Finite Model Theory and Graph Isomorphism. III.

Anuj Dawar

University of Cambridge Computer Laboratory
visiting RWTH Aachen

Beroun, 14 December 2013

Recapitulation

We obtain *stratifications* of the relation of graph isomorphism by considering *game equivalences* arising in finite model theory.

These give rise to *polynomial time decidable* approximations of graph isomorphism.

These can be computed by a *iterated refinement* of partitions of the k-tuples of vertices.

Recapitulation. II

The family of equivalences $\equiv^{C^k}$ has many equivalent characterisations in terms of

- *Combinatorics*: as the k-dimensional Weisfeiler-Lehman method.
- *Logic:* equivalence in the logic with k variables and counting.
- *Games:* bijection games, counting games

and, to come

- *Linear Programming* relaxations of the isomorphism problem.
- *Algebra*: Coherent algebras.

Graph Isomorphism Integer Program

Yet another way of approximating the *graph isomorphism relation* is obtained by considering it as a *0/1 linear program*.

If A and B are adjacency matrices of graphs G and H, then $G \cong H$ if, and only if, there is a *permutation matrix* P such that:

$$PAP^{-1} = B \quad \text{or, equivalently} \quad PA = BP$$

Introducing a variable x_{ij} for each entry of P and adding the constraints:

$$\sum_i x_{ij} = 1 \quad \text{and} \quad \sum_j x_{ij} = 1$$

we get a system of equations that has a *0-1 solution* if, and only if, G and H are isomorphic.

Fractional Isomorphism

To the system of equations:

$$PA = BP; \quad \sum_i x_{ij} = 1 \quad \text{and} \quad \sum_j x_{ij} = 1$$

add the inequalities

$$0 \leq x_{ij} \leq 1.$$

Say that G and H are *fractionally isomorphic* ($G \cong^f H$) if the resulting system has *any real solution*.

$G \cong^f H$ if, and only if, $G \equiv^{C^2} H$.

(Ramana, Scheiermann, Ullman 1994)

Equitable Partitions

An equivalence relation $\equiv$ on the vertices of a graph $G = (V, E)$ induces an *equitable partition* if

for all $u, v \in V$ *with* $u \equiv v$ *and each* $\equiv$*-equivalence class* S,

$$|\{w \in S \mid (u, w) \in E\}| = |\{w \in S \mid (v, w) \in E\}|.$$

The *naive vertex classification* algorithm finds the *coarsest* equitable partition of the vertices of G.

Equitable Partition to Fractional Isomorphism

Let $G \equiv^{C^2} H$ and let A and B be the respective *adjacency matrices* of G and H.

For each $u \in V(G)$, let $d_u = |\{v \in V(G) \mid u \equiv^{C^2} v\}|$.

Define the matrix P by

$$P_{uv} = \begin{cases} \frac{1}{d_u} & \text{if } u \equiv^{C^2} v \\ 0 & \text{otherwise} \end{cases}$$

Then, $PA = BP$.

Fractional Isomorphism to Equitable Partition

Suppose $G \cong^f H$ and this is witnessed by a *doubly stochastic* matrix P such that:

$$PA = BP$$

For $u, v \in V(G)$, let $u \sim v$ if there is some $w \in V(H)$ such that

$$P_{uw} P_{wv} > 0.$$

Then, we can show that the partition induced by the relation $\sim$ is an *equitable partition*.

Sherali-Adams Hierarchy

If we have any *linear program* for which we seek a *0-1 solution*, we can relax the constraint and admit *fractional solutions*.

The resulting linear program can be solved in *polynomial time*, but admits solutions which are not solutions to the original problem.

Sherali and Adams (1990) define a way of *tightening* the linear program by adding a number of *lift and project* constraints.

Sherali-Adams Hierarchy

The kth *lift-and-project* of a linear program is defined as follows:

For each constraint $\mathbf{a}^T\mathbf{x} = b$ in the linear program, and each set I of variables with $|I| < k$ and $J \subseteq I$, multiply the constraint by

$$\prod_{i \in I \setminus J} x_i \prod_{j \in J} (1 - x_j)$$

and then *linearize* by replacing x_i^2 by x_i and $\prod_{j \in K} x_j$ by a new variable y_K for each set K (along with constraints: $y_\emptyset = 1$, $y_{\{x\}} = x$ and $y_K \leq y_{K'}$ for $K' \subseteq K$).

Say that $G \cong^{f,k} H$ if the kth lift-and-project of the *isomorphism program* on G and H admits a solution.

Sherali-Adams Isomorphism

For each k

$$\equiv^{C^{k+1}} \subseteq \cong^{f,k} \subseteq \equiv^{C^k}$$

(Atserias, Maneva 2012)

For $k > 2$, the inclusions are strict. **(Grohe, Otto 2012)**

Grohe, Otto also describe versions of the k-pebble game corresponding exactly to $\cong^{f,k}$ and variations on *Sherali-Adams* relaxations of isomorphism corresponding exactly to $\equiv^{C^k}$.

Limitations of FPC

There are polynomial-time decidable properties of graphs that are not definable in FPC. **(Cai, Fürer, Immerman, 1992)**

More precisely, we can construct a sequence of pairs of graphs $G_k, H_k (k \in \omega)$ such that:

- $G_k \equiv^{C^k} H_k$ for all k.
- There is a polynomial time decidable class of graphs that includes all G_k and excludes all H_k.

Still, FPC is a *natural* level of expressiveness within P.

Restricted Graph Classes

If we restrict the class of structures we consider, FPC may be powerful enough to express all polynomial-time decidable properties.

1. FPC captures P on *trees*. **(Immerman and Lander 1990)**.
2. FPC captures P on any class of graphs of *bounded treewidth*. **(Grohe and Mariño 1999)**.
3. FPC captures P on the class of *planar graphs*. **(Grohe 1998)**.
4. FPC captures P on any *proper minor-closed class of graphs*. **(Grohe 2010)**.

In each case, the proof proceeds by showing that for any G in the class, a *canonical, ordered* representaton of G can be interpreted in G using FPC.

Definable Canonization

We say that a class of graphs $\mathcal{C}$ admits *definable canonization* if there is a formula $\eta(\nu_1, \nu_2)$ of FPC with free numeric variables such that for any graph $G \in \mathcal{C}$

$$G \cong ([n], \eta^G)$$

and, if $G, H \in \mathcal{C}$ are isomorphic, then:

$$([n], <, \eta^G) \cong ([n], <, \eta^H).$$

If $\mathcal{C}$ admits definable canonization, then there is a k such that $\equiv^{C^k}$ coincides with isomorphism on $\mathcal{C}$.

Isomorphism on Trees

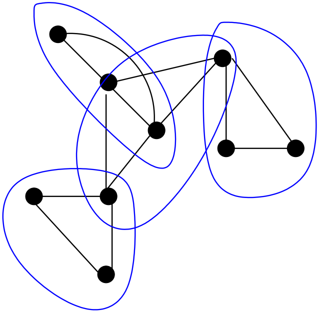
To see that, on *directed trees*, $\equiv^{C^2}$ coincides with isomorphism, note that the following conditions are equivalent

1. Two trees T_u, T_v rooted at u and v respectively are isomorphic.
2. There is a *bijection* h between the children of u in T_u and the children of v in T_v such that for each a, the trees rooted at a and $h(a)$ are isomorphic.

If there is no isomorphism taking u to v, we can use (3) to describe a winning strategy for *Spoiler* in the 2-pebble *bijection game*.

TreeWidth

The *treewidth* of a graph is a measure of how tree-like the graph is. A graph has treewidth k if it can be covered by subgraphs of at most $k + 1$ nodes in a tree-like fashion.



TreeWidth

Formal Definition:

For a graph $G = (V, E)$, a *tree decomposition* of G is a relation $D \subset V \times T$ with a tree T such that:

- for each $v \in V$, the set $\{t \mid (v, t) \in D\}$ forms a connected subtree of T; and
- for each edge $(u, v) \in E$, there is a $t \in T$ such that $(u, t), (v, t) \in D$.

We call $\beta(t) := \{v \mid (v, t) \in D\}$ the *bag* at t.

The *treewidth* of G is the least k such that there is a tree T and a tree-decomposition $D \subset V \times T$ such that for each $t \in T$,

$$|\{v \in V \mid (v, t) \in D\}| \leq k + 1.$$

Isomorphism for Graphs of Bounded Treewidth

The argument showing that on trees, $\equiv^{C^2}$ coincides with isomorphism extends to showing that if

- we expand graphs G and H to G^* and H^* *encoding* a tree-decomposition of width k; *and*
- $G^* \equiv^{C^{2k}} H^*$, then

$G \cong H$.

Unfortunately, tree decompositions are not *unique* and G^* is not determined by G up to isomorphism.

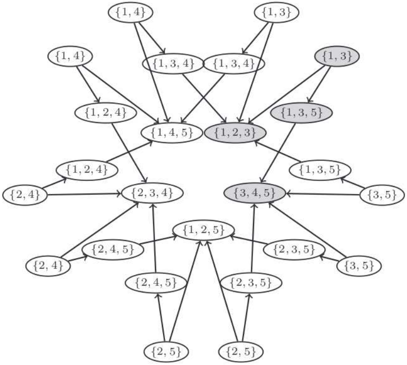
Treelike Decompositions

A *treelike decomposition* of a graph G is a *directed acyclic graph* D, with a *bag* $\beta(d) \subseteq V(G)$ of vertices associated with each node of D and satisfying certain *connectedness* and *consistency* conditions.

A treelike decomposition of G can be obtained (for instance) from a *tree decomposition* by closing it under the *automorphisms* of G—starting at leaves and working upwards.

Treelike Decomposition of a 5-cycle

The picture shows a treelike decomposition of a 5-cycle C_5.
The *grey nodes* form a tree decomposition.



picture credit: M. Grohe: JACM, 59(5), 27.

Definable Treelike Decompositions

Grohe shows that for each k there is an FPC-definable tree-like decomposition of width k on the class of graphs of tree-width at most k.

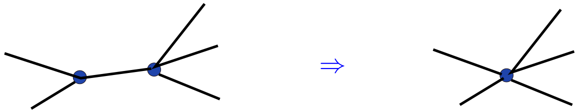
This can be used to establish that $\equiv^{C^{2k}}$ coincides with isomorphism on the class of graphs of treewidth at most k.

A similar result for *planar graphs* is obtained by showing a *definable decomposition* of graphs into their *3-connected* components.

Graph Minors

We say that a graph G is a minor of graph H (written $G \preceq H$) if G can be obtained from H by repeated applications of the operations:

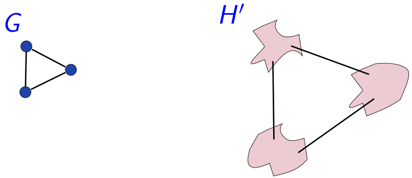
- *delete an edge*;
- *delete a vertex* (and all incident edges); and
- *contract an edge*



Graph Minors

Alternatively, $G = (V, E)$ is a minor of $H = (U, F)$, if there is a graph $H' = (U', F')$ with $U' \subseteq U$ and $F' \subseteq F$ and a surjective map $M : U' \to V$ such that

- for each $v \in V$, $M^{-1}(v)$ is a connected subgraph of H'; and
- for each edge $(u, v) \in E$, there is an edge in F' between some $x \in M^{-1}(u)$ and some $y \in M^{-1}(v)$.



Facts about Graph Minors

- G is planar if, and only if, $K_5 \not\preceq G$ and $K_{3,3} \not\preceq G$.
- If $G \subset H$ then $G \preceq H$.
- The relation $\preceq$ is transitive.
- If $G \preceq H$, then $\mathrm{tw}(G) \leq \mathrm{tw}(H)$.
- If $\mathrm{tw}(G) < k - 1$, then $K_k \not\preceq G$.

Say that a class of graphs $\mathcal{C}$ *excludes* H *as a minor* if $H \not\preceq G$ for all $G \in \mathcal{C}$.
$\mathcal{C}$ has *excluded minors* if it excludes some H as a minor (equivalently, it excludes some K_k as a minor).

- $\mathcal{T}_k$ excludes K_{k+2} as a minor.

More Facts about Graph Minors

Theorem (Robertson-Seymour)
In any infinite collection $\{G_i \mid i \in \omega\}$ of graphs, there are i, j with $G_i \preceq G_j$.

Corollary
For any class $\mathcal{C}$ *closed under minors*, there is a finite collection $\mathcal{F}$ of graphs such that $G \in \mathcal{C}$ *if, and only if,* $F \not\preceq G$ for all $F \in \mathcal{F}$.

The proof relies on Robertson and Seymour's *structure theorem*:

> *A graph G that excludes a minor K_k admits a tree-decomposition in which each bag is almost embeddable in a surface of genus k'*

Isomorphism on Excluded Minor Classes

Grohe lifts the decomposition of planar graphs into *3-connected components* to graphs *embeddable* in an arbitrary surface.

More heavy lifting is required to obtain a *definable treelike decomposition* of the class of graphs *excluding a K_k-minor* into components that can be almost embedded in a surface.

The final result is that for each k, there is a k' such that on graphs excluding K_k as a minor, $\equiv^{C^{k'}}$ coincides with isomorphism.

Cai-Fürer-Immerman Graphs

To show that $\equiv^{C^k}$ does not capture isomorphism everywhere we construct a sequence of pairs of graphs $G_k, H_k (k \in \omega)$ such that:

- $G_k \equiv^{C^k} H_k$ for all k.
- There is a polynomial time decidable class of graphs that includes all G_k and excludes all H_k.

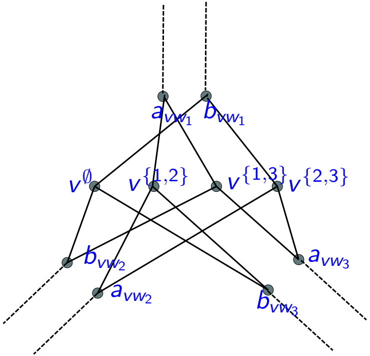
Constructing G_k and H_k

Given any graph G, we can define a graph X_G by replacing every edge with a pair of edges, and every vertex with a gadget.

The picture shows the gadget for a vertex v that is adjacent in G to vertices w_1, w_2 and w_3.

The vertex v^S is adjacent to $a_{vw_i} (i \in S)$ and $b_{vw_i} (i \notin S)$ and there is one vertex for all *even size* S.

The graph $\tilde{X}_G$ is like X_G except that at *one vertex* v, we include v^S for *odd size* S.



Properties

If G is *connected* and has *treewidth* at least k, then:

1. $X_G \not\cong \tilde{X}_G$; and
2. $X_G \equiv^{C^k} \tilde{X}_G$.

(1) allows us to construct a polynomial time property separating X_G and $\tilde{X}_G$.
(2) is proved by a game argument.

> *The original proof of* **(Cai, Fürer, Immerman)** *relied on the existence of balanced separators in* G*. The characterisation in terms of treewidth is from* **(D., Richerby 07)**.

Cops and Robbers

A game played on an undirected graph $G = (V, E)$ between a player controlling k cops and another player in charge of a robber.

At any point, the cops are sitting on a set $X \subseteq V$ of the nodes and the robber on a node $r \in V$.
A move consists in the cop player removing some cops from $X' \subseteq X$ nodes and announcing a new position Y for them. The robber responds by moving along a path from r to some node s such that the path does not go through $X \setminus X'$.
The new position is $(X \setminus X') \cup Y$ and s. If a cop and the robber are on the same node, the robber is caught and the game ends.

Strategies and Decompositions

Theorem (Seymour and Thomas 93):
There is a winning strategy for the *cop player* with k cops on a graph G if, and only if, the tree-width of G is at most $k - 1$.

It is not difficult to construct, from a tree decomposition of width k, a winning strategy for $k + 1$ cops.

Somewhat more involved to show that a winning strategy yields a decomposition.

Cops and Robbers on the Grid

If G is the $k \times k$ toroidal grid, than the *robber* has a winning strategy in the k-*cops and robbers* game played on G.

To show this, we note that for any set X of at most k vertices, the graph $G \setminus X$ contains a connected component with at least half the vertices of G.

If all vertices in X are in distinct rows then $G \setminus X$ is connected. Otherwise, $G \setminus X$ contains an entire row column and in its connected component there are at least $k - 1$ vertices from at least $k/2$ columns.

Robber's strategy is to stay in the large component.

Cops, Robbers and Bijections

We use this to construct a winning strategy for Duplicator in the k-pebble bijection game on X_G and $\tilde{X}_G$.

- A bijection $h : X_G \to \tilde{X}_G$ is *good bar* v if it is an isomorphism everywhere except at the vertices v^S.
- If h is good bar v and there is a path from v to u, then there is a bijection h' that is good bar u such that h and h' differ only at vertices corresponding to the path from v to u.
- Duplicator plays bijections that are good bar v, where v is the *robber position* in G when the cop position is given by the currently pebbled elements.

Bounding Degree and Colour-Class Size

In the construction of **Cai, Fürer and Immerman**, we can choose our graphs G_k, H_k (for which $G_K \equiv^{C^k} H_k$) to have:

- degree bounded by 3;
- colour-class size bounded by 4.

The latter restriction means that we can make them *coloured graphs* in which no more than 4 vertices have the same colour.

It is known that any class of graphs of *bounded degree* admits a polynomial-time isomorphism test. **(Luks 1982)**

It is known that any class of graphs of *bounded colour-class size* admits a polynomial-time isomorphism test. **(Furst, Hopcroft, Luks 1980)**