Preservation Theorems in Finite Model Theory

Anuj Dawar
University of Cambridge

Panhellenic Logic Symposium, Volos, 7 July 2007

Finite Model Theory – Early Days

In the 1980s, the term *finite model theory* came to be used to describe the study of the expressive power of logics (from first-order to second-order logic and in between), on the class of all finite structures.

The motivation for the study is that problems in computer science (especially in *complexity theory* and *database theory*) are naturally expressed as questions about the expressive power of logics.

And, the structures involved in computation are finite.

Finite Model Theory – Early Trends

Kolaitis (LICS 93) identified trends in the results in finite model theory.

- *Negative*: showing the failure of classical model-theoretic results on finite structures.

  Compactness. Completeness. Interpolation and preservation theorems.

- *Conservative*: showing that certain classical model-theoretic results continue to hold on finite structures.

  Some consequences of compactness. Monotone vs. positive inductions. Locality.

- *Positive*: exploring concepts and results which are specific to finite structures.

  Descriptive complexity. 0–1 laws.

Preservation Theorems

Preservation theorems for first-order logic provide a correspondence between syntactic and semantic restrictions.

A sentence φ is equivalent to an existential sentence if, and only if, the models of φ are closed under extensions.

Łoś-Tarski

A sentence φ is equivalent to one that is positive in the relation symbol R if, and only if, it is monotone in the relation R. **Lyndon**.

Proving Preservation

In each of the cases, it is trivial to see that the syntactic restriction implies the semantic restriction.

The other direction, of *expressive completeness*, is usually proved using compactness.

For example, if φ is closed under extensions:

Take Φ to be the existential consequences of φ and show $\Phi \models \varphi$ by:

$$\begin{array}{ccc} \mathbb{A} \models \Phi \cup \{\varphi\} & \preceq & \mathbb{A}^* \\ & & \cap \\ \mathbb{B} \models \Phi \cup \{\neg\varphi\} & \preceq & \mathbb{B}^* \end{array}$$

Relativised Preservation

We are interested in relativisations of expressive completeness to classes of structure $\mathcal{C}$:

> If φ satisfies the semantic condition restricted to $\mathcal{C}$, it is equivalent (on $\mathcal{C}$) to a sentence in the restricted syntactic form.

If $\mathcal{C}$ satisfies compactnes, then the preservation property necessarily holds in $\mathcal{C}$.

Restricting the class $\mathcal{C}$ in this statement weakens both the hypothesis and the conclusion.

Both Łoś-Tarski and Lyndon are known to fail when $\mathcal{C}$ is the class of all finite structures.

Preservation under Extensions in the Finite

(Tait 1959) showed that there is a φ preserved under extensions on finite structures, but not equivalent to an existential sentence.

- Either $\leq$ is not a linear order;
- or $R(x, z)$ for some x, y, z with $x < y < z$;
- or R contains a cycle.

For any existential sentence whose finite models include all of the above, we can find a model that does not satisfy these conditions.

Preservation Theorems in the Finite

A sentence φ is equivalent to an existential positive sentence if, and only if, the models of φ are closed under homomorphisms.

This has recently been shown to hold in the finite.

(Rossman 2005)

A first-order formula φ is equivalent to a *modal* formula if, and only if, φ is closed under bisimulations.

(van Benthem 1983)

This also holds in the finite.

(Rosen 1995)

Homomorphism Preservation in the Finite

The overall architecture of Rossman's proof of the finite homomorphism preservation theorem is as follows.

Writing $\mathbb{A} \preceq_l \mathbb{B}$ to denote that all existential positive formulas of quantifier rank at most l that are true in $\mathbb{A}$ are also true in $\mathbb{B}$:

$\forall m \exists l$

$$\begin{array}{ccc} \mathbb{A} & \preceq_l & \mathbb{B} \\ \updownarrow & & \uparrow \\ \mathbb{A}^* & \equiv_m & \mathbb{B}^* \end{array}$$

Restricted Classes

In this talk, we look at classes of finite structures restricted by the form of their *adjacency* (or *Gaifman*) graph.

> This is the graph on the universe of the structure where two nodes are adjacent if they appear together in some relation.

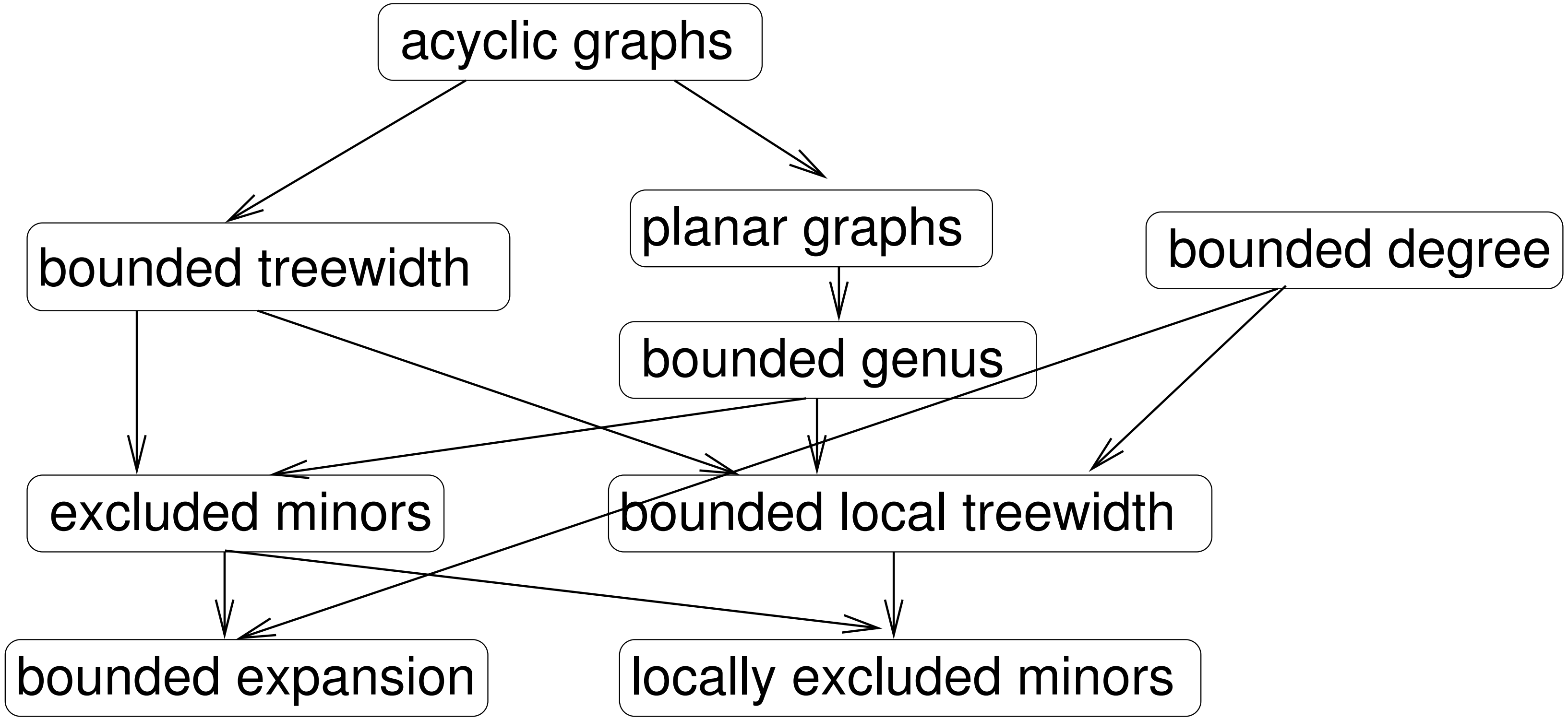
The classes are based on classes of graphs with good *algorithmic* properties.

Do they also have good *model-theoretic* properties?

While this talk focuses on preservation properties, many other natural model-theoretic questions arise.

Well-Behaved Classes of Finite Structures

Certain classes of finite structures have been recognized as *well-behaved* (to different degrees):



Well-Behavedness

Often, the good algorithmic properties of a class are explained by, or related to, a logical result.

On any class of structures of bounded tree-width, $\mathbb{A} \models \varphi$ for an MSO formula φ is decidable in time $f(\varphi)O(|\mathbb{A}|)$. (MSO is fixed-parameter tractable on classes of bounded tree-width.)

(Courcelle 1990).

Any first-order definable set-optimisation problem has a *polynomial approximation scheme* on any class of graphs that excludes a minor.

(D., Grohe, Kreutzer, Schweikardt 2006).

LFP$+$counting captures P on the class of planar graphs.

(Grohe 2000).

Preservation on Well-Behaved Classes

Note: While *acyclicity* and *planarity* are restrictions on graphs, the others in the list are only restrictions on *classes of graphs*.

Knowing that a preservation theorem holds (or fails) on a class $\mathcal{C}$ and that $\mathcal{C}' \subseteq \mathcal{C}$ does not allow us to conclude anything about the corresponding preservation property for $\mathcal{C}'$.

The question arises anew for every class.

We look specifically at the homomorphism and extension preservation properties.

Minimal Models

If φ is a first-order sentence whose models are closed under homomorphisms, then we say that $\mathbb{A}$ is a *minimal* model of φ if,

$\mathbb{A} \models \varphi$ and no proper substructure $\mathbb{B}$ of $\mathbb{A}$ is a model of φ

If φ is a first-order sentence whose models are closed under extensions, then we say that $\mathbb{A}$ is a *minimal* model of φ if,

$\mathbb{A} \models \varphi$ and no proper *induced* substructure $\mathbb{B}$ of $\mathbb{A}$ is a model of φ

Minimal Models

A sentence φ that is invariant under homomorphisms is equivalent to an existential positive sentence if, and only if, it has finitely many minimal models.

A sentence φ that is invariant under extensions is equivalent to an existential sentence if, and only if, it has finitely many minimal models.

As a general strategy, to prove (extension or homomorphism) preservation theorems on a class of finite structures $\mathcal{C}$, we aim to show that for any sentence φ, there is an N such that all minimal models of φ in $\mathcal{C}$ have at most N elements.

Extension Preservation on Forests

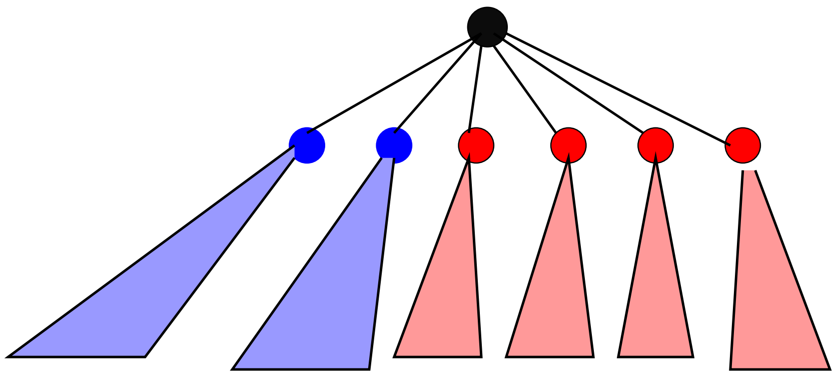
Theorem

The extension preservation property holds in the class of finite acyclic graphs.

(Atserias, D., Grohe, 2005)

Let φ with quantifier rank m be closed under extensions in this class.

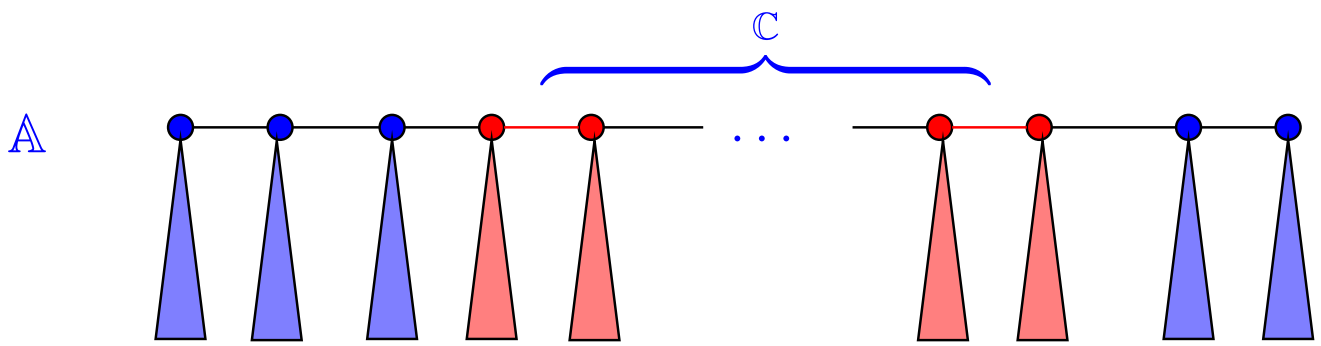
As $\equiv_m$ has finite index, a minimal model of φ cannot have arbitrarily large degree.



If there are m copies of $\equiv_m$-equivalent subtrees, delete one to get a proper submodel.

Extension Preservation on Forests

Also, a minimal model of φ cannot have an arbitrarily long path.



$\mathbb{C}$ is a segment of frequently occurring types of subtree.

$$\mathbb{A} - \mathbb{C} \equiv_m \mathbb{A} + \mathbb{C}$$

The extension preservation property holds in any class of acyclic structures closed under substructures and disjoint unions.

Note: non-elementary blow-up in formula size

(D., Grohe, Kreutzer, Schweikardt 2007).

Tree-Width

Tree-width is a measure of how *tree-like* a structure is.

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 $(u, t), (v, t) \in D$.

The *tree-width* 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 $t \in T$,

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

Examples

- Trees have tree-width 1.
- Cycles have tree-width 2.
- The clique K_k has tree-width $k - 1$.
- The $m \times n$ grid has tree-width $\min(m, n)$.

Bounded Tree-Width

Let $\mathcal{T}_k$ be the class of structures of tree-width at most k.

- The extension-preservation property holds. (Atserias, D., Grohe)
- The homomorphism-preservation property holds. (Atserias, D., Kolaitis)

Let $\mathcal{C}$ be a subclass of $\mathcal{T}_k$.

- The extension-preservation property fails, in general. In partiular, it fails on the class of planar graphs of treewidth at most 4. (Atserias, D., Grohe)
- The homomorphism-preservation property holds (provided $\mathcal{C}$ is closed under substructures and disjoint unions). (Atserias, D., Kolaitis)

Wide Classes

Definition

A class of structures $\mathcal{C}$ is said to be *wide* if for every d and m there is an N such that any structure in $\mathcal{C}$ with more than N elements contains a d-scattered set of size m.

Example: Classes of structures of bounded degree.

Definition

A class of structures $\mathcal{C}$ is *almost wide* if there is an s such that for every d and m there is an N such that any structure in $\mathcal{C}$ with more than N elements contains s elements whose removal leaves a d-scattered set of size m.

Example: Trees.

Preservation on Wide Classes

The extension preservation theorem holds in any class $\mathcal{C}$ that is

- *wide*
- closed under taking substructures
- closed under disjoint unions

(Atserias, D., Grohe 2005).

The homomorphism preservation theorem holds in any class $\mathcal{C}$ that is

- *almost wide*
- closed under taking substructures
- closed under disjoint unions

(Atserias, D., Kolaitis 2004).

Gaifman Locality

Gaifman locality is a key ingredient of the proofs.

A basic local formula $\psi^r(x)$ is a formula in which all quantifiers are relativised to $\mathrm{Nbd}^r(x)$.

A basic local sentence is one of the form

$$\exists x_1 \ldots \exists x_n \bigwedge_{i \neq j} \mathrm{Nbd}^r(x_i) \cap \mathrm{Nbd}^r(x_j) = \emptyset \wedge \bigwedge_i \psi^r(x_i)$$

Theorem (Gaifman 1982)

Every first-order sentence is equivalent to the Boolean combination of basic local sentences.

Homomorphism Preservation on Wide Classes

Show that if φ is preserved by homomorphisms, then a minimal model cannot have a large scattered set. **(Ajtai-Gurevich 1994)**

If $\mathbb{A}$ contains a large enough scattered set, it contains two elements a and a' such that the basic local formulas (up to some suitable quantifier rank) satisfied in $\mathrm{Nbd}^r(a)$ and $\mathrm{Nbd}^r(a')$ are the same.

Let $\mathbb{A}'$ be $\mathbb{A}$ with one tuple containing a removed.

$$\begin{array}{ccc} \mathbb{A} & \xrightarrow{\text{hom}} & \mathbb{A} + n \cdot \mathbb{A}' \\ & & \equiv_q \\ \mathbb{A}' & \xleftarrow{\text{hom}} & n \cdot \mathbb{A}' \end{array}$$

Extension Preservation on Wide Classes

The proof of *extension preservation* is more involved, but again relies on Gaifman locality.

In any model $\mathbb{A}$ with large enough scattered sets, we find a substructure $\mathbb{A}'$ and an extension $\mathbb{B}$ of $\mathbb{A}$ such that

$$\mathbb{A}' \equiv_m \mathbb{B}.$$

We construct $\mathbb{A}'$ in a series of stages by including all *neighbourhoods* of *rare type* and certain neighbourhoods of frequent type.

$\mathbb{B}$ is then obtained as the disjoint union of $\mathbb{A}$ with the neighbourhoods in $\mathbb{A}'$ of frequent type.

The *types* used are MSO types of neighbourhoods.

Graph Minors

We say that a graph $G = (V, E)$ is a minor of graph $H = (U, F)$, (written $G \prec H$) 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, $K_5 \not\prec G$ and $K_{3,3} \not\prec G$.

If $G \prec H$, then tree-width$(G) \leq$ tree-width(H).

The relation $\prec$ is transitive.

If tree-width$(G) < k - 1$, then $K_k \not\prec G$.

$K_k \prec K_{k-1,k-1}$.

A class of graphs $\mathcal{C}$ has bounded treewidth if, and only if, there is some grid G such that $G \not\prec H$ for any $H \in \mathcal{C}$.

Theorem (Robertson-Seymour)

In any infinite collection $\{G_i \mid i \in \omega\}$ of graphs, there are i, j with $G_i \prec G_j$.

Excluded Minor Classes are Almost Wide

A combinatorial construction (based on **(Kreidler and Seese 1999)**) shows that if $\mathcal{C}$ *excludes a graph minor* then $\mathcal{C}$ is almost wide.

I.e.,

> For each k, d and m there is an N such that if G is a graph with $K_k \not\prec G$, and $|G| > N$, then there is a set $B \subset G$ with $|B| < k - 1$ such that $G - B$ has a d-scattered set of size m.

This is established by starting with a large set S in G so that we can repeatedly find a large enough subset S' and expand the radius of neighbourhoods of elements in S' while keeping them disjoint. To do this we may have to delete some elements of G, but we do not delete more than $k - 1$ elements in total.

This involves an iterated Ramsey argument.

Quasi-wide Classes of Structures

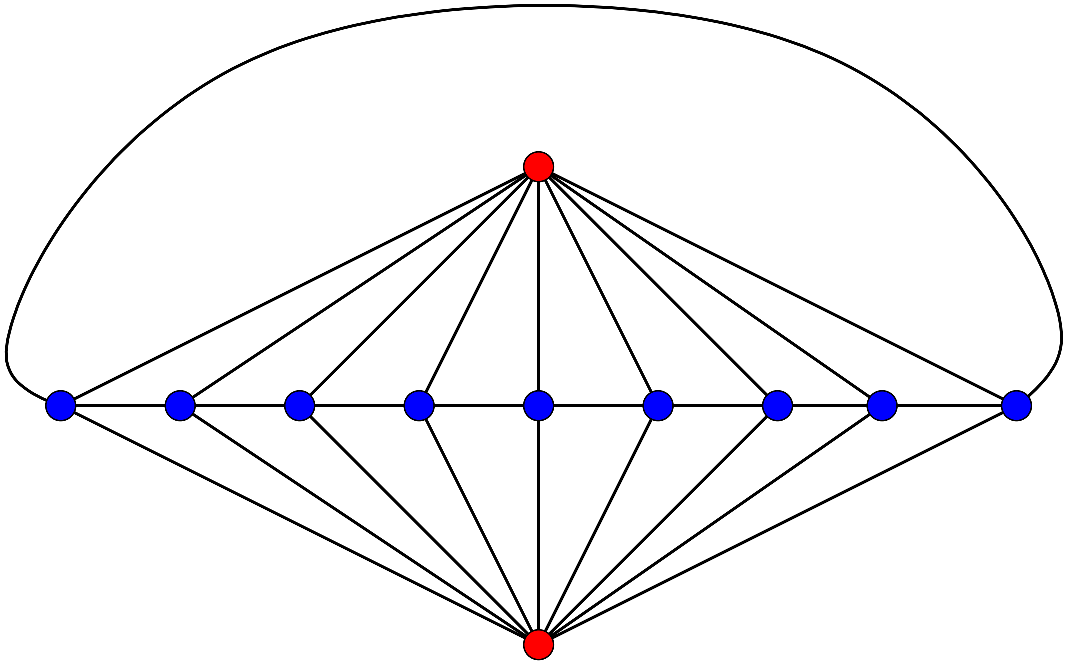
Say that a class of structures $\mathcal{C}$ is *quasi-wide* if

$$\forall d \exists s \forall m \exists N \quad \mathbb{A} \in \mathcal{C} \text{ and } |\mathbb{A}| > N \Rightarrow$$

$\mathbb{A}$ contains a set of s elements whose removal leaves a d-scattered set of size m.

- Classes of *bounded expansion* (as defined by **(Nesetril and Ossona de Mendez 2005)** are quasi-wide.
- Classes that *locally exclude a minor* (defined in **D., Grohe, Kreutzer 2007)** are quasi-wide.
- The homomorphism preservation theorem holds in any quasi-wide class that is closed under disjoint unions and substructures (by a strengthening of the Ajtai-Gurevich lemma). **(D., Malod, forthcoming)**.

Extension Preservation Fails on Planar Graphs



"There are two red vertices such that *if* every other vertex is a neighbour of both *then* every vertex has at least two blue neighbours".

Conclusion

The class of *all* finite structures is not well-behaved in a model-theoretic sense.

Putting additional structural restrictions allows us to recover some interesting model theory.

The restrictions often coincide with those giving interesting algorithmic properties.

Besides *Preservation* Theorems, many other properties remain to be explored.

In the absence of *Compactness*, the proof methods are varied and often highly combinatorial, though *Locality* plays an important role.