with atmost one cycle.

The input graphs look like this:

The result of the transduction is:

$$n_1 = n_2 \qquad\qquad n_1 \neq n_2$$

THEOREM: Graphs without cycles of odd (even) length

are not **CFOL**-definable even in the presence of a linear order.


Corollary: Not definable in **CFOL** with order are

 (i) Forests, bipartite graphs, chordal graphs, perfect graphs

 (ii) interval graphs (cycles are not interval graphs)

(iii) Block graphs (every biconnected component is a clique)

(iv) Parity graphs (any two induced paths joining the same pair of vertices
      have the same parity)

THEOREM: Trees or connected graphs are not **CFOL**-definable even in the presence of linear order.

The transduction

$$\Phi_T\left((v_1, v_2), (u_1, u_2)\right) \;=\; \begin{aligned}&(E_1(v_1, u_1) \wedge E_2(v_2, u_2)) \vee \\ &(v_1 = u_1 = start_1 \wedge E(v_2, u_2)) \vee \\ &(v_1 = u_1 = end_1 \wedge E(v_2, u_2)),\end{aligned}$$

combined with Φ_{sym} transforms the cartesian product of two directed paths into the structures below:

$$n_1 > n_2 \qquad\qquad n_1 = n_2 \qquad\qquad n_1 < n_2$$
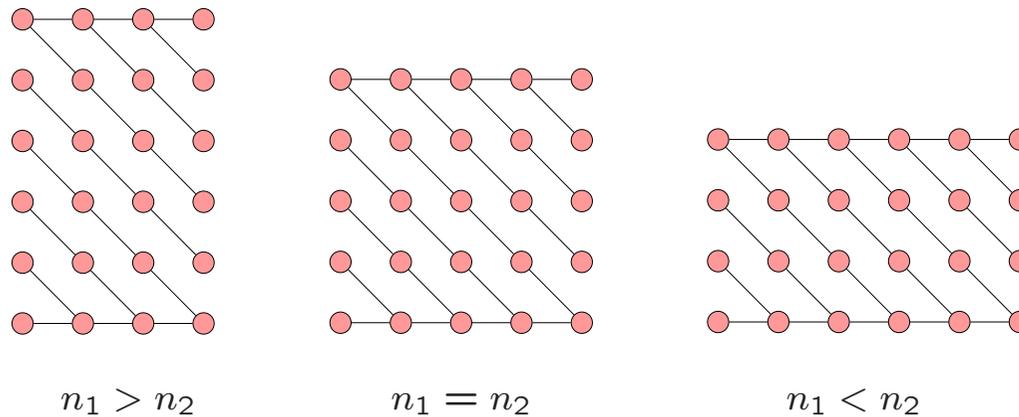
Tree: $n_1 = n_2$. Connected: $n_1 \geq n_2$

File:cmatrices

k-graphs and k-sums

A k-graph is a graph $G = (V(G), E(G))$

with k distinct vertices labeled with $0, 1, \ldots, k-1$.

Given two k-graphs G_1, G_2 we define the k-sum

$$G_1 \sqcup_k G_2$$

 as the disjoint union of G_1 and G_2 where we

identify correspondingly labeled vertices.

Theorem: The k-sum is smooth for $\mathbf{FOL}, \mathbf{CFOL}, \mathbf{MSOL}$ and $\mathbf{CMSOL}$.
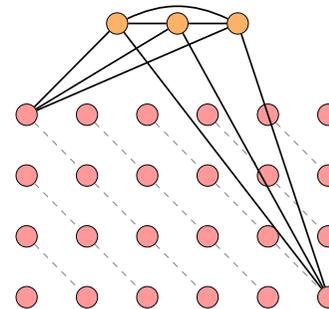
THEOREM: Planar graphs are not **CFOL**-definable even on ordered

connected graphs

For our next connection matrix we use the 2-sum of the following two 2-graphs:

- the 2-graph (G, a, b) obtained from from K_5 by choosing two vertices a and b and removing the edge between them

- the cartesian product of the two graphs $P_{n_1}^1$ and $P_{n_2}^2$:
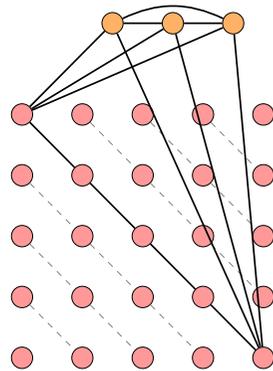
$$n_1 = n_2 \qquad\qquad n_1 \neq n_2$$

The result of this construction has a clique of size 5 as a minor iff $n_1 = n_2$.

It can never have a $K_{3,3}$ as a minor.

A modification ....

If we modify the above construction by taking K_3 instead of K_5 and making $(start_1, start_2)$ and (end_1, end_2) adjacent, we get

Proposition: The following classes of graphs are not **CFOL**-definable even on ordered connected graphs.

(i) Cactus graphs, i.e. graphs in which any two cycles have at most one vertex in common.

(ii) Pseudo-forests, i.e. graphs in which every connected components has at most one cycle.

Non-definability in $\mathbf{CMSOL}$ for graphs $G = (V, E)$

Using the join operation

The join operation of graphs $G = (V, E)$, where E is the edge relation, is defined by
$$(V_1, E_1) \bowtie (V_2, E_2) = (V_1 \sqcup V_2, E_1 \sqcup E_2 \cup \{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2\}$$
This is a quantifier free transduction of the dosjoint union, hence smooth for $\mathbf{CMSOL}$.

Consider the connection matrix where the rows and columns are labeled by the graphs on n vertices but without edges E_n.

The graph $E_i \bowtie E_j = K_{i,j}$ is

- hamiltonian iff $i = j$;

- has a perfect matching iff $i = j$;

- is a cage graph (a regular graph with as few vertices as possible for its girth) iff $i = j$;

- is a well-covered graph (every minimal vertex cover has the same size as any other minimal vertex cover) iff $i = j$.

All of these connection matrices have infinite rank.

Proposition: None of the properties above are $\mathbf{CMSOL}$-definable as graphs even in the presence of an order.

File:cmatrices

CMSOL for hyper-graphs $G = (V, E; R)$

A hyper-graphs $G = (V, E; R)$ has vertices V and edges E

and an incidence relationa R between the two.

- **CMSOL** for hyper-graphs $G = (V, E; R)$ allows quantification over edge sets.

- For the language of hypergraphs the join operation is

  neither **MSOL**- nor **CMSOL**-smooth,

  since it increases the number of edges.

- Note also that hamiltonicity and having a perfect matching are both definable in **CMSOL** in the language of hypergraphs.

In the many papers of B. Courcelle, **MSOL** on graphs is called $\mathbf{MSOL}_1$

and for hyper-graphs it is called $\mathbf{MSOL}_2$.

Non-definability in $\mathbf{CMSOL}$ for hyper-graphs $G = (V, E; R)$

Using the disjoint union

Using the connection submatrices of the disjoint union we still get:

- Regular: $K_i \sqcup K_j$ is regular iff $i = j$;

- A generalization of regular graphs are *bidegree* graphs, i.e., graphs where every vertex has one of two possible degrees. $K_i \sqcup (K_j \sqcup K_1)$ is a bidegree graph iff $i = j$.

- The average degree of $K_i \sqcup E_j$ is at most $\frac{|V|}{2}$ iff $i = j$;

- A digraph is *aperiodic* if the common denominator of the lengths of all cycles in the graph is 1. We denote by C_i^d the directed cycle with i vertices. For prime numbers p, q the digraphs $C_p \sqcup C_q$ is aperiodic iff $p \neq q$.

- A graph is *asymmetric (or rigid)* if it has no non-trivial automorphisms. It was shown by P. Erdös and A. Rényi (1963) that almost all finite graphs are asymmetric. So there is an infinite set $I \subseteq \mathbb{N}$ such that for $i \in I$ there is an asymmetric graph R_i of cardinality i. $R_i \sqcup R_j$ is asymmetric iff $i \neq j$.

Proposition: None of the properties above are $\mathbf{CMSOL}$-definable as hyper-graphs even in the presence of an order.

Proving non-definability with connection matrices: Merits

The advantages of the Finite Rank Theorem for tame $\mathcal{L}$ in proving that a property is not definable in $\mathcal{L}$ are the following:

(i) It suffices to prove that certain binary operations on graphs (τ-structures) are $\mathcal{L}$-smooth operation.

(ii) Once the $\mathcal{L}$-smoothness of a binary operation has been established, proofs of non-definability become surprisingly simple and transparent.

One of the most striking examples is the fact that asymmetric (rigid) graphs are not definable in **CMSOL**.

(iii) Many properties can be proven to be non-definable using the same or similar submatrices, i.e., matrices with the same row and column indices. This was well illustrated in the shown examples.

Proving non-definability with connection matrices: Limitations

The classical method of proving non-definability in **FOL** using pebble games is complete in the sense that a property is $\mathbf{FOL}(\tau)_q$-definable iff the class of its models is closed under game equivalence of length q.

Using pebble games one proves easily that the class of structures without any relations of even cardinality, EVEN, is not **FOL**-definable.

However, one cannot prove that EVEN is not **FOL**-definable using infinite rank connection matrices, in the following sense:

Proposition: Let Φ a quantifierfree transduction between τ-structures and let $\square_\Phi$ be the binary operation on τ-structures:

$$\square_\Phi(\mathfrak{A}, \mathfrak{B}) = \Phi^\star(\mathfrak{A} \sqcup_{rich} \mathfrak{B})$$

Then the connection matrix $M(\square_\Phi, \mathsf{EVEN})$ satisfies:

(i) There is a finite partition $\{U_1, \ldots, U_k\}$ of the (finite) τ-structures such that the submatrices obtained from restricting $M(\square, \psi)$ to $M(\mathsf{EVEN}, \square_\Phi)^{[U_i, U_j]}$ have constant entries.

(ii) In particular, the infinite matrix $M(\mathsf{EVEN}, \square_\Phi)$ has finite rank over any field $\mathcal{F}$.

(iii) $M(\mathsf{EVEN}, \square_\Phi)$ has an infinite submatrix of rank at most 1.

Note that EVEN is trivially definable in **CFOL**.

File:cmatrices

Lecture 2

Counting weighted homomorphisms

aka

Partition functions

aka

Vertext coloring models

Lecture 2: Overview

- Counting weighted homomorphisms

- The original connection matrices

- Representability

- Representation theorems

Counting weighted homomorphisms

Let $H = (V(H), E(H))$ be a fixed graph, possibly with loops.
Let $\alpha : V(H) \to \mathbb{R}^+$ and $\beta : E(H) \to \mathbb{R}$ be weight functions of vertices and edges respectively.
Let $h : G \to H$ be a homomorphism.

Define weights of h by

$$\alpha_h = \prod_{u \in V(G)} \alpha(h(u)) \quad \text{and} \quad \beta_h = \prod_{u,v \in E(G)} \beta(h(u), h(v))$$

Finally, we sum over all homomorphisms

$$Z_{H,\alpha,\beta}(G) = \sum_{h:G \to H} \alpha_h \cdot \beta_h$$

$Z_{H,\alpha,\beta}(G)$ is often called a **partion function** or a **vertex coloring model**.

- $Z_{H,\alpha,\beta}(G)$ is multiplicative.

- For α, β constant $= 1$, $Z_{H,\alpha,\beta}(G)$ counts the number of homomorphisms.

Examples of partition functions

- For $H = K_m$, a clique with m vertices,

$$Z_{K_m,1,1}(G) = \chi(G, m)$$

  which counts the number of proper m-colorings.

- For $H = L_1$, an isolated loop, $\alpha = \lambda$, $\beta = \mu$,

$$Z_{L_1,\lambda,\mu}(G) = \lambda^{|V(G)|} \cdot \mu^{|E(G)|}$$

- For $H = L_m$ consisting of m isolated loops, $\alpha = \lambda$, $\beta = \mu$,

$$Z_{L_m,\lambda,\mu}(G) = m^{k(G)} \cdot \lambda^{|V(G)|} \cdot \mu^{|E(G)|}$$

- For $H = K_1 \bowtie L_1$ with vertices v, ℓ respectively, and $\alpha(v) = X, \alpha(\ell) = 1$, $\beta = 1$ we get

$$Z_{K_1 \bowtie L_1,\alpha,\beta}(G) = \sum_i ind_i(G) \cdot X^i$$

  where $ind_i(G)$ is the number of independents sets of size i in G.

k-connection matrices

The name connection matrices originates here.

M. Freedman, L. Lovász and A. Schrijver coined the term

with k-unions in mind, thinking of k-connections.

k-graphs and their connection matrices

A k-graph is a graph $G = (V(G), E(G))$ with k distinct vertices labeled with $0, 1, \ldots, k-1$.

Let $G_0, G_1, \ldots$ be an enumeration of all k-graphs up to label preserving isomorphisms.

Given two k-graphs G_1, G_2 we define the k-sum $G_1 \sqcup_k G_2$ as the disjoint union of G_1 and G_2 where we identify correspondingly labeled vertices.

Given a numeric graph invariant f, we define the k-**connection matrix**

$$M(f, k) = m_{i,j}(f, k)$$

by

$$m_{i,j}(f, k) = f(G_i \sqcup_k G_j)$$

We denote by $r(f, k)$ the **rank of** $M(f, k)$.

Properties of k-connection matrices, I

Examples:

(i) $M(\omega(G), k)$ has infinite rank.

(ii) $M(|V(G)|, k)$ has rank 2 for every $k \geq 0$.

(iii) $M(2^{|V(G)|}, k)$ has rank $r(2^{|V(G)|}, k) = 1$ for every $k \geq 0$.

(iv) $M(2^{k(G)}, k)$ has rank $r(2^{k(G)}, k) = 2^k$ for every $k \geq 0$.

Proposition:([FLS] Proposition 2.2)

Let f be a multiplicative numeric graph invariant, and $k, \ell \geq 0$. Then

$$r(f, k + \ell) \geq r(f, k) \cdot r(f, \ell)$$

Properties of k-connection matrices, II

Let f be a graph parameter.

(i) f is k-additive if for all k-labeled graphs $G_1, G_2 \in \mathcal{G}_k$ we have

$$f(G_1 \sqcup_k G_2) = f(G_1) + f(G_2).$$

(ii) f is k-multiplicative if for all k-labeled graphs $G_1, G_2 \in \mathcal{G}_k$ we have

$$f(G_1 \sqcup_k G_2) = f(G_1) \cdot f(G_2).$$

(iii) f is k-maximizing if for all k-labeled graphs $G_1, G_2 \in \mathcal{G}_k$ we have

$$p(G_1 \sqcup G_2, \bar{X}) = \max\{p(G_1, \bar{X}), p(G_2, \bar{X})\},$$

and similarly for k-minimizing.

Properties of k-connection matrices, III

Proposition:

(i) If f is k-multiplicative, $r(f, k) = 1$.

(ii) If f is k-additive, $r(f, k) = 2$.

(iii) If f is k-maximizing or k-minimizing, $r(f, k)$ is infinite.

(iv) If $p(G, X)$ is a multiplicative (additive) graph polynomial,
and $d_p(G)$ is its degree,
then $d_p(G)$ is additive (resp. maximizing).

Properties of k-connection matrices, IV

Theorem:([FLS] Lemma 2.3 and [L2] Corollary 2.3)

For every weighted graph (H, α, β)

$$r(Z_{H,\alpha,\beta}(G), k) \leq |V(H)|^k$$

If (H, α, β) has no automorphisms and no twins, then

$$r(Z_{H,\alpha,\beta}(G), k) = |V(H)|^k$$

Definitions:

Autmorphisms here are weight preserving.

Two vertices $u, v \in V(H)$ of (H, α, β) are **twins**

if for every $w \in V(H)$ we have that $\beta(u, w) = \beta(v, w)$.

Being twins does not depend on α.

Properties of k-connection matrices, V

Theorem:([FLS] Lemma 2.3)

For every weighted graph (H, α, β)
the matrix $M(Z_{H,\alpha,\beta}(G), k)$ is positive semi-definite.

Examples:([FLS] Section 3)

(i) Let $pm(G)$ denote the number of perfect matchings of G.
$pm(G)$ is multiplicative and $r(pm, k) = 2^k$,
but $M(pm, 1)$ is not positive definite.

(ii) For $\chi(-, \lambda)$, $\lambda \in \mathbb{Z}$ we have:
$M(\chi(-, \lambda), k)$ is positive-semi-definite,
and $r(\chi(-, \lambda), k)$ is finite,
but exponentially bounded only for $\lambda \in \mathbb{Z}^+$,
otherwise it grows superexponentially.

Representation theorems

The FLS-representation theorem

We say that a a numeric graph invariant is **hom-presentable** if there is a weighted graph (H, α, β) such that for every G

$$f(G) = Z_{H,\alpha,\beta}(G)$$

Examples:

(i) $|V(G)|, |E(G)|, k(G)$ are not hom-presentable, but
$2^{|V(G)|}, 2^{|E(G)|}, 2^{k(G)}$ are hom-presentable.

(ii) $\chi(-, \lambda)$ is hom-representable for every $\lambda \in \mathbb{Z}^+$,
but the **choice of** (H, α, β) **depends on** λ.

Theorem:([FLS] Theorem 2.4)

f is hom-presentable iff for every $k \in \mathbb{N}$

(i) $M(f, k)$ is positive semi-definite, and

(ii) $r(f, k) \leq q^k$ for some $q \in \mathbb{N}^+$.

Other representation theorems

There are generalizations:

- B. Szegedy [Sz] considers **edge coloring models** and connection matrices $S(f,k)$ based on identification of k unfinished edges.

- A. Schrijver [Schr-08, Schr-08a] unifies the proofs of [FLS] and [Sz] using further variations of connection matrices $S_i(f,k)$ with $i = 1,\ldots,10$ including also hyper-graphs and directed graphs.

****

In principle one can define connection matrices

$$M(f,\otimes_k)$$

for any binary operation $\otimes_k$ on k-graphs.

Lecture 3:

Parametrized numeric graph invariants

and

graph polynomials

Lecture 3: Overciew

- Parametrized graph parameters and graph polynomials

- The prominet graph polynomials:

  - Characteristic polynomial,

  - Tutte polynomial,

  - Cover polynomial

- Definability of graph polynomials

- Conclusions

Parametrized numeric graph invariants

A **parametrized numeric graph invariant** is a function

$$f : \mathcal{G} \times R \to \mathbb{R}$$

which is invariant under graph isomorphism.
Here R can be $\mathbb{N}, \mathbb{Z}, \mathbb{R}$ or any ring.

Examples:

(i) $ind_k(G)$ the number of independent sets of size k.

(ii) $ind(G, X) = \sum_i ind(G, i) \cdot X^i$, the independent set polynomial.

(iii) The chromatic polynomial $\chi(G, \lambda)$.

(iv) Any graph polynomial from the literature, like matching polynomials, Tutte polynomial, interlace polynomial, cover polynomial of directed graphs, etc.

Coding many graph parameters into a graph polynomial

A particular graph polynomial is considered

interesting

if it encodes many useful graph parameters.

The characteristic polynomial

Let $G = (V(G), E(G))$ be a graph.

The characteristic polynomial $P(G, X)$ is defined as the characteristic polynomial (in the sense of linear algebra) of the adjacency matrix A_G of G defined as

$$det(X \cdot \mathbf{1} - A_G) = \sum_{i=0}^{n} c_i(G) \cdot X^{n-i}.$$

It is well known that

(i) $n = |V(G)|$

(ii) $-c_2(G) = |E(G)|$

(iii) $-c_3(G)$ equals twice the number of triangles of G.

(iv) The second largest zero $\lambda_2(G)$ of $P(G; X)$ gives a lower bound to the conductivity of G

The Tutte polynomial

The Tutte polynomial of G is defined as

$$T(G; X, Y) = \sum_{F \subseteq E(G)} (X - 1)^{r\langle E \rangle - r\langle F \rangle} (Y - 1)^{n\langle F \rangle}$$

where $k\langle F \rangle$ is the number of connected components of the spanning subgraph defined by F,

$r\langle F \rangle = |V| - k\langle F \rangle$ is its rank

and $n\langle F \rangle = |F| - |V| + k\langle F \rangle$ is its nullity.

Evaluations of the Tutte polynomial

See D. Welsh, Complexity: Knots, colourings and counting, Cambridge, 1993, and M. Korn and I. Pak, Tilings of rectangles with T-tetrominoes, TCS 319, 2004

(i) $T(G; 1, 1)$ is the number of spanning trees of G,

(ii) $T(G; 1, 2)$ is the number of connected spanning subgraphs of G,

(iii) $T(G; 2, 1)$ is the number of spanning forests of G,

(iv) $T(G; 2, 2)$ is the number of spanning subgraphs of G,

(v) $T(G; 1 - k, 0)$ is the number of proper k-vertex colorings of G,

(vi) $T(G; 2, 0)$ is the number of acyclic orientations of G,

(vii) $T(G; 0, -2)$ is the number of Eulerian orientations of G.

(viii) $2 \cdot T(Grid_{4x,4y}; 3, 3)$ is the number of tilings of the $(4x \times 4y)$- grid graph with T-tetrominoes

The cover polynomial

Chung and Graham, 1995 and D'Antona and Munarini, 2000

Let $D = (V, E)$ be a directed graph.

$C \subseteq E$ is a **path-cycle cover of** G if C is a subgraph with maximal in-degree ≤ 1 and maximal out-degree ≤ 1 and C is a vertex disjoint decomposition of E with $p(C)$ paths and $c(C)$ cycles.

The **(factorial) cover polynomial** is the polynomial

$$C(D, x, y) = \sum_C (x)_{p(C)} \cdot y^{c(C)}$$

The **(geometric) cover polynomial** is the polynomial

$$C_{geom}(D, x, y) = \sum_C (x)^{p(C)} \cdot y^{c(C)}$$

Here $(x)_n = x \cdot (x - 1) \cdot \ldots \cdot (x - n + 1)$ is the falling factorial.

Evaluations of the cover polynomial

(i) $C(D, 0, 1)$ is the number of cycle covers of D, which is the permanent of the adjacency matrix of D.

(ii) $C(D, 0, -1)$ is the determinant of the adjacency matrix of D.

(iii) $C(D, 1, 0)$ is the number of hamiltonian paths of D.

(iv) $C(D, x, 1)$ is the *factorial rook polynomial* of D.

Evaluations, coefficients and zeros of graph polynomials

How could one prove that a graph parameter f is

not coded

in a given graph polynomial from an infinite class of graph polynomials $\mathcal{P}$ as

- an evaluation?

- a coefficient?

- a zero?

We will again use **connection matrices**!

Definability of graph polynomials

in Monadic Second Order Logic MSOL

Simple MSOL-definable graph polynomials

The graph polynomial $ind(G, X) = \sum_i ind(G, i) \cdot X^i$, can be written also as

$$ind(G, X) = \sum_{I \subseteq V(G)} \prod_{v \in I} X$$

where I ranges over all independent sets of G.
To be an independent set is definable by a formula of Monadic Second Order Logic (MSOL) $\phi(I)$.

A simple MSOL-definable graph polynomial $p(G, X)$ is a polynomial of the form

$$p(G, X) = \sum_{A \subseteq V(G) : \phi(A)} \prod_{v \in I} X$$

where A ranges over all subsets of $V(G)$ satisfying $\phi(A)$ and $\phi(A)$ is a MSOL-formula.

File:msol

General MSOL-definable graph polynomials

For the general case

- One allows several indeterminates $X_1, \ldots, X_t$.

- One gives an inductive definition.

- One allows an ordering of the vertices.

- One requires the definition to be invariant under the ordering, i.e., different orderings still give the same polynomial.

- This also allows to define the modular counting quantifiers
  $C_{m,q}$ "there are, modulo q exactly m elements..."

The general case includes the Tutte polynomial, the cover polynomial, and virtually all graph polynomials from the literature.

Finiteness of $r(f, \Box)$ for $\mathcal{L}$-smooth operations $\Box$.

Finite Rank Theorem:([GKM])

Let $p(G, \bar{X})$ be a general $\mathcal{L}$-definable graph polynomial
in t indeterminates, and $\bar{x} \in \mathbb{R}^t$.

For $f(G) = p(G, \bar{x})$ the rank $r(f, \Box)$ is finite, provide $\Box$ is $\mathcal{L}$-smooth.

Proof: We use the bilinear version of the Feferman-Vaught Theorem for graph polynomials from [M] to estimate the $r(f, \Box)$.

The estimates are very bad, and just suffice to establish finiteness of $r(f, \Box)$,

but they grow with multiple exponentials in k.

File:msol

Applications of the Finite Rank Theorem, I

Corollary:[GKM] The following numeric graph invariants f have $r(f, k) < \infty$:

(i) The number of acyclic orientations

(ii) The number of eulerian orientations

Proof: They are both instances of the Tutte polynomial.

Applications of the Finite Rank Theorem, II

Corollary:([GKM])

$\omega(G)$ is not an instance of any MSOL-definable graph polynomial, but is the degree of some MSOL-definable graph polynomial,

Proof: $\omega(G_1 \sqcup G_2) = \max\{\omega(G_1), \omega(G_2)\}$. So $r(\omega, 0) = \infty$.

$\omega(G)$ can be obtained as degree of the graph polynomial

$$clique(G, X) = \sum_i clique_i(G)X^i = \sum_{C \subseteq V} \prod_{v \in C} X$$

where $clique_i(G)$ is the number of cliques of size i, and C varies over all cliques of G.

Clearly, $clique(G, X)$ is a simple MSOL-definable graph polynomial. $\square$

Applications of the Finite Rank Theorem, III

Corollary:([GKM])

(i) If f satisfies $f(G_1 \sqcup G_2) = \max\{f(G_1), f(G_2)\}$, then f is not an instance of an MSOL-definable graph polynomial.

(ii) If f satisfies $f(G_1 \sqcup G_2) = \min\{f(G_1), f(G_2)\}$, then f is not an instance of an MSOL-definable graph polynomial.

Applications of the Finite Rank Theorem, IV

Let $d(G)$ denote the average degree of G. We have

$$d(G) = \frac{1}{|V(G)|} \cdot \sum_{v \in V(G)} d_G(v),$$

where $d_G(v)$ denotes the degree of a vertex v of G.

Corollary:([GKM])

$d(G)$ is not an instance of an MSOL-definable graph polynomial.

Proof: For $f = d(G)$ we have

$$M(d, 0) = 2\frac{|E_1| + |E_2|}{|V_1| + |V_2|}.$$

This contains, for graphs with a fixed number e of edges, the Cauchy matrix

$$(\frac{2e}{i+j}),$$

hence $r(d, 0)$ is infinite.   □

Three graph polynomials, I

Rainbow polynomial $\chi_{rainbow}(G, k)$ is the number of path-rainbow connected k-colorings, which are functions $c : E(G) \to [k]$ such that between any two vertices $u, v \in V(G)$ there exists a path where all the edges have different colors.

MCC-polynomial For every fixed $t \in \mathbb{N}$, $\chi_{mcc(t)}(G, k)$ is the number of vertex k-colorings $f : V(G) \to [k]$ for which every color induces a subgraph with a **c**onnected **c**omponent of **m**aximal size t.

Convex coloring polynomial $\chi_{convex}(G, k)$ is the number of convex colorings, i.e., vertex k-colorings $f : V(G) \to [k]$ such that every color induces a connected subgraph of G.

Makowsky and B. Zilber (2005) showed that $\chi_{rainbow}(G, k)$, $\chi_{mcc(t)}(G, k)$, and $\chi_{convex}(G, k)$ are graph polynomials with k as the variable.

Path-rainbow connected colorings were introduced by G. Chartrand et al. in 2008. Their complexity was studied in S. Chakraborty et. al in 2008. $mcc(t)$-colorings were first studied by N. Alon et al. in 2003.
Note $\chi_{mcc(1)}(G, k)$ is the chromatic polynomial.

Convex colorings were studied by S. Moran in 2007.

File:msol

Three graph polynomials, II

Proposition: The following connection matrices have infinite rank:

(i) $M(\sqcup_1, \chi_{rainbow}(G,k)$;

(ii) $M(\sqcup_1, \chi_{convex}(G,k)$;

(iii) For every $t > 0$ the matrix $M(\bowtie, \chi_{mcc(t)}(G,k)$;

Proof:

$\chi_{rainbow}(G,k)$: We use that the 1-sum of paths with one end labeled is again a path with $P_i \sqcup_1 P_j = P_{i+j-1}$ and that $\chi_{rainbow}(P_r,k) = 0$ iff $r > k+3$.

$\chi_{convex}(G,k)$: We use edgeless graphs and disjoint union $E_i \sqcup E_j = E_{i+j}$ and that $\chi_{convex}(E_r,k) = 0$ iff $r > k$.

$\chi_{mcc(t)}(G,k)$: We use the join and cliques, $K_i \bowtie K_j = K_{i+j}$ and that $\chi_{mcc(t)}(K_r,k) = 0$ iff $r > kt$.

Three graph polynomials, III

Corollary:

(i) $\chi_{rainbow}(G,k)$ and $\chi_{convex}(G,k)$ are not an **CMSOL**-definable in the language of graphs and hypergraphs.

(ii) $\chi_{mcc(t)}(G,k)$ (for any fixed $t > 0$) is not **CMSOL**-definable in the language of graphs.

(iii) In particular the chromatic polynomial is not **CMSOL**-definable in the language of graphs.

Note: It is however **CMSOL**-definable in the language of ordered hypergraphs.

Proof:
(i) The 1-sum and the disjoint union are **CMSOL**-sum-like and **CMSOL**-smooth for hypergraphs.
(ii) The join is only **CMSOL**-sum-like and **CMSOL**-smooth for graphs.

Specific coefficients of graph polynomials

Let $p(G, \bar{X})$ be an $MSOL$-definable graph polynomial with values in $\mathbb{R}[\bar{X}]$ with m indeterminates $X_1, \ldots, X_m$, and let

$$X_1^{\alpha_1} \cdot X_2^{\alpha_2} \cdot \ldots \cdot X_m^{\alpha_m}$$

be a specific monomial of $p(G, \bar{X})$ with coefficient $c_\alpha(G)$, where $\alpha = (\alpha_1, \ldots, \alpha_m)$.

Theorem:[GKM] Then there is an invariantly $MSOL$-definable graph polynomial $p_\alpha(G, \bar{X})$ such that $c_\alpha(G)$ is an evaluation of $p_\alpha(G, \bar{X})$.

Remark: The theorem remains valid for monomials of the form

$$X_1^{n_1(G)-\alpha_1} \cdot X_2^{n_2(G)-\alpha_2} \cdot \ldots \cdot X_m^{n_m(G)-\alpha_m},$$

where $n_i(G) = |V(G)|$ or $n_i(G) = |E(G)|$.

This can be used to treat the coefficient of $X^{|V(G)|-3}$ of the characteristic polynomial.

Open problems

(i) For what f is $r(f, k)$ always finite?

(ii) For what f is $M(f, k)$ always positive semi-definite?

(iii) What can we say in general about the various connection matrices of additive graph parameters?

In particular, are there representation theorems for additive graph parameters?

(iv) What can we say in general about the various connection matrices of maximizing (minimizing) graph parameters?

Thank you for your attention!

References, I

[KMZ-2011] T. Kotek and J.A. Makowsky and B. Zilber,
    On Counting Generalized Colorings,
    in: Model Theoretic Methods in Finite Combinatorics,
    Contemporary Mathematics, vol 558 (2011) pp. 207-242
    American Mathematical Society

[Kotek-Thesis] T. Kotek,
    Definability of combinatorial functions,
    Ph.D. Thesis, Submitted: March 2012,
    Technion - Israel Institute of Technology, Haifa, Israel

[FLS] M. Freedman, L. Lovász and A. Schrijver:
    Reflection positivity, rank connectivity, and homomorphism of graphs,
    J. AMS 20 (2007), 37-51

[L1] L. Lovász: Connection matrices,
    in: Combinatorics, Complexity, and Chance, A Tribute to Dominic Welsh,
    Oxford Univ. Press (2007), 179-190

References, II

[CMR] B. Courcelle, J.A. Makowsky and U. Rotics:
On the Fixed Parameter Complexity of Graph Enumeration Problems
Definable in Monadic Second Order Logic,
Discrete Applied Mathematics, 108.1-2 (2001) 23-52

[M] J.A. Makowsky:
Algorithmic uses of the Feferman-Vaught theorem,
Annals of Pure and Applied Logic, 126.1-3 (2004) 159-213

[M-zoo] J.A. Makowsky:
From a Zoo to a Zoology: Towards a general theory of graph
polynomials, Theory of Computing Systems, vol. 43 (2008) pp. 542-562.

[GKM08] B. Godlin, T. Kotek and J.A. Makowsky:
Evaluation of graph polynomials,
34th International Workshop on Graph-Theoretic Concepts in Computer
Science, WG08, LNCS 5344 (2008) 183-194

File:ref