IC/94/273
INTERNAL REPORT
(Limited Distribution)

International Atomic Energy Agency
and
United Nations Educational Scientific and Cultural Organization
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

PRESERVATION THEOREMS ON FINITE STRUCTURES

Michel Hébert[1]
International Centre for Theoretical Physics, Trieste, Italy.

ABSTRACT

This paper concerns classical Preservation results applied to finite structures. We consider binary relations for which a strong form of preservation theorem (called strong interpolation) exists in the usual case. This includes most classical cases: embeddings, extensions, homomorphisms into and onto, sandwiches, etc. We establish necessary and sufficient syntactic conditions for the preservation theorems for sentences and for theories to hold in the restricted context of finite structures. We deduce that for all relations above, the restricted theorem for theories hold provided the language is finite. For the sentences the restricted version fails in most cases; in fact the "homomorphism into" case seems to be the only possible one, but the efforts to show that have failed ([2]). We hope our results may help to solve this frustrating problem; in the meantime, they are used to put a lower bound on the level of complexity of potential couterexamples.

MIRAMARE – TRIESTE
September 1994

[1]Permanent Address: Department of Science/Math. Unit, The American University in Cairo, 113 Sharia Kasr El-Aini, P.O. Box 2511, Cairo, Egypt.
e-mail: mhebert@egaucacs.bitnet; mhebert@auc-acs.eun.eg

In this paper, "sentence" means "finitary first-order sentence", but we will meet infinitary disjunctions and conjunctions of such. A theory is a set of (finitary) sentences. We repeat a few definitions from [5] and [6] to make the present paper self-contained.

A relation R (of any arity) on structures gives rise to a so-called (binary) q-relation $\mathbf{R^*}$ on theories defined by $T_1\mathbf{R^*}T_2$ iff for every $\mathfrak{B}\models T_2$ there exists a set $\{\mathfrak{A}_i\}_I$ of models of T_1 such that $\{\mathfrak{A}_i\}_I R\mathfrak{B}$. For example the product relation P, where $\{\mathfrak{A}_i\}_I P\mathfrak{B}$ means $\mathfrak{B}\cong\prod_I\mathfrak{A}_i$. If R is binary, then we write $\mathfrak{A}R\mathfrak{B}$ instead of $\{\mathfrak{A}\}R\mathfrak{B}$. The relations considered here will be the following:

S: $\mathfrak{A}S\mathfrak{B}$ means there exists an embedding $\mathfrak{B}\rightarrowtail\mathfrak{A}$.

H_0: $\mathfrak{A}H_0\mathfrak{B}$ means there exists a homomorphism onto $\mathfrak{A}\twoheadrightarrow\mathfrak{B}$.

H_i: $\mathfrak{A}H_i\mathfrak{B}$ means there exists a homomorphism (into) $\mathfrak{A}\rightarrow\mathfrak{B}$.

Sw: $\mathfrak{A}Sw\mathfrak{B}$ means there exist embeddings $f:\mathfrak{A}\rightarrowtail\mathfrak{B}'$ and $g:\mathfrak{B}'\rightarrowtail\mathfrak{A}'$ such that gf is an elementary embedding and $\mathfrak{B}'$ is an elementary substructure of $\mathfrak{B}$.

A set of sentences Δ also gives rise to a (q-) relation $\mathbf{\Delta^*}$ between theories, defined by $T_1\mathbf{\Delta^*}T_2$ iff $[\ T_1\vdash\delta\in\Delta \implies T_2\vdash\delta\ ]$. We say that *R admits strong Δ-interpolation* if $\mathbf{R^*}=\mathbf{\Delta^*}$. If R is binary, then by Compactness one can actually assume the T_i's to be complete, and then the equality $\mathbf{R^*}=\mathbf{\Delta^*}$ amounts to:

$$[\ \mathfrak{A}\equiv\mathfrak{A}'R\mathfrak{B}'\equiv\mathfrak{B}\ ] \text{ iff } [\ \mathfrak{A}\models\delta\in\Delta \implies \mathfrak{B}\models\delta\ ].$$

$\mathbf{R^*}=\mathbf{\Delta^*}$ implies in particular that R *admits global relativized Δ-preservation*:

For every theories S and T,

$$[\mathfrak{A},\mathfrak{B}\models S,\ \mathfrak{A}\models T \text{ and } \mathfrak{A}R\mathfrak{B} \implies \mathfrak{B}\models T] \text{ iff } [T\equiv_S\Delta' \text{ for some } \Delta'\subset\Delta],$$

where $\equiv_S$ is the restriction of $\equiv$ to models of S (i.e. $T\equiv_S\Delta'$ means $T\cup S\equiv\Delta'\cup S$). It also implies its "sentence" version, the *local relativized Δ-preservation*:

For all sentences ψ and ϕ,

$$[\mathfrak{A},\mathfrak{B}\models\psi,\ \mathfrak{A}\models\phi \text{ and } \mathfrak{A}R\mathfrak{B} \implies \mathfrak{B}\models\phi] \text{ iff } [\phi\equiv_\psi\delta \text{ for some } \delta\in\Delta].$$

(The usual preservation theorems for theories and sentences correspond of course to the cases where S and ψ above are respectively the universally valid theory and the universally valid sentence).

Note that if a binary relation R admits strong Δ-interpolation, then R^{-1} admits strong $\neg\Delta$-interpolation (where $\neg\Delta=\{\neg\delta \mid \delta\in\Delta\}$), so that relativized preservations hold for R^{-1} too.

Examples of relations which do not admit strong interpolation are the intersection relation and the (categorical) limit relation (see [5] for more details and many more examples of all kinds). But most classical cases of preservation are in fact consequences of strong interpolation: for example, if $\forall$, $\exists\forall$, Pos and $\exists$Pos are respectively the sets of universal, existential-universal, positive and existential-positive sentences, then $\mathbf{S^* = \forall^*}$, $\mathbf{Sw^* = \exists\forall^*}$, $\mathbf{H_0{}^* = Pos^*}$ and $\mathbf{H_1{}^* = \exists Pos^*}$.

To consider the restriction of our relations to finite structures, we first remark that strong interpolation implies *ordered* relativized preservation. For binary relations, this translates as follow:

Global version (i.e. for theories): For any two theories S_1 and S_2 such that $S_1 \vdash S_2$, the statements (i) and (ii) are equivalent:

(i)[For every theory T $[\mathfrak{A},\mathfrak{B} \models S_1,\ \mathfrak{A} \models T$ and $\mathfrak{A} R \mathfrak{B} \Rightarrow \mathfrak{B} \models T]]$

(ii) $[\Delta_1 \equiv_{S_1} T \equiv_{S_2} \Delta_2$ for some $\Delta_1, \Delta_2 \subset \Delta$ such that $\Delta_1 \vdash \Delta_2]$.

The local version is of course obtained by replacing theories S_1, S_2, T, Δ_1 and Δ_2 by sentences.

Proposition 1. If R admits strong Δ-interpolation, then it admits ordered relativized Δ-preservations (local and global).

Proof. For theories: (i) being given, relativized preservation tells us that there exist Δ'_1, $\Delta'_2 \subset \Delta$ such that $\Delta'_1 \equiv_{S_1} T \equiv_{S_2} \Delta'_2$. One then checks easily that $\Delta_i = \{\delta \in \Delta \mid T \cup S_i \vdash \delta\}$, $i=1,2$, will satisfy (ii).

For sentences: Replacing T and S_i by ϕ and ψ_i respectively in the theory case, one obtains $\Delta_1, \Delta_2 \subset \Delta$ such that $\Delta_1 \equiv_{\psi_1} \phi \equiv_{\psi_2} \Delta_2$ and $\Delta_1 \vdash \Delta_2$. By Compactness, $\delta_i' \wedge \psi_i \vdash \phi \wedge \psi_i$ for some $\delta_i' \in \Delta$, $i=1,2$. By Compactness again, there exists $\delta_1'' \in \Delta_1$ such that $\delta_1'' \vdash \delta_2'$. Then for $\delta_1 := \delta_1' \wedge \delta_1''$ and $\delta_2 := \delta_2'$, one get $\delta_1 \equiv_{\psi_1} \phi \equiv_{\psi_2} \delta_2$ and $\delta_1 \vdash \delta_2$, as required. ■

In what follows, we will say that a sentence ϕ *is preserved by R on finite structures* if $[\mathfrak{A},\mathfrak{B}$ finite, $\mathfrak{A} \models \phi$ and $\mathfrak{A} R \mathfrak{B} \Rightarrow \mathfrak{B} \models \phi]$. Similarly for theories. The symbols $\models_f$ and $\equiv_f$ will denote the restrictions of $\models$ and $\equiv$ to finite structures. Finally, T_f is the theory of finite structures: $T_f = \{\mu \mid \mathfrak{A} \models \mu$ for every finite $\mathfrak{A}\}$.

Theorem 2. Let R be a binary relation on structures which admits strong Δ-interpolation. Then the following are equivalent for every sentence ϕ:

(a) R preserves ϕ on finite structures

(b) $\phi \equiv_f \bigvee_{n>0} \delta'_n$ for some $\delta'_n \in \Delta$, $n \in \mathbb{N}$

(c) $\phi \equiv_f \bigwedge_{n>0} \delta_n$ for some $\delta_n \in \Delta$, $n \in \mathbb{N}$

(d) If $\mathfrak{B} \models T_f$ and $\mathfrak{A}$ is finite, then $[\mathfrak{A} \models \phi$ and $\mathfrak{A} R \mathfrak{B} \Rightarrow \mathfrak{B} \models \phi]$

(e) If $\mathfrak{A} \models T_f$ and $\mathfrak{B}$ is finite, then $[\mathfrak{A} \models \phi$ and $\mathfrak{A} R \mathfrak{B} \Rightarrow \mathfrak{B} \models \phi]$.

Proof. (a) $\Rightarrow$ (b). We apply Proposition 1 (local part) to the sequence $\psi_1 \vdash \psi_2 \vdash \psi_3 \vdash \ldots$, where $\psi_n := \forall x_1 \ldots x_{n+1}(\bigvee_{i \neq j}(x_i = x_j))$. Then there is a sequence $\delta'_1 \vdash \delta'_2 \vdash \delta'_3 \vdash \ldots$ of sentences in Δ such that $\phi \equiv \delta'_n$ on structures of cardinality $\leq n$. $\phi \models_f \bigvee_{n>0} \delta'_n$ is clear. Conversely, if $\mathfrak{A} \models \delta'_n$ for some n and $\mathfrak{A}$ is of cardinality m, then $\mathfrak{A} \models \delta'_k$ for $k = \max(n,m)$, so that $\mathfrak{A} \models \phi$.

(b) $\Rightarrow$ (a) Easy.

(a) $\Leftrightarrow$ (c) R admits strong Δ-interpolation iff R^{-1} admits strong $\neg\Delta$-interpolation, and we can apply (a) $\Leftrightarrow$ (b) to $\neg\phi$ and R^{-1}. Then $\neg\phi \equiv_f \bigvee_{n>0}(\neg\delta_n)$ for some $\delta_n \in \Delta$, so that $\phi \equiv_f \bigwedge_{n>0} \delta_n$.

(b) $\Rightarrow$ (d). From (b), $(\delta'_n \rightarrow \phi) \in T_f$ for each n, so that $T_f \cup \{\bigvee_{n>0} \delta'_n\} \models \phi$. Hence if $\mathfrak{A}$ is finite and $\mathfrak{A} \models \phi$, then $\mathfrak{A} \models \bigvee_{n>0} \delta'_n$, and $\mathfrak{A} R \mathfrak{B} \Rightarrow \mathfrak{B} \models \bigvee_{n>0} \delta'_n$ $\Rightarrow \mathfrak{B} \models \phi$ if $\mathfrak{B} \models T_f$.

(c) $\Rightarrow$ (e). Similar to (b) $\Rightarrow$ (d). Note that (b) implies $T_f \cup \{\phi\} \models \bigwedge_{n>0} \delta_n$.

(e) $\Rightarrow$ (a). Trivial. ■

Note that for R as in Theorem 2, and for any sentence ϕ, we have

$[\phi \equiv_f \delta$ for some $\delta \in \Delta]$ iff $[\phi$ preserves R on models of $T_f]$

since R admits relativized Δ-preservation. This is in particular the case if $T_f \cup \{\bigwedge_{n>0} \delta_n\} \models \phi$ or, equivalently, if $T_f \cup \{\phi\} \models \bigvee_{n>0} \delta'_n$.

W.W. Tait [8] showed that the restriction to finite structures of the usual local preservation theorem for S does not hold, and Gurevich and Shelah did the same for H_0: one can find in [4] examples of sentences ϕ_1 and ϕ_2 which respectively preserve substructures and homomorphic images on finite structures, but are not equivalent (on finite structures) to any sentence respectively in $\forall$ and Pos.

The Sw relation is trivial in the finite case, since $\mathfrak{A}\, Sw\, \mathfrak{B} \Rightarrow \mathfrak{A} \cong \mathfrak{B}$ if $\mathfrak{A}$ and $\mathfrak{B}$ are finite. Hence "sandwiching" finite structures preserves all sentences. The same is of course true for Sw^{-1}, so that we get the following as a consequence of Theorem 2:

Corollary 3. For every sentence ϕ, $\bigwedge_{n>0}\delta_n \equiv_f \phi \equiv_f \bigvee_{n>0}\delta'_n$ for some $\delta_n, \delta'_n \in \forall\exists$, $n \in \mathbb{N}$. The same is true if one replaces $\forall\exists$ by $\exists\forall$. ■

Of course one can easily find a sentence which is not finitely equivalent to any $\exists\forall$-sentence (for example $\forall x \exists y(R(x,y))$: if $\psi = \exists u_1 \ldots u_n \forall Y(\mu)$, μ open, consider the structure $\mathfrak{A}$ on $A = \{a_1, \ldots, a_{n+1}\}$ defined by $R^{\mathfrak{A}} = \{(a_1,a_2),(a_2,a_3),\ldots,(a_{n+1},a_1)\}$; then $\mathfrak{A} \models \phi$, but if $\mathfrak{A} \models \psi$, then $\mathfrak{A} \models \forall Y \mu[a'_1, \ldots, a'_n, Y]$, for some $a'_1, \ldots, a'_n \in A$, and the substructure $\mathfrak{A}'$ on $A-\{a\}$, where $a \notin \{a'_1, \ldots, a'_n\}$, satisfies ψ but not ϕ). Hence the restriction to finite structures of the usual local preservation theorem for S_W does not hold. The same remarks hold for each "n-sandwich" relation S_{W_n} (see [1]). Recall that $\mathbf{S_{W_n}}^* = \exists\forall\exists\forall\ldots\exists\forall^*$ (n appearances of $\exists\forall$).

The H_i relation constitutes a difficult (open) problem. Gurevich and Shelah believed for some time they had proven that the restriction to finite structures of the local preservation theorem held in this case ([4]), but they later found an error in their proof ([2]). The following (very partial) result uses Theorem 2 to determine a lower bound on the complexity of potential counterexamples:

Proposition 4. Let ϕ be a sentence in a finite relational language such that $\phi \in \exists\forall$ or ϕ is of the form $\forall X \exists y(\mu(X,y))$, μ open (y is a single variable).
Then ϕ is preserved by homomorphisms between finite structures iff $\phi \equiv_f \delta$ for some $\delta \in \exists Pos$.

Proof. One first remarks the following two facts:

(i) If $\mathfrak{A} \models \phi \in \exists\forall$, then there exists an embedding

$$\mathfrak{C} \rightarrowtail \mathfrak{A}$$

for some finite model $\mathfrak{C}$ of ϕ.

(ii) If $\mathfrak{B} \models \phi' = \exists X \forall y(\mu'(X,y))$, μ' open, then there exists an homomorphism

$$\mathfrak{B} \twoheadrightarrow \mathfrak{C}$$

onto some finite model $\mathfrak{C}$ of ϕ'.

(i) is very easy. For (ii), if $b_1, \ldots, b_n \in B$ are such that $\mathfrak{B} \models \forall y(\mu'[b_1, \ldots, b_n, y])$, define an equivalence relation Θ on B by $b\Theta b'$ iff $\mathfrak{B} \models \zeta[b_1, \ldots, b_n, b] \leftrightarrow \zeta[b_1, \ldots, b_n, b']$ for every atomic formula ζ. Let C be the (finite) quotient set $\{[b]_\Theta \mid b \in B\}$ of equivalence classes. Note that $[b_i]_\Theta = \{b_i\}$ for each i. Define the structure $\mathfrak{C}$ on C by $\mathfrak{C} \models \zeta[[b_1]_\Theta, \ldots, [b_n]_\Theta, [b]_\Theta]$ iff $\mathfrak{B} \models \zeta[b_1, \ldots, b_n, b]$ for some $b \in [b]_\Theta$ (ζ atomic) and $\mathfrak{C} \models \zeta[c_1, \ldots, c_k]$ for all other atomic

formulas without equality (that is, all $P(c_1, \ldots, c_k)$'s, P a relation symbol, where at least two of the c_i's are distinct and different from each $[b_i]_\Theta$). Then $\mathfrak{C} \models \phi'$ and the function $b \mapsto [b]_\Theta$ defines a homomorphism $\mathfrak{B} \twoheadrightarrow \mathfrak{C}$.

(i) and (ii), together with parts (d) and (e) of Theorem 2, give the result: given any homomorphism $\mathfrak{A} \to \mathfrak{B}$ between models of T_f with $\mathfrak{A} \models \phi$, we have to show (according to the remark after the theorem) that $\mathfrak{B} \models \phi$. If $\phi \in \exists\forall$, then $\mathfrak{C} \rightarrowtail \mathfrak{A} \to \mathfrak{B}$ for some finite $\mathfrak{C} \models \phi$, which implies $\mathfrak{B} \models \phi$ by the theorem. For the second type of ϕ, take $\phi' = \neg\phi$. Then $\mathfrak{B} \models \neg\phi$ would imply the existence of some $\mathfrak{B} \twoheadrightarrow \mathfrak{C}$ for some finite $\mathfrak{C} \models \neg\phi$, and the composed homomorphism $\mathfrak{A} \to \mathfrak{B} \twoheadrightarrow \mathfrak{C}$ implies $\mathfrak{A} \models \neg\phi$, a contradiction. ■

Note that the same proof shows that the $\exists\forall$-sentences (respectively the sentences of the form $\forall X \exists y(\mu(X,y))$, μ open) which are preserved by extensions (respectively homomorphic images) on finite structures are finitely equivalent to existential (resp. positive) sentences. The result for extensions was proved by Kevin Compton (see [3]).

All this means that for no one of the relations considered yet we have a syntactic characterization, in the finite context, of the sentences preserved under the relation. Our last result show that the situation is radically different for theories:

Theorem 5. Assume the language to be finite, and let R be a binary relation admitting strong Δ-interpolation. Then R admits global Δ-preservation on finite structures: for every theory T, we have

$$[[\mathfrak{A}, \mathfrak{B} \text{ finite}, \mathfrak{A} \models T \text{ and } \mathfrak{A} R \mathfrak{B}] \Rightarrow \mathfrak{B} \models T] \text{ iff } [T \equiv_f \Delta' \text{ for some } \Delta' \subset \Delta]$$

Proof. From Proposition 1, one gets, for $\psi_n = \forall x_1 \ldots x_{n+1}(\bigvee_{i \neq j}(x_i = x_j))$, $T \cup \{\psi_n\} \equiv \Delta_n \cup \{\psi_n\}$ for every $n>0$, $\Delta_n \subset \Delta$ and $\Delta_n \vdash \Delta_{n+1}$. But in a finite language, there are only a finite number of structures of cardinality $\leq n$. This implies we can replace each Δ_n by a sentence $\delta_n \in \Delta$, so that $T \cup \{\psi_n\} \equiv \delta_n \wedge \psi_n$ for each n, and $\delta_n \vdash \delta_{n+1}$. Hence $T \equiv_f \bigvee_{n>0}\delta_n$. Now $\neg(\bigvee_{n>0}\delta_n) = \bigwedge_{n>0}(\neg\delta_n)$ preserves R^{-1} on finite structures, and we can apply again our argument, this time to R^{-1} and the theory $T' = \{\neg\delta_n \mid n>0\}$. Then $T' \equiv_f \bigvee_{n>0}\delta'_n$ for some $\delta'_n \in \neg\Delta$, and this implies $T \equiv_f \{\neg\delta'_n \mid n>0\} \subset \Delta$, as required. ■

Libo Lo ([7]) has announced the result of Theorem 5 for the relations S and H_i^{-1}. We have not been able to contact him to see his proof, but it must be different from ours, since the full statement of his result clearly implies that he could not obtain the result for S^{-1} and H_i.

Acknowledgments

The author would like to thank Professor Abdus Salam, the International Atomic Energy Agency and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

References

[1] Chang, C.C., Keisler, H.J., *Model Theory*, North-Holland, Amsterdam 1990.

[2] Gurevich, Y., *Finite model theory*, Abstracts of papers: Veszprém, 1992, J. Symb. Logic *58* (1993) p. 1101.

[3] Gurevich, Y., *Toward logic tailored for computational complexity*, in Computation and Proof Theory, Lect. Notes in Math. 1104, Springer-Verlag, Berlin 1984, 175–216.

[4] Gurevich, Y., *On finite model theory*, in Feasible Mathematics (S.R. Buss and P.J. Scott, editors), Birkhauser, Boston 1990, 211-219.

[5] Hatcher, W.S. and Hébert, M., *Preservation Theory*, Manuscript, 1992.

[6] Hébert, M., *Preservation and interpolation through binary relations between theories*, Z. fur math. Logik Grundl. d. Math. 35 (1989), 169-182.

[7] Lo, L., Preservation theorems of finite models, Abstracts of papers: Durham 1992, J. Symb. Logic [illegible] (1993),

[8] Tait, W.W., *A counterexample to a conjecture of Scott and Suppes*, J. Symb. Logic 24 (1959), 15-16.