Preservation theorems on sparse classes revisited

Anuj Dawar[1] and Ioannis Eleftheriadis[1*]

University of Cambridge, Cambridge, United Kingdom
anuj.dawar@cl.cam.ac.uk
ie257@cam.ac.uk

Abstract

We revisit the work studying homomorphism preservation in sparse classes of structures initiated in [Atserias et al., JACM 2006] and [Dawar, JCSS 2010]. These established that first-order logic has the homomorphism preservation property in any sparse class that is monotone and addable. It turns out that the assumption of addability is not strong enough for the proofs given. We demonstrate this by constructing classes of graphs of bounded treewidth which are monotone and addable but fail to have homomorphism preservation. We also show that homomorphism preservation fails on the class of planar graphs. On the other hand, the proofs can be recovered by replacing addability by a stronger condition of amalgamation over bottlenecks. This is analogous to a similar condition formulated for extension preservation in [Atserias et al., SiCOMP 2008].

1 Introduction

Preservation theorems have played an important role in the development of finite model theory. They provide a correspondence between the syntactic structure of first-order sentences and their semantic behaviour. In the early development of finite model theory it was noted that many classical preservation theorems fail when we limit ourselves to finite structures. An important case in point is the Łoś-Tarski or *extension* preservation theorem, which asserts that a first-order formula is preserved by embeddings between all structures if, and only if, it is equivalent to an existential formula. Interestingly, this was shown to fail on finite structures [9] much before the question attracted interest in finite model theory [6]. On the other hand, the *homomorphism* preservation theorem, asserting that formulas preserved by homomorphisms are precisely those equivalent to existential-positive ones, was remarkably shown to hold on finite structures by Rossman [8], spurring applications in constraint satisfaction and database theory.

However, even before Rossman's result, these preservation properties were investigated on subclasses of the class of finite structures. In this context, restricting to a subclass weakens both the hypothesis and the conclusion, therefore leading to an entirely new question. Thus, while the class of all finite structures is combinatorially wild, it contains *tame* classes which are both algorithmically and model-theoretically better behaved [4]. A study of preservation properties for such restricted classes of finite structures was initiated in [3] and [2], which looked at homomorphism preservation and extension preservation respectively. The focus was on tame classes defined by *wideness* conditions, allowing for methods based on the *locality* of first-order logic.

The main result asserted in [3] is that homomorphism preservation holds in any class $\mathcal{C}$ which is *almost wide* and is *monotone* and *addable*. From this, it is concluded that homomorphism preservation holds for any class $\mathcal{C}$ whose Gaifman graphs exclude some graph G as a minor, as long as $\mathcal{C}$ is monotone and addable. The result was extended from almost wide to *quasi-wide* classes in [5], from which homomorphism preservation was deduced for classes that locally

*Supported by a George and Marrie Vergotis Scholarship awarded by Cambridge Trust, an Onassis Foundation Scholarship, and a Robert Sansom Studentship

exclude minors and classes that have bounded expansion, again subject to the proviso that they are monotone and addable. Quasi-wide classes were later identified with *nowhere dense* classes, which are now central in structural and algorithmic graph theory [7].

The main technical construction in [3] is concerned with showing that classes of graphs which exclude a minor are indeed almost wide. The fact that homomorphism preservation holds in monotone and addable almost wide classes is deduced from a construction of Ajtai and Gurevich [1] which shows the "density" of minimal models of a first-order sentence preserved under homomorphisms, and the fact that in an almost wide class a collection of such dense models must necessarily be finite. While the Ajtai and Gurevich construction is carried out within the class of all finite structures, it is argued in [3] that it can be carried out in any monotone and addable class because of "the fact that disjoint union and taking a substructure are the only constructions used in the proof" [3, p. 216].

The starting point of the present paper is that this argument is flawed. The construction requires us to take not just disjoint unions, but unions that identify certain elements: in other words *amalgamations* over sets of points. On the other hand, we can relax the requirement of monotonicity to just hereditariness. The conclusion is that homomorphism preservation holds in any class $\mathcal{C}$ that is quasi-wide, hereditary and closed under amalgamation over bottleneck points. The precise statement is given in Theorem 4.1 below. We also show that the requirements formulated in [3] are insufficient by constructing a class that is almost wide (indeed, has bounded treewidth), is monotone and addable, but fails to have the homomorphism preservation property. The class of planar graphs is an interesting case as it is used in [2] as an example of a hereditary, addable class with excluded minors in which extension preservation fails. We show that homomorphism preservation also fails in this class, strengthening the result of [2].

2 Preliminaries

We fix a finite relational vocabulary τ; by a structure we implicitly mean a τ-structure. Given two structures A, B, a homomorphism $f : A \to B$ is a map such that for all relation symbols R and tuples $\bar{a}$ from A we have $\bar{a} \in R^A \implies f(\bar{a}) \in R^B$. If moreover $f(\bar{a}) \in R^B \implies \bar{a} \in R^A$ then f is said to be *strong*. An injective strong homomorphism is called an *embedding*.

A structure B is said to be a *weak substructure* of a structure A if $B \subseteq A$ and the inclusion map $\iota : B \hookrightarrow A$ is a homomorphism. Likewise, B is an *induced substructure* of A if the inclusion map is an embedding. An induced substructure B of A is said to be *free in A* if there is some structure C such that A is the disjoint union $B + C$. Finally, a substructure B of A is said to be *proper* if the inclusion map is not full. We say that a class of structures is *monotone* if it is closed under weak substructures, and it is *hereditary* if it is closed under induced substructures. Moreover a class is called *addable* if it is closed under taking disjoint unions.

Given structures A, B, S and embeddings $f : S \to A$ and $g : S \to B$, we write $A \oplus_{S,f,g} B$ for the quotient of the disjoint union $A + B$ by the equivalence relation generated by $\{(f(s), g(s)) : s \in S\}$. Whenever $S \subseteq A \cap B$, we write $A \oplus_S B$ for $A \oplus_{S,\iota_A,\iota_B} B$ where ι_A, ι_B are the corresponding inclusion maps, and call this the *free amalgam of A and B over S*.

Fixing a graph H, we say that a graph G is H-free and H-minor-free if it does not contain H as an induced subgraph and minor respectively. By Wagner's Theorem, a graph is planar if and only if it is K_5-minor-free and $K_{3,3}$-minor-free. Finally, a class of graphs $\mathcal{C}$ is said to be *quasi-wide* if for every $r \in \mathbb{N}$ there exist $s_r \in \mathbb{N}$ and $f_r : \mathbb{N} \to \mathbb{N}$ such that for every $m \in \mathbb{N}$ and every $G \in \mathcal{C}$ there exist disjoint sets $A, S \subseteq V(G)$ such that A is r-independent in $G \setminus S$.

We say that a formula ϕ is preserved by homomorphisms (respectively extensions) over a class of structures $\mathcal{C}$ if for all $A, B \in \mathcal{C}$ such that there is a homomorphism (respectively embedding) from A to B, $A \models \phi$ implies that $B \models \phi$. We say that a class of structures $\mathcal{C}$ has the *homomorphism preservation property* (HPP) (respectively *extension preservation property*, EPP) if for every formula ϕ preserved by homomorphisms (respectively extensions) over $\mathcal{C}$ there is an existential-positive (respectively existential) formula ψ such that $M \models \phi \iff M \models \psi$ for all $M \in \mathcal{C}$.

Given a formula ϕ and a class of structures $\mathcal{C}$, we say that $M \in \mathcal{C}$ is a *minimal induced model* of ϕ in $\mathcal{C}$ if $M \models \phi$ and for any proper induced substructure N of M with $N \in \mathcal{C}$ we have $N \not\models \phi$. The relationship between minimal models and preservation is highlighted by the following theorem.

Theorem 2.1. *Let $\mathcal{C}$ be a hereditary class of finite structures. The $\mathcal{C}$ has the HPP (respectively EPP) if and only if every formula preserved by homomorphisms (respectively extensions) over $\mathcal{C}$ has finitely many minimal induced models in $\mathcal{C}$. So, if $\mathcal{C}$ has the EPP then it has the HPP.*

3 Preservation can fail on classes of small treewidth

Theorem 4.4 of [3] can be paraphrased in the language of this paper as saying that *homomorphism preservation holds over any monotone and addable class of bounded treewidth.* Here, we provide a simple counterexample to this, exhibiting a monotone and addable class of graphs of treewidth 3 where homomorphism preservation fails.

Definition 3.1. Fix $k \in \mathbb{N}$ and $n_i \geq 3$ for every $i \in [k]$. We define the *bouquet of cycles of type* $(n_1, \dots, n_k)$, denoted by $W_{n_1,\dots,n_k}$, as the graph obtained by taking the disjoint union of k cycles of length $n_1, \dots, n_k$ respectively, and adding an apex vertex, i.e. a vertex adjacent to every vertex in these cycles. Whenever $k = 1$, we refer to the graph W_n as the *wheel of order n*.

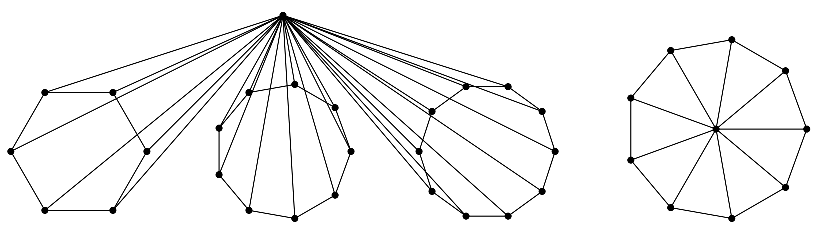

Figure 1: The bouquet of cycles of type $(6, 9, 10)$ and the wheel of order 9 respectively.

First, observe that each bouquet has treewidth 3. Indeed, taking a tree decomposition of each cycle of width 2, and adding the apex to every bag in the decomposition gives the required tree decomposition. The advantage of working with bouquets of cycles is that, unlike single cycles, there is a formula that defines their existence as free induced subgraphs. To see this, we let

$$\psi(x, z) := \exists u \exists v [u \neq v \land u \neq x \land v \neq x \land E(z, u) \land E(z, v) \land \forall w (E(w, z) \to w = u \lor w = v \lor w = x)],$$

$$\text{and } \phi := \exists x \exists y [E(x, y) \land \forall z (z \neq x \land \text{dist}(x, z) \leq 2 \to E(x, z) \land \psi(x, z)].$$

Intuitively, ϕ asserts the following: "there is a vertex x of degree at least one such that every other vertex reachable from x by a path of length two is adjacent to x and has exactly two distinct neighbours which are not x".

Lemma 3.2. *Let G be an arbitrary finite graph. Then $G \vDash \phi$ if, and only if, it contains a bouquet of cycles as a free induced subgraph.*

It is evident that ϕ is not preserved by homomorphisms over the class of all undirected graphs. However, when restricting to subgraphs of disjoint unions of wheels we no longer have non-free-occurring bouquets of cycles in the class. This is precisely the core of the following theorem.

Theorem 3.3. *The monotone and addable closure of $\{W_{2n+1} : n \in \mathbb{N}\}$ does not have the HPP.*

4 Preservation under bottleneck amalgamation

The main result of this section is the corrected version of Theorem 4.4 in [3] and its generalisation, Theorem 9 in [5]. More precisely, we establish homomorphism preservation on hereditary quasi-wide classes which are closed under certain free amalgams. While the existence of arbitrary amalgams certainly suffices, it prohibits any sort of sparsity in the class. Indeed, any hereditary class of undirected graphs with the free amalgamation property contains arbitrarily large 1-subdivided cliques, and hence, cannot be quasi-wide.

The proof proceeds by obtaining a concrete bound on the size of minimal models of ϕ in $\mathcal{C}$, and concluding by Theorem 2.1. The existence of this bound is guaranteed by quasi-wideness, as any large enough structure contains a large scattered set after removing a small number of bottleneck points. To isolate the bottleneck points $\bar{p}$ of M we consider a structure $\bar{p}M$ in an expanded language which is bi-interpretable with M, and work with the corresponding interpretation ϕ^k of ϕ; in particular $\bar{p}M$ contains a large scattered set itself and it models ϕ^k. Then, by removing a carefully chosen point from the scattered set of $\bar{p}M$, we obtain a proper induced substructure $\bar{p}N$ of $\bar{p}M$ such that $N \in \mathcal{C}$ by hereditariness. To argue that this still models ϕ^k, we use a relativisation of the locality argument of Ajtai and Gurevich from [1]. While in its original version the argument only considers disjoint copies of M, working with the interpretation $\bar{p}M$ of M corresponds to taking free amalgams of M over the set of bottleneck points; this is precisely the subtlety that was missed in [3] and [5].

Theorem 4.1. *Let $\mathcal{C}$ be a hereditary class such that for every $r \in \mathbb{N}$ there exist $k_r \in \mathbb{N}$ and $f_r : \mathbb{N} \to \mathbb{N}$ satisfying that for every $m \in \mathbb{N}$ and $M \in \mathcal{C}$ of size at least $f_r(m)$ there exist disjoint sets $A, S \subseteq M$ such that $|A| \geq m$, $|S| \leq k_r$, A is r-independent in $M \setminus S$, and $\oplus^n_S M \in \mathcal{C}$ for every $n \in \mathbb{N}$. Then homomorphism preservation holds over $\mathcal{C}$.*

Obtaining homomorphism preservation for quasi-wide classes therefore amounts to verifying closure under amalgams over bottleneck points. This is precisely the case for K_4-minor-free and outerplanar graphs. Another class with this property is already known to exist by [2], that is, the class $\mathcal{T}_k$ of all graphs of treewidth bounded by k, for any $k \in \mathbb{N}$.

Theorem 4.2. *The classes of K_4-minor-free graphs and outerplanar graphs have the HPP.*

5 Preservation fails on planar graphs

In this section we witness that homomorphism preservation fails on the class of planar graphs. Previously, it was established [2] that the extension preservation property fails on planar graphs. Since extension preservation implies homomorphism preservation on hereditary classes by Theorem 2.1, our result strengthens the above. Our construction will in fact also reveal that homomorphism preservation fails on the class of K_5-minor-free graphs.

Definition 5.1. Fix $n \in \mathbb{N}$. Define D_n as the undirected graph on vertex set

$$V(D_n) = \{v_1, v_2\} \cup \{a_i : i \in [n]\} \cup \{b_i : i \in [n]\}, \text{ and edge set}$$

$$E(D_n) = \{(v_1, a_i) : i \in [n]\} \cup \{(v_2, b_i) : i \in [n]\} \cup \{(a_i, b_i) : i \in [n]\} \cup \{(a_i, a_{i+1}) : i \in [n-1]\}$$
$$\cup \{(b_i, b_{i+1}) : i \in [n-1]\} \cup \{(a_{i+1}, b_i) : i \in [n-1]\} \cup \{(a_1, a_n), (b_1, b_n), (a_1, b_n)\}.$$

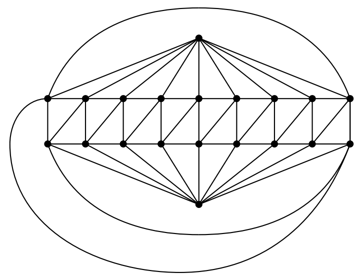

Figure 2: A planar embedding of D_9.

We proceed to characterise the K_5-minor-free homomorphic images of D_n.

Theorem 5.2. *Fix $n \geq 4$. Then any K_4-free and K_5-minor-free homomorphic image of D_n contains an induced copy of D_m for some $m \geq 4$ such that $m \mid n$.*

We then show that the existence of the graphs D_n as induced subgraphs is definable among K_4-free K_5-minor-free graphs by a simple first-order formula. Indeed, consider the formula

$$\chi(x_1, x_2, y_1, z_1, y_2, z_2) = E(x_1, y_2) \wedge E(y_1, y_2) \wedge E(z_1, y_2) \wedge E(z_1, z_2) \wedge E(y_2, z_2) \wedge E(z_2, x_2),$$

$$\text{and } \phi = \exists x_1, x_2, y, z[E(x_1, y) \wedge E(y, z) \wedge E(z, x_2) \wedge \forall a, b(E(x_1, a) \wedge E(a, b) \wedge E(b, x_2))$$
$$\rightarrow \exists c, d\ \chi(x_1, x_2, a, b, c, d))]$$

Proposition 5.3. *Let H be a finite K_4-free and K_5-minor free graph. Then $H \models \phi$ if and only if, there is some $n \geq 4$ such that H contains D_n as an induced subgraph.*

Putting the above together, we deduce the main theorem of this section.

Theorem 5.4. *The class of planar graphs does not have the HPP.*

Proof. Let $\hat{\phi}$ be the disjunction of ϕ with the formula that induces a copy of K_4, i.e. $\hat{\phi} := \phi \vee \exists x_1, x_2, x_3, x_4 \bigwedge_{i \neq j} E(x_i, x_j)$. We argue that $\hat{\phi}$ is preserved by homomorphisms over the class of planar graphs. Indeed, let $f : G \to H$ be a homomorphism with G, H planar such that $G \models \hat{\phi}$. Clearly, if H contains a copy of K_4 then $H \models \hat{\phi}$. Without loss of generality we may assume that $G \models \phi$ and G, H are K_4-free. It follows by Proposition 5.3 that there exists some $n \geq 4$ such that G contains D_n as a subgraph. Theorem 5.2 thus implies that H that there is some $m \geq 4$ such that H contains D_m as a subgraph. Proposition 5.3 then implies that that $H \models \phi$, and thus $H \models \hat{\phi}$ as required. To conclude, observe that the minimal models of $\hat{\phi}$ over the class of planar graphs are K_4 and the graphs D_n for $n \geq 4$; since these are infinitely many Theorem 2.1 implies that $\hat{\phi}$ is not equivalent to an existential-positive formula over the class of planar graphs. □

Since we only use exclusion of K_5-minors, the same proof relativises to the following theorem.

Theorem 5.5. *The class of all K_5-minor-free graphs does not have the HPP.*

References

[1] Miklos Ajtai and Yuri Gurevich. Datalog vs first-order logic. *Journal of Computer and System Sciences*, 49(3):562–588, 1994. 30th IEEE Conference on Foundations of Computer Science.

[2] Albert Atserias, Anuj Dawar, and Martin Grohe. Preservation under extensions on well-behaved finite structures. *SIAM Journal on Computing*, 38(4):1364–1381, 2008.

[3] Albert Atserias, Anuj Dawar, and Phokion G Kolaitis. On preservation under homomorphisms and unions of conjunctive queries. *Journal of the ACM (JACM)*, 53(2):208–237, 2006.

[4] Anuj Dawar. Finite model theory on tame classes of structures. In *International Symposium on Mathematical Foundations of Computer Science*, pages 2–12. Springer, 2007.

[5] Anuj Dawar. Homomorphism preservation on quasi-wide classes. *Journal of Computer and System Sciences*, 76(5):324–332, 2010.

[6] Yuri Gurevich. Toward logic tailored for computational complexity. *Computation and Proof Theory*, pages 175–216, 1984.

[7] Jaroslav Nešetřil and Patrice Ossona De Mendez. *Sparsity: graphs, structures, and algorithms*, volume 28. Springer Science & Business Media, 2012.

[8] Benjamin Rossman. Homomorphism preservation theorems. *Journal of the ACM (JACM)*, 55(3):1–53, 2008.

[9] W. W. Tait. A counterexample to a conjecture of Scott and Suppes. *Journal of Symbolic Logic*, 24(1):15–16, 1959.