Morphisms and Isomorphisms in Categories of Relational Structures

Anuj Dawar

Department of Computer Science and Technology, University of Cambridge

Resources and Co-Resources
Cambridge, 19 July 2023

Homomorphisms and Isomorphisms

Two classic algorithmic problems on *finite relational structures*:

Given two structures $\mathbb{A}$ and $\mathbb{B}$, determine if there is a homomorphism $\mathbb{A} \longrightarrow \mathbb{B}$.

- *Problem is NP-complete.*
- *Much studied in the form of the constraint satisfaction problem*
- *For any fixed $\mathbb{B}$, $\mathrm{CSP}(\mathbb{B}) = \{\mathbb{A} \mid \mathbb{A} \longrightarrow \mathbb{B}\}$ is either in P or NP-complete* **(Bulatov; Zhuk 2017)**

Isomorphism

Given two structures $\mathbb{A}$ and $\mathbb{B}$, determine if there is a isomorphism $\mathbb{A} \cong \mathbb{B}$.

- *Not known to be in P or NP-complete*
- *Much studied in the form of the equivalent graph isomorphism problem.*
- *Solvable in quasi-polymomial time* **(Babai 2015)**

Tractable Approximations

For both *homomorphism* and *isomorphism* problems there has been an extensive study of *islands of tractability*.

This can take a number of forms including:

1. Identifying classes of structures $\mathcal{C}$ such that the problem is tractable when restricted to $\mathcal{C}$.
2. Giving efficient algorithms which determine relatons that are *appoximate* versions of the homomophism and isomorphism relations.

These approaches are often related in the sense that the approximation algorithm may be *exact* on a restricted class of structures.

Combinatorial Algorithms

The efficient algorithms that solve the problems in restricted classes are broadly classified in two sorts: *combinatorial* and *algebraic*.

The k-dimensional *Weisfeiler-Leman* test gives an *approximation* of isomorphism:

- $\mathbb{A} \equiv^k \mathbb{B}$ if, and only if, $\mathbb{A}$ and $\mathbb{B}$ are not distinguished in the k-variable infinitary logic with counting ($\mathcal{C}^k_{\infty\omega}$).
- If $\mathbb{A}$ and $\mathbb{B}$ have *treewidth* less than k, then $\mathbb{A} \equiv^k \mathbb{B}$ implies $\mathbb{A} \cong \mathbb{B}$.
- More generally, for any class of graphs *excluding some minor*, there is a k such that $\equiv^k$ is isomorphism **(Grohe)**.

Local Consistency

The k-local consistency algorithm (for $k \in \omega$) gives an *approximation* of homomorphism : $\mathbb{A} \Rrightarrow^k \mathbb{B}$

- $\mathbb{A} \Rrightarrow^k \mathbb{B}$ iff every k-variable formula of *existential positive k-variable infinitary logic* that is true in $\mathbb{A}$ is true in $\mathbb{B}$.
- $\mathrm{CSP}(\mathbb{B})$ has *bounded width* (is solvable) by a Datalog program iff for some k:

$$\mathbb{A} \longrightarrow \mathbb{B} \quad \Leftrightarrow \quad \mathbb{A} \Rrightarrow^k \mathbb{B}.$$

- If $\mathbb{A}$ has *treewidth* less than k then $\mathbb{A} \Rrightarrow^k \mathbb{B}$ implies $\mathbb{A} \longrightarrow \mathbb{B}$.

Pebbling Comonad

The *graded pebbling comonad* $\mathbb{P}_k (k \in \omega)$ and its associated Kleisli category establish a formal connection between *local consistency* and the *Weisfeiler-Leman* algorithms.

$\mathbb{A} \xrightarrow{\mathcal{K}(\mathbb{P}_k)} \mathbb{B}$ if, and only if, $\mathbb{A} \Rrightarrow^k \mathbb{B}$.

$\mathbb{A} \overset{\mathcal{K}(\mathbb{P}_k)}{\cong} \mathbb{B}$ if, and only if, $\mathbb{A} \equiv^k \mathbb{B}$.

Moreover, equivalence in the infinitary logic $\mathcal{L}^k_{\infty\omega}$ *(without counting)* can be obtained as *bisimulation* in $\mathcal{EM}(\mathbb{P}_k)$.
(Abramsky-Shah; Abramsky-Reggio)

The *coalgebras* of the comonad also explain the key role of *treewidth*.

From Me <anuj.dawar@cl.cam.ac.uk>
To Samson Abramsky <samson.abramsky@cs.ox.ac.uk> 12/12/2016, 02:04
Bcc Me <anuj.dawar@cl.cam.ac.uk>
Subject **Re: strategies** OpenPGP

Hi Samson,

> When were discussing your notion of strategies, I asked if one could view a strategy from structure A to structure B as a homomorphism from A to some structure T(B). You indicated that one could, where T(B) was an unfolding of B.
> Could you spell this out a bit more?

I may have misunderstood the question when you asked it. What I believe is that we can view a strategy from A to B as a homomorphism from T(A) to B, where T(A) is an unfolding of A.

Let me define the structure T(A) more formally. For simplicity, assume the structures are graphs (i.e. they have one binary, symmetric relation E). The elements of T(A) are sequences s1...sn (for any natural number n) of pairs (i,a) where i is a number in [1,...,k] and a is an element of A. Informally, we think of the sequence s1...sn as representing a play of n moves by Spoiler where the pair (i,a) represents a move of pebble i to position a.

The edge relation on T(A) is defined as follows: We have an edge between s = s1...sm and s' = s1...sn if the following conditions hold:
1. s is a prefix of s'
2. if sm = (i,a) then i does not occur as the first component of any sl for l > m
3. sn = (j,b) and (a,b) is an edge in A.

The idea is that at the stage of the game represented by the sequence s', Spoiler has just placed pebble j on b, and any response by Duplicator must respect all edges to positions a where there is already a pebble. These are represented by the pairs (i,a) which are the last occurrence of i in the sequence s'.

Now, we can show that a winning strategy for Duplicator in the k-pebble game between A and B gives a homomorphism from T(A) to B, and conversely.

I hope this is clear. Please let me know if not.

Best,
-Anuj.

Lovász-like Theorems

Theorem (Lovász)
$\mathbb{A} \cong \mathbb{B}$ if, and only if, $\forall \mathbb{C}\, |\hom(\mathbb{C}, \mathbb{A})| = |\hom(\mathbb{C}, \mathbb{B})|$

Theorem (Dvořak)
$\mathbb{A} \equiv^k \mathbb{B}$ if, and only if,
$\forall \mathbb{C}\, \mathsf{tw}(\mathbb{C}) < k \Rightarrow |\hom(\mathbb{C}, \mathbb{A})| = |\hom(\mathbb{C}, \mathbb{B})|$

Theorem (Grohe)
$\mathbb{A} \equiv_m \mathbb{B}$ if, and only if,
$\forall \mathbb{C}\, \mathsf{td}(\mathbb{C}) < m \Rightarrow |\hom(\mathbb{C}, \mathbb{A})| = |\hom(\mathbb{C}, \mathbb{B})|$

These are all instances of the same theorem in different categories:
$\mathcal{R}(\sigma)$, $\mathcal{EM}(\mathbb{P}_k(\mathcal{R}(\sigma)))$, $\mathcal{EM}(\mathbb{E}_k(\mathcal{R}(\sigma)))$.

Abstract Lovász Theorem

Definition.
A category $\mathcal{C}$ is *combinatorial* if for all $A, B \in \mathcal{C}$ we have

$$A \cong_{\mathcal{C}} B \quad \text{iff} \quad \forall C \in \mathcal{C} : |\operatorname{hom}(C, A)| = |\operatorname{hom}(C, B)|$$

Theorem (D,Jakl,Reggio)
Any locally finite category with pushouts and a proper factorization system is combinatorial.

Algebraic Algorithms

There is no systematic picture linking *algebraic* approaches to *homomorphism* and *isomorphism*.

Algebraic approaches for *isomorphism* are usually based on *permutation group methods*.

e.g. **Luks'** *algorithm for bounded degree graphs.*

Algebraic analysis of *CSP* is based on *clones* and has yielded algorithms generalizing linear algebra.

Generalized Gaussian elimination

Zhuk's *algorithm repeatedly solves systems of equations over (direct products of) finite fields.*

Note: It is not clear that we can extract from Zhuk's algorithm a general *approximation* of homomorphism in the way that k-local consistency yields $\Rrightarrow^k$.

Invertible Map Equivalences

The *invertible map equivalences* (**D, Holm**) provide an approximation of *isomorphism* that incorporates tests based on linear algebra over *finite fields*.

$\mathbb{A} \equiv^k_{\mathsf{IM}} \mathbb{B}$ if $\mathbb{A}$ and $\mathbb{B}$ are not distinguished by any sentence of LA^k—the extension of infinitary logic with k variables with all *linear algebraic Lindström quantifiers*.

We have a comonad that gives an account of infinitary logic with *all* n-ary Lindström quantifiers. **(D, Ó Conghaile)**.

Linear Algebraic Quantifiers

A quantifier is *linear algebraic* if it is invariant under invertible linear maps over some prime field.

Do these approximations of isomorphism also induce a corresponding approximation of *homomorphism*?

The result of **Lichter, Pago, Seppelt** tells us that this can not be achieved by a comonadic construction like for local consistency:

> *There is no class of graphs $\mathcal{F}$ such that $\equiv^k_{IM}$ is characterized by homomorphism counts from $\mathcal{F}$.*

Preservation Theorems

Theorem (Łoś, Lyndon, Tarski)
Every first-order sentence which is *preserved along homomorphisms* is equivalent to an *existential positive* sentence.

Theorem (Rossman)
Every first-order sentence of quantifier rank m which is *preserved along homomorphisms* is equivalent to an *existential positive* sentence of quantifer rank m

Theorem (Rossman)
Every first-order sentence φ whose *finite* models are closed under homomorphisms is equivalent *on finite structures* to an existential positive sentence.

Note: A finite equirank version is unknown.

Comonadic Formulation

Recall that $(\mathcal{E}_m)_{m \in \omega}$ is the *graded EF comonad*.

Equirank Homomorphism Preservation Theorem (comonadic formulation)
If $\mathcal{C}$ is a *morphism-closed* subcategory of $\mathcal{R}(\sigma)$, and $\mathbb{E}_m(\mathcal{C})$ is *bisimulation-closed*, then $\mathbb{E}_m(\mathcal{C})$ is morphism-closed (in $\mathcal{K}(\mathbb{E}_m(\mathcal{R}(\sigma)))$).

Note: A more general formulation is proved by **Abramsky, Reggio**.

Finite Homomorphism Preservation

Finite Homomorphism Preservation Theorem (comonadic formulation)
If $\mathcal{C}$ is a *morphism-closed* subcategory of $\mathcal{R}_{\mathsf{fin}}(\sigma)$, and for some m $\mathbb{E}_m(\mathcal{C})$ is *bisimulation-closed*, then there is an m' such that $\mathbb{E}_{m'}(\mathcal{C})$ is morphism-closed (in $\mathcal{K}(\mathbb{E}_{m'}(\mathcal{R}_{\mathsf{fin}}(\sigma)))$).

An abstract proof remains a challenge. It would have to make crucial use of the *grading*.

Since we can remove counting quantifiers from a first-order formula at the expense of an increase in *quantifier rank*, we can also formulate it in terms of *isomorphism*

Finite Homomorphism Preservation Theorem (comonadic isomorphism formulation)
If $\mathcal{C}$ is a *morphism-closed* subcategory of $\mathcal{R}_{\mathsf{fin}}(\sigma)$, and for some m $\mathbb{E}_m(\mathcal{C})$ is *isomorphism-closed*, then there is an m' such that $\mathbb{E}_{m'}(\mathcal{C})$ is morphism-closed (in $\mathcal{K}(\mathbb{E}_{m'}(\mathcal{R}_{\mathsf{fin}}(\sigma)))$).

Preservation in the Pebbling Comonad

We can ask whether similar statements hold of the *pebbling comonad*.

Homomorphism Preservation Theorem for Infinitary Logic
If $\mathcal{C}$ is a *morphism-closed* subcategory of $\mathcal{R}_{\mathsf{fin}}(\sigma)$, and for some k $\mathbb{P}_k(\mathcal{C})$ is *bisimulation-closed*, then there is a k' such that $\mathbb{P}_{k'}(\mathcal{C})$ is morphism-closed (in $\mathcal{K}(\mathbb{P}_{k'}(\mathcal{R}_{\mathsf{fin}}(\sigma)))$).

Homomorphism Preservation Theorem for Counting Infinitary Logic
If $\mathcal{C}$ is a *morphism-closed* subcategory of $\mathcal{R}_{\mathsf{fin}}(\sigma)$, and for some k $\mathbb{P}_k(\mathcal{C})$ is *isomorphism-closed* (in $\mathcal{K}(\mathbb{P}_k(\mathcal{R}_{\mathsf{fin}}(\sigma)))$), then there is a k' such that $\mathbb{P}_{k'}(\mathcal{C})$ is morphism-closed (in $\mathcal{K}(\mathbb{P}_{k'}(\mathcal{R}_{\mathsf{fin}}(\sigma)))$).

The statements are not *equivalent*. The second is *stronger*. Both are open questions.

Infinitary Preservation

Question (logical formulation)

Every sentence of $\mathcal{L}^{\omega}_{\infty\omega}$ whose finite models are closed under homomorphisms is equivalent, over finite structures, to an existential positive sentence of $\mathcal{L}^{\omega}_{\infty\omega}$.

Theorem (Feder-Vardi 2003)

Every *existential* sentence of $\mathcal{L}^{\omega}_{\infty\omega}$ whose finite models are closed under homomorphisms is equivalent, over finite structures, to an existential positive sentence of $\mathcal{L}^{\omega}_{\infty\omega}$.

Extending this to a logic with *universal quantifiers* is a real challenge.

They also show:

Theorem (Feder-Vardi 2003)

Every Datalog program with *negation* whose finite models are closed under homomorphisms is equivalent, over finite structures to a Datalog program.

And this, we know, does not extend to the case with *universal*

Infinitary Preservation with Counting

Question
Every sentence of $\mathcal{C}^{\omega}_{\infty\omega}$ whose finite models are closed under homomorphisms is equivalent, over finite structures, to an *existential positive* sentence of $\mathcal{L}^{\omega}_{\infty\omega}$.

We know this is true if the sentence defines a class of the form $\{\mathbb{A} \mid \mathbb{A} \not\longrightarrow \mathbb{B}\}$ for some fixed $\mathbb{B}$.

This follows from results **(Atserias, Bulatov, D.)** and **(Barto, Kozik)**.

Conclusion

It was a great project, and there is so much still to do.