Tractable Approximations of Graph Isomorphism.

Anuj Dawar

University of Cambridge Computer Laboratory

joint work with Bjarki Holm

Algebra, Logic and Algorithms
Leeds, 12 November 2014

Graph Isomorphism

A *graph* $G = (V, E)$ is a set of vertices V and a set of edges E (i.e. an *irreflexive, symmetric* relation $E \subseteq V \times V$).

A graph is often represented by its *adjacency matrix*. This is a $V \times V$ matrix of 0-1 entries.

Graph Isomorphism: Given graphs G, H, decide whether $G \cong H$.

$G \cong H$ if there is a *bijection* $h : V(G) \to V(H)$ such that $(u, v) \in E(G)$ *if, and only if,* $(h(u), h(v)) \in E(H)$

In other words, can the adjacency matrix of G be *re-ordered* to get that for H?

Complexity of Graph Isomorphism

The *graph isomorphism problem* has an *unusual* status in terms of *computational complexity*

It is:

- not known to be in P;
- in NP:
- not expected to be NP-complete.

In practice and *on average*, graph isomorphism is efficiently decidable.

Tractable Approximations of Isomorphism

A *tractable approximation* of graph isomorphism is a *polynomial-time decidable* equivalence $\equiv$ on graphs such that:

$$G \cong H \quad \Rightarrow \quad G \equiv H.$$

Practical algorithms for testing graph isomorphism typically decide such an approximation.

If this fails to distinguish a pair of graphs G and H, more discriminating tests are deployed.

A complete isomorphism test might consist of a *family* of ever tighter approximations of isomorphism.

Vertex Classification

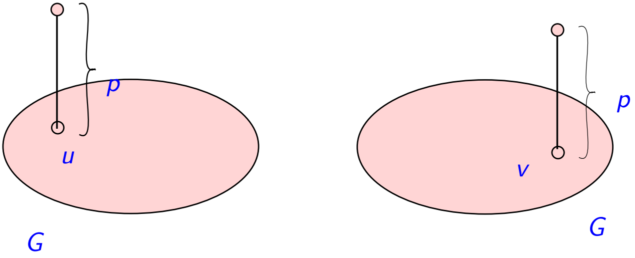
The following problem, which we call the *vertex classification problem* is easily seen to be computationally equivalent to graph isomorphism:

> *Given a graph G and a pair of vertices u and v, decide if there is an automorphism of G that takes u to v.*

That is to say, there is a *polynomial-time reduction* from the graph isomorphism problem to the vertex classification problem and *vice versa*.

Reducing Classification to Isomorphism

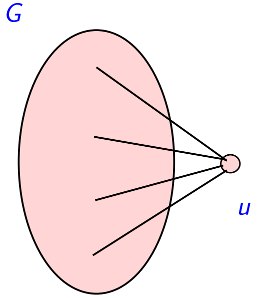
Given a graph G and two vertices $u, v \in V(G)$, we construct a pair of graphs which are isomorphic *if, and only if,* some automorphism of G takes u to v.



where p is a simple path longer than any simple path in G.

Reducing Isomorphism to Classification

Conversely, given two graphs G and H, we construct a graph with two distinguished vertices u, v so that there is an automorphism taking u to v *iff* $G \cong H$.



Equivalence Relations

The algorithms we study decide equivalence relations on *vertices* (or tuples of vertices) that approximate the *orbits* of the automorphism group.

$$(G, u) \cong (G, v) \quad \Rightarrow \quad u \equiv v$$

For such an equivalence relation, there is a *corresponding* equivalence relation on graphs that approximates *isomorphism*.

We abuse notation and use the same notation $\equiv$ for the equivalence relation on vertices, on tuples of vertices and on graphs.

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,

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

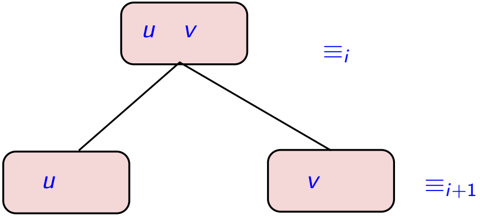
The *naive vertex classification* algorithm finds the *coarsest* equitable partition of the vertices of G.

Colour Refinement

Define, on a graph $G = (V, E)$, a series of equivalence relations:

$$\equiv_0 \supseteq \equiv_1 \supseteq \cdots \supseteq \equiv_i \cdots$$

where $u \equiv_{i+1} v$ if they have the same number of neighbours in each $\equiv_i$-equivalence class.



This converges to the coarsest equitable partition of G: $u \sim v$.

Almost All Graphs

Naive vertex classification provides a simple test for isomorphism that works (in a precise sense) on *almost all graphs*:

> *For graphs G on n vertices with vertices u and v, the probability that $u \sim v$ goes to 0 as $n \to \infty$.*

This also provides an algorithm with good *average case* performance:

> *Check if the input graphs are distinguished by naive vertex classification. In the small number of cases where it fails, try more sophisticated tests.*

But the test fails miserably on *regular graphs* (i.e. graphs where all vertices have the same number of neighbours).

Weisfeiler-Lehman Algorithms

The *k-dimensional Weisfeiler-Lehman* test for isomorphism (as described by **Babai**), generalises naive vertex classification to k-tuples.

We obtain, by successive refinements, an equivalence relation $\equiv^k$ on k-tuples of vertices in a graph G:

$$\equiv^k_0 \supseteq \equiv^k_1 \supseteq \cdots \supseteq \equiv^k_i \quad \cdots$$

$\mathbf{u} \equiv^k_0 \mathbf{v}$ if the two tuples induce isomorphic k-vertex graphs.

The refinement is defined by an *easily checked* condition on tuples.
The refinement is guaranteed to terminate within n^k iterations.

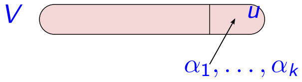
Induced Partitions

Given an equivalence relation $\equiv_i^k$, each k-tuple $\mathbf{u}$ induces a *labelled partition* of the vertices V, where each vertex u is labelled by the k-tuple

$$\alpha_1, \ldots, \alpha_k$$

of $\equiv_i^k$-equivalence classes obtained by substituting u in each of the k positions in $\mathbf{u}$.

Define $\equiv_{i+1}^k$ to be the equivalence relation where $\mathbf{u} \equiv_{i+1}^k \mathbf{v}$ if, in the partitions they induce, the correponding labelled parts *have the same cardinality*.



Families of Tractable Approximations

If G, H are n-vertex graphs and $k < n$, we have:

$$G \cong H \quad \Leftrightarrow \quad G \equiv^n H \quad \Rightarrow \quad G \equiv^{k+1} H \quad \Rightarrow \quad G \equiv^k H.$$

$G \equiv^k H$ is decidable in time $n^{O(k)}$.

The equivalence relations $\equiv^k$ form a *family* of tractable approximations of graph isomorphism.

They have many equivalent characterisations arising from *logic*; *algebra* and *combinatorics*.

Restricted Graph Classes

If we restrict the class of graphs we consider, $\equiv^k$ may coincide with isomorphism.

1. On *trees*, isomorphism is the same as $\equiv^2$. **(Immerman and Lander 1990)**.
2. There is a k such that on the class of *planar graphs* isomorphism is the same as $\equiv^k$. **(Grohe 1998)**.
3. There is a k' such that on the class of graphs of *treewidth* at most k, isomorphism is the same as $\equiv^{k'}$. **(Grohe and Mariño 1999)**.
4. For any *proper minor-closed class of graphs*, $\mathcal{C}$, there is a k such that isomorphism is the same as $\equiv^k$. **(Grohe 2010)**.

These results emerged in the course of establishing *logical characterizations* of polynomial-time computability.

Infinite Hierarchy

(Cai, Fürer, Immerman, 1992) show that there are polynomial-time decidable properties of graphs that are not definable in *fixed-point logic with counting*.

There is no fixed k for which $\equiv^k$ coincides with isomorphism.

(Cai, Fürer, Immerman 1992).

They give a construction of a sequence of pairs of graphs $G_k, H_k (k \in \omega)$ such that for all k:

- $G_k \not\cong H_k$
- $G_k \equiv^k H_k$.

Moreover, G_k, H_k can be chosen to be 3-regular and of colour-class size 4.

Counting Logic

C^k is the logic obtained from *first-order logic* by allowing:

- *counting quantifiers*: $\exists^i x\,\varphi$; and
- only the variables $x_1, \ldots . x_k$.

Every formula of C^k is equivalent to a formula of first-order logic, albeit one with more variables.

We write $G \equiv^{C^k} H$ to denote that no sentence of C^k distinguishes G from H.

It is not difficult to show that $G \equiv^{C^{k+1}} H$ if, and only if, $G \equiv^k H$.

Counting Tuples of Elements

Consider extending the counting logic with quantifiers that count *tuples* of elements.

This does not add further expressive power.

$$\exists^i \overline{xy}\ \varphi$$

is equivalent to

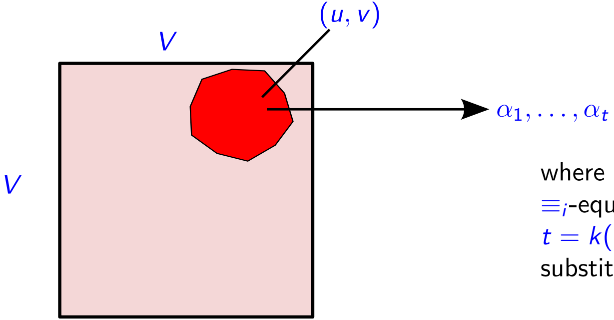
$$\bigvee_{f \in F} \bigwedge_{j \in \mathrm{dom}(f)} \exists^{f(j)} x\ \exists^j y\ \varphi$$

where F is the set of finite partial functions f on $\mathbb{N}$ such that $(\sum_{j \in \mathrm{dom}(f)} jf(j)) = i$.

In other words, in the characterisation of $\equiv^k$ in terms of induced partitions, there is no gain in considering partitions of V^m instead of V.

Induced Partitions of V^2

Can we get a more *refined equivalence* if we use tuples $\mathbf{v}$ to induce partitions of V^2 instead of V?



where $\alpha_1, \ldots, \alpha_t$ are the $\equiv_i$-equivalence classes of the $t = k(k-1)$ ways of substituting (u, v) into $\mathbf{v}$.

No. The *sizes* of these classes are determined by the sizes in the induced partition of V.

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:

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

A *permutation matrix* is a 0-1-matrix which has exactly one 1 in each row and column.

Integer Program

Introducing a variable x_{ij} for each entry of P , the equation $PA = BP$ becomes a system of *linear equations*

$$\sum_k x_{ik} a_{kj} = \sum_k b_{ik} x_{kj}$$

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, $G \sim H$.

(Ramana, Scheiermann, Ullman 1994)

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.

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

$$G \equiv^k H \quad \Rightarrow \quad G \cong^{f,k} H \quad \Rightarrow \quad G \equiv^{k-1} H$$

(Atserias, Maneva 2012)

For $k > 2$, the reverse implications fail. **(Grohe, Otto 2012)**

Rank Logics

The Cai-Fürer-Immerman construction can be reduced to the *solvability of systems of equations* over a 2-element field.

This motivates an extension of first-order logic with operators for the the *rank* of a matrix over a *finite field*.

(D., Grohe, Holm, Laubner, 2009)

For each prime p and each arity m, we have an operator rk_m^p which binds $2m$ variables and defines the rank (over $\mathbb{GF}(p)$) of the $V^m \times V^m$ matrix defined by a formula $\varphi(\mathbf{x}, \mathbf{y})$.

Equivalences and Rank Logic

The definition of rank logics yields a *family* of approximations of isomorphism.

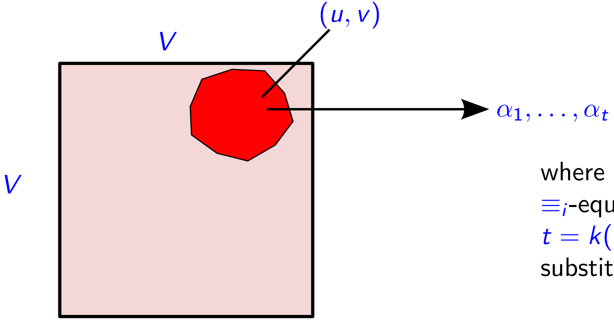
$G \equiv^{R}_{k,\Omega,m} H$ if G and H are not distinguished by any formula of FOrk with at most k variables using operators rk^{p}_{m} for p in the finite set of primes Ω.

We do not know if these relations are *tractable*.
But, we can *refine* them further to obtain a tractable family: $\equiv^{\mathrm{IM}}_{k,\Omega,m}$.
(**D., Holm 2012**)

Induced Partitions

For simplicity, consider the case when $m = 1$ and $\Omega = \{p\}$.

Given an equivalence relation $\equiv$ on V^k, each k-tuple $\mathbf{u}$ induces a *labelled partition* of $V \times V$.



where $\alpha_1, \ldots, \alpha_t$ are the $\equiv_i$-equivalence classes of the $t = k(k-1)$ ways of substituting (u, v) into $\mathbf{u}$.

Induced Partitions

Let $P_1^{\mathbf{u}}, \ldots, P_s^{\mathbf{u}}$ be the parts of this partition (seen as $0\text{-}1$ $V \times V$ matrices) and $P_1^{\mathbf{v}}, \ldots, P_s^{\mathbf{v}}$ be the corresponding parts for a tuple $\mathbf{v}$.

Let $\mathbf{u} \equiv_{i+1} \mathbf{v}$ if, $\mathbf{u} \equiv_i \mathbf{v}$ and for any tuple $\mu \in \{0, \ldots, p-1\}^{[s]}$, we have

$$\mathrm{rk}\Big(\sum_j \mu_i P_j^{\mathbf{u}}\Big) = \mathrm{rk}\Big(\sum_j \mu_j P_j^{\mathbf{v}}\Big).$$

where the rank is in the field $\mathbb{GF}(p)$.

$\equiv^R_{k,\{p\},1}$ is the relation obtained by starting with $\equiv^k_0$ and iteratively repeating this refinement.

For general m and Ω, we need to consider partitions of $V^m \times V^m$ and *rank* and *linear combinations* $\pmod p$ for all $p \in \Omega$.

Complexity of Refinement

We do not know if the relations $\equiv^{R}_{k,\Omega,m}$ are *tractable*

To check $\mathbf{u} \equiv_{i+1} \mathbf{v}$, we have to check the rank of a *potentially exponential* number of linear combinations.

But, we can *refine* the relations further to obtain a tractable family: $\equiv^{\mathrm{IM}}_{k,\Omega,m}$.

(**D., Holm 2012**)

Invertible Map Equivalence

The relation $\equiv^{\mathrm{IM}}_{k,\{p\},1}$ is obtained as the limit of the sequence of equivalence relations where:

> $\mathbf{u} \equiv_{i+1} \mathbf{v}$ *if* $\mathbf{u} \equiv_i \mathbf{v}$ *and there is an invertible matrix* S *(modulo* p*) such that we have for all* j
>
> $$SP^{\mathbf{u}}_j S^{-1} = P^{\mathbf{v}}_j.$$
>
> *This implies, in particular, that all linear combinations have the same rank.*

A result of **(Chistov, Karpinsky, Ivanyov 1997)** guarantees that *simultaneous similarity* of a collection of matrices is decidable in polynomial time. So we get a family of polynomial-time equivalence relations $\equiv^{\mathrm{IM}}_{k,\Omega,m}$.

Could there be a fixed k, m, Ω for which $\equiv^{\mathrm{IM}}_{k,\Omega,m}$ is the same as isomorphism?

Coherent Algebras

Weisfeiler and Lehman presented their algorithm in terms of *cellular algebras*.

These are algebras of matrices on the *complex numbers* defined in terms of *Schur multiplication*:

$$(A \circ B)(i,j) = A(i,j)B(i,j)$$

They are also called *coherent configurations* in the work of **Higman**.

Definition:
A *coherent algebra* with index set V is an algebra $\mathcal{A}$ of $V \times V$ matrices over $\mathbb{C}$ that is:

> *closed under Hermitian adjoints; closed under Schur multiplication; contains the identity I and the all 1's matrix J.*

Coherent Algebras

One can show that a coherent algebra has a *unique basis* $A_1, \ldots, A_m$ of *0-1* matrices which is closed under *adjoints* and such that

$$\sum_i A_i = J.$$

One can also derive *structure constants* p_{ij}^k such that

$$A_i A_j = \sum_k p_{ij}^k A_k.$$

Associate with any graph G, its *coherent invariant*, defined as the smallest coherent algebra $\mathcal{A}_G$ containing the adjacency matrix of G.

Weisfeiler-Lehman method

Say that two graphs G_1 and G_2 are *WL*-equivalent if there is an isomorphism between their *coherent invariants* $\mathcal{A}_{G_1}$ and $\mathcal{A}_{G_2}$.

G_1 and G_2 are *WL*-equivalent if, and only if, $G_1 \equiv^2 G_2$.

Friedland (1989) has shown that two coherent algebras with standard bases $A_1, \ldots, A_m$ and $B_1, \ldots, B_m$ are isomorphic if, and only if, there is an invertible matrix S such that

$$SA_jS^{-1} = B_i \quad \text{for all } 1 \leq j \leq m.$$

Complex Invertible Map Equivalence

Define $\equiv^{\mathrm{IM}}_{\mathbb{C},k}$ as the limit of the sequence of equivalence relations $\equiv_i$ where:

> $\mathbf{u} \equiv_{i+1} \mathbf{v}$ *if there is an invertible linear map* S *on the vector space* $\mathbb{C}^V$ *such that we have for all* i
>
> $$SP^{\mathbf{u}}_j S^{-1} = P^{\mathbf{v}}_j.$$

We can show $\equiv^{\mathrm{IM}}_{\mathbb{C},k+1} \subseteq \equiv^k \subseteq \equiv^{\mathrm{IM}}_{\mathbb{C},k-1}$.

Research Directions

We can show that $\equiv^{\mathrm{IM}}_{4,\{2\},1}$ is the same as isomorphism on graphs of *colour class size 4*.

- For all t, are there fixed k, Ω and m such that $\equiv^{\mathrm{IM}}_{k,\Omega,m}$ is isomorphism on graphs of colour class size t?
- What about graphs of degree at most 3? or t?

Is the *arity hierarchy* really strict on graphs? Could it be that $\equiv^{\mathrm{IM}}_{k,\Omega,m}$ is subsumed by $\equiv^{\mathrm{IM}}_{k',\Omega,1}$ for sufficiently large k'?

Show that no fixed $\equiv^{\mathrm{IM}}_{k,\Omega,m}$ is the same as isomorphism on graphs.

> *Note: we can show that $\equiv^{IM}_{k,\Omega,1}$ is not the same as isomorphism for any fixed k and Ω.*

Summary

The *Weisfeiler-Lehman* family of approximations of graph isomorphism have a number of equivalent characterisations in terms of:

> *complex algebras*; *combinatorics*; *counting logics*; *bijection games*; *linear programming relaxations of isomorphism.*

We have introduced a new and stronger family of approximations of graph isomorphism based on algebras over *finite fields*, and these capture isomorphism on some interesting classes.
There remain many questions about the strength of these approximations and their relations to logical definability.