HANF NUMBERS AND PRESENTATION THEOREMS IN AECS

JOHN BALDWIN
DEPARTMENT OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE
UNIVERSITY OF ILLINOIS AT CHICAGO

WILL BONEY
MATHEMATICS DEPARTMENT
HARVARD UNIVERSITY

1. Introduction

This paper addresses a number of fundamental problems in logic and the philosophy of mathematics by considering some more technical problems in model theory and set theory. The interplay between syntax and semantics is usually considered the hallmark of model theory. At first sight, Shelah's notion of abstract elementary class shatters that icon. As in the beginnings of the modern theory of structures ([Cor92]) Shelah studies certain classes of models and relations among them, providing an axiomatization in the Bourbaki ([Bou50]) as opposed to the Gödel or Tarski sense: mathematical requirements, not sentences in a formal language. This formalism-free approach ([Ken13]) was designed to circumvent confusion arising from the syntactical schemes of infinitary logic; if a logic is closed under infinite conjunctions, what is the sense of studying types? However, Shelah's presentation theorem and more strongly Boney's use [Bon] of aec's as theories of $L_{\kappa,\omega}$ (for κ strongly compact) reintroduce syntactical arguments. The issues addressed in this paper trace to the failure of infinitary logics to satisfy the *upward* Löwenheim-Skolem theorem or more specifically the compactness theorem. The compactness theorem allows such basic algebraic notions as amalgamation and joint embedding to be easily encoded in first order logic. Thus, all complete first order theories have amalgamation and joint embedding in all cardinalities. In contrast these and other familiar concepts from algebra and model theory turn out to be heavily cardinal-dependent for infinitary logic and specifically for abstract elementary classes. This is especially striking as one of the most important contributions of modern model theory is the freeing of first order model theory from its entanglement with axiomatic set theory ([Bal15a], chapter 7 of [Bal15b]).

Two main issues are addressed here. We consider not the interaction of syntax and semantics in the usual formal language/structure dichotomy but methodologically. What are reasons for adopting syntactic and/or semantic approaches to a particular topic? We compare methods from the very beginnings of model theory with semantic methods powered by large cardinal hypotheses. Secondly, what then are the connections of large cardinal axioms with the cardinal dependence of algebraic properties in model theory. Here we describe the opening of the gates for potentially large interactions between set theorists (and incidentally graph theorists) and model theorists. More precisely, can the combinatorial properties of small large cardinals be coded as structural properties of abstract elementary classes so as to produce Hanf numbers intermediate in cardinality between 'well below the first inaccessible' and 'strongly compact'? More specifically, as in Section 3.3, are there new model theoretic characterizations of small large cardinals?

Most theorems in mathematics are either true in a specific small cardinality (at most the continuum) or in all cardinals. For example all, *finite* division rings are commutative, thus all finite Desarguesian planes are Pappian. But *all* Pappian planes are Desarguean and not conversely. Of course this stricture does not apply to set theory, but the distinctions arising in set theory are combinatorial. First order model theory, to some extent, and Abstract Elementary Classes (AEC) are beginning to provide a deeper exploration of Cantor's paradise: algebraic properties that are cardinality dependent. In this article, we explore whether certain key properties (amalgamation, joint embedding, and their relatives) follow this line. These algebraic properties are structural in the sense of [Cor04].

Much of this issue arises from an interesting decision of Shelah. Generalizing Fraïssé [Fra54] who considered only finite and countable stuctures, Jónsson laid the foundations for AEC by his study of universal and homogeneous relation systems [Jón56, Jón60]. Both of these authors assumed the amalgamation property (AP) and the joint embedding property (JEP), which in their context is cardinal independent. Variants such as disjoint or free amalgamation (DAP) are a well-studied notion in model theory and universal algebra. But Shelah omitted the requirement of amalgamation in defining AEC. Two reasons are evident for this: it is cardinal dependent in this context; Shelah's theorem (under weak diamond) that categoricity in κ and few models in κ^+ implies amalgamation in κ suggests that amalgamation might be a dividing line.

Grossberg [Gro02, Conjecture 9.3] first raised the question of the existence of Hanf numbers for joint embedding and amalgamation in Abstract Elementary Classes (AEC). We define four kinds of amalgamation properties (with various cardinal parameters) in Subsection 1.1 and a fifth at the end of Section 3.1. The first three notions are staples of the model theory and universal algebra since the fifties and treated for first order logic in a fairly uniform manner by the methods of Abraham Robinson. It is a rather striking feature of Shelah's presentation theorem that issues of disjointness require careful study for AEC, while disjoint amalgamation is trivial for complete first order theories..

Our main result is the following:

Theorem 1.0.1. *Let κ be strongly compact and $\boldsymbol{K}$ be an AEC with Löwenheim-Skolem number less than κ. If $\boldsymbol{K}$ satisfies*[1] *AP/JEP/DAP/DJEP/NDJEP for models of size $[\mu, < \kappa)$, then $\boldsymbol{K}$ satisfies AP/JEP/DAP/DJEP/NDJEP for all models of size $\geq \mu$.*

We conclude with a survey of results showing the large gap for many properties between the largest cardinal where an 'exotic' structure exists and the smallest where eventual behavior is determined. Then we provide specific question to investigate this distinction.

Our starting place for this investigation was second author's work [Bon] that emphasized the role of large cardinals in the study of AEC. A key aspect of the definition of AEC is as a mathematical definition with no formal syntax - class of structures satisfying certain closure properties. However, Shelah's Presentation Theorem says that AECs are expressible in infinitary languages, $L_{\kappa,\omega}$, which allowed a proof via sufficiently complete

[1] This alphabet soup is decoded in Definition 1.1.1.

ultraproducts that, assuming enough strongly compact cardinals, all AEC's were eventually tame in the sense of [GV06].

Thus we approached the problem of finding a Hanf number for amalgamation, etc. from two directions: using ultraproducts to give purely semantic arguments and using Shelah's Presentation Theorem to give purely syntactic arguments. However, there was a gap: although syntactic arguments gave characterizations similar to those found in first order, they required looking at the *disjoint* versions of properties, while the semantic arguments did not see this difference.

The requirement of disjointness in the syntactic arguments stems from a lack of canonicity in Shelah's Presentation Theorem: a single model has many expansions which means that the transfer of structural properties between an AEC $\boldsymbol{K}$ and it's expansion can break down. To fix this problem, we developed a new presentation theorem, called the *relational presentation theorem* because the expansion consists of relations rather than the Skolem-like functions from Shelah's Presentation Theorem.

Theorem 1.0.2 (The relational presentation theorem, Theorem 3.2.3)**.** *To each AEC $\boldsymbol{K}$ with $LS(K) = \kappa$ in vocabulary τ, there is an expansion of τ by predicates of arity κ and a theory T^* in $\mathbb{L}_{(2^\kappa)^+,\kappa^+}$ such that $\boldsymbol{K}$ is exactly the class of τ reducts of models of T^*.*

Note that this presentation theorem works in $\mathbb{L}_{(2^\kappa)^+,\kappa^+}$ and has symbols of arity κ, a far cry from the $\mathbb{L}_{(2^\kappa)^+,\omega}$ and finitary language of Shelah's Presentation Theorem. The benefit of this is that the expansion is canonical or functorial (see Definition 3.0.6). This functoriality makes the transfer of properties between $\boldsymbol{K}$ and $(\operatorname{Mod} T^*, \subset_{\tau^*})$ trivial (see Proposition 3.0.7). This allows us to formulate natural syntactic conditions for our structural properties.

Comparing the relational presentation theorem to Shelah's, another well-known advantage of Shelah's is that it allows for the computation of Hanf numbers for existence (see Section 4) because these exist in $\mathbb{L}_{\kappa,\omega}$. However, there is an advantage of the relational presentation theorem: Shelah's Presentation Theorem works with a sentence in the logic $\mathbb{L}_{(2^{LS(\boldsymbol{K})})^+,\omega}$ and there is little hope of bringing that cardinal down[2]. On the other hand, the logic and size of theory in the relational presentation theorem can be brought down by putting structure assumptions on the class $\boldsymbol{K}$, primarily on the number of nonisomorphic extensions of size $LS(\boldsymbol{K})$, $|\{(M,N)/\cong : M \prec_{\boldsymbol{K}} N \text{ from } \boldsymbol{K}_{LS(\boldsymbol{K})}\}|$. We can get slightly better amalgamation results by using weaker large cardinals, such as accessible cardinals with the weakly compact embedding property; see Section 3.3.

We would like to thank Spencer Unger and Sebastien Vasey for helpful discussions regarding these results.

1.1. **Preliminaries.**

We discuss the relevant background of AECs, especially for the case of disjoint amalgamation.

Definition 1.1.1. *We consider several variations on the joint embedding property, written JEP or JEP$[\mu, \kappa)$.*

[2]Indeed an AEC $\boldsymbol{K}$ where the sentence is in a smaller logic would likely have to have satisfy the very strong property that there are $< 2^{LS(\boldsymbol{K})}$ many $\tau(\boldsymbol{K})$ structures that **are not** in $\boldsymbol{K}$

(1) *Given a class of cardinals $\mathcal{F}$ and an AEC $\boldsymbol{K}$, $\boldsymbol{K}_{\mathcal{F}}$ denotes the collection of $M \in \boldsymbol{K}$ such that $|M| \in \mathcal{F}$. When $\mathcal{F}$ is a singleton, we write $\boldsymbol{K}_\kappa$ instead of $\boldsymbol{K}_{\{\kappa\}}$. Similarly, when $\mathcal{F}$ is an interval, we write $< \kappa$ in place of $[\mathrm{LS}(\boldsymbol{K}), \kappa)$; $\leq \kappa$ in place of $[\mathrm{LS}(\boldsymbol{K}), \kappa]$; $> \kappa$ in place of $\{\lambda \mid \lambda > \kappa\}$; and $\geq \kappa$ in place of $\{\lambda \mid \lambda \geq \kappa\}$.*
(2) *An AEC $(\boldsymbol{K}, \prec_{\boldsymbol{K}})$ has the joint embedding property, JEP, (on the interval $[\mu, \kappa)$) if any two models (from $\boldsymbol{K}_{[\mu,\kappa)}$) can be $\boldsymbol{K}$-embedded into a larger model.*
(3) *If the embeddings witnessing the joint embedding property can be chosen to have disjoint ranges, then we call this the* disjoint embedding property *and write $DJEP$.*
(4) *An AEC $(\boldsymbol{K}, \prec_{\boldsymbol{K}})$ has the amalgamation property, AP, (on the interval $[\mu, \kappa)$) if, given any triple of models $M_0 \prec M_1, M_2$ (from $\boldsymbol{K}_{[\mu,\kappa)}$), M_1 and M_2 can be $\boldsymbol{K}$-embedded into a larger model by embeddings that agree on M_0.*
(5) *If the embeddings witnessing the amalgamation property can be chosen to have disjoint ranges except for M_0, then we call this the* disjoint amalgamation property *and write DAP.*

Definition 1.1.2. (1) *A* finite diagram *or* $EC(T, \Gamma)$-class *is the class of models of a first order theory T which omit all types from a specified collection Γ of complete types in finitely many variables over the empty set.*
(2) *Let Γ be a collection of first order types in finitely many variables over the empty set for a first order theory T in a vocabulary τ_1. A $PC(T, \Gamma, \tau)$* class *is the class of reducts to $\tau \subset \tau_1$ of models of a first order τ_1-theory T which omit all members of the specified collection Γ of partial types.*

2. Semantic arguments

It turns out that the Hanf number computation for the amalgamation properties is immediate from Boney's "Łoś' Theorem for AECs" [Bon, Theorem 4.3]. We will sketch the argument for completeness. For convenience here, we take the following of the many equivalent definitions of strongly compact; it is the most useful for ultraproduct constructions.

Definition 2.0.3 ([Jec06].20)**.** *The cardinal κ is strongly compact iff for every S and every κ-complete filter on S can be extended to a κ-complete ultrafilter. Equivalently, for every $\lambda \geq \kappa$, there is a fine*[3]*, κ-complete ultrafilter on $P_\kappa\lambda = \{\sigma \subset \lambda : |\sigma| < \kappa\}$.*

For this paper, "essentially below κ" means "$LS(K) < \kappa$."

Fact 2.0.4 (Łoś' Theorem for AECs)**.** *Suppose K is an AEC essentially below κ and U is a κ-complete ultrafilter on I. Then K and the class of K-embeddings are closed under κ-complete ultraproducts and the ultrapower embedding is a K-embedding.*

The argument for Theorem 2.0.5 has two main steps. First, use Shelah's presentation theorem to interpret the AEC into $L_{\kappa,\omega}$ and then use the fact that $L_{\kappa,\omega}$ classes are closed under ultraproduct by *κ-complete* ultraproducts.

[3] U is fine iff $G(\alpha) := \{z \in P_\kappa(\lambda) | \alpha \in z\}$ is an element of U for each $\alpha < \lambda$.

Theorem 2.0.5. *Let κ be strongly compact and $\boldsymbol{K}$ be an AEC with Löwenheim-Skolem number less than κ.*

- *If $\boldsymbol{K}$ satisfies $AP(<\kappa)$ then $\boldsymbol{K}$ satisfies AP.*
- *If $\boldsymbol{K}$ satisfies $JEP(<\kappa)$ then $\boldsymbol{K}$ satisfies JEP.*
- *If $\boldsymbol{K}$ satisfies $DAP(<\kappa)$ then $\boldsymbol{K}$ satisfies DAP.*

Proof: We first sketch the proof for the first item, AP, and then note the modifications for the other two.

Suppose that $\boldsymbol{K}$ satisfies $AP(<\kappa)$ and consider a triple of models (M, M_1, M_2) with $M \prec_{\boldsymbol{K}} M_1, M_2$ and $|M| \leq |M_1| \leq |M_2| = \lambda \geq \kappa$. Now we will use our strongly compact cardinal. An *approximation* of (M, M_1, M_2) is a triple $\boldsymbol{N} = (N^{\boldsymbol{N}}, N_1^{\boldsymbol{N}}, N_2^{\boldsymbol{N}}) \in (K_{<\kappa})^3$ such that $N^{\boldsymbol{N}} \prec M, N_\ell^{\boldsymbol{N}} \prec M_\ell, N^{\boldsymbol{N}} \prec N_\ell^{\boldsymbol{N}}$ for $\ell = 1, 2$. We will take an ultraproduct indexed by the set X below of approximations to the triple (M, M_1, M_2). Set

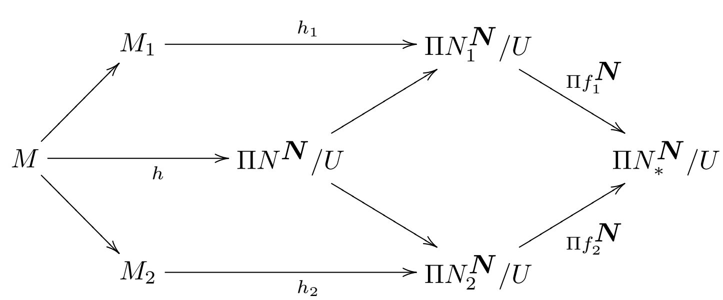
$$X := \{\boldsymbol{N} \in (\boldsymbol{K}_{<\kappa})^3 : \boldsymbol{N} \text{ is an approximation of } (M, M_1, M_2)\}$$

For each $\boldsymbol{N} \in X$, $AP(<\kappa)$ implies there is an amalgam of this triple. Fix $f_\ell^{\boldsymbol{N}} : N_\ell^{\boldsymbol{N}} \to N_*^{\boldsymbol{N}}$ to witness this fact. For each $(A, B, C) \in [M]^{<\kappa} \times [M_1]^{<\kappa} \times [M_2]^{<\kappa}$, define

$$G(A, B, C) := \{\boldsymbol{N} \in X : A \subset N^{\boldsymbol{N}}, B \subset N_1^{\boldsymbol{N}}, C \subset N_2^{\boldsymbol{N}}\}$$

These sets generate a κ-complete filter on X, so it can be extended to a κ-complete ultrafilter U on X; note that this ultrafilter will satisfy the appropriate generalization of fineness, namely that $G(A, B, C)$ is always a U-large set.

We will now take the ultraproduct of the approximations and their amalgam. In the end, we will end up with the following commuting diagram, which provides the amalgam of the original triple.



First, we use Łoś' Theorem for AECs to get the following maps:

$$h : M \to \Pi N^{\boldsymbol{N}}/U$$
$$h_\ell : M_\ell \to \Pi N_\ell^{\boldsymbol{N}}/U \qquad \text{for } \ell = 1, 2$$

h is defined by taking $m \in M$ to the equivalence class of constant function $\boldsymbol{N} \mapsto x$; this constant function is not always defined, but the fineness-like condition guarantees that it is defined on a U-large set (and h_1, h_2 are defined similarly). The uniform definition of these maps imply that $h_1 \mid M = h \mid M = h_2 \mid M$.

Second, we can average the $f_\ell^{\boldsymbol{N}}$ maps to get ultraproduct maps

$$\Pi f_\ell^{\boldsymbol{N}} : \Pi N_\ell^{\boldsymbol{N}}/U \to \Pi N_*^{\boldsymbol{N}}/U$$

These maps agree on $\Pi N^{\boldsymbol{N}}/U$ since each of the individual functions do. As each M_ℓ embeds in $\Pi N_\ell^{\boldsymbol{N}}/U$ the composition of the f and h maps gives the amalgam.

There is no difficulty if one of M_0 or M_1 has cardinality $< \kappa$; many of the approximating triples will have the same first or second coordinates but this causes no harm. Similary, we get the JEP transfer if $M_0 = \emptyset$. And we can transfer disjoint amalgamation since in that case each $N_1^{\boldsymbol{N}} \cap N_2^{\boldsymbol{N}} = N^{\boldsymbol{N}}$ and this is preserved by the ultraproduct. $\dagger_{2.0.5}$

3. Syntactic Approaches

The two methods discussed in this section both depend on expanding the models of $\boldsymbol{K}$ to models in a larger vocabulary. We begin with a concept introduced in Vasey [Vasa, Definition 3.1].

Definition 3.0.6. *A* functorial expansion *of an AEC $\boldsymbol{K}$ in a vocabulary τ is an AEC $\hat{\boldsymbol{K}}$ in a vocabulary $\hat{\tau}$ extending τ such that*

(1) *each $M \in \boldsymbol{K}$ has a* unique *expansion to a $\hat{M} \in \hat{\boldsymbol{K}}$,*
(2) *if $f : M \cong M'$ then $f : \hat{M} \cong \hat{M}'$, and*
(3) *if M is a strong substructure of M' for $\boldsymbol{K}$, then $\hat{M}$ is strong substructure of $\hat{M}'$ for $\hat{\boldsymbol{K}}$.*

This concept unifies a number of previous expansions: Morley's adding a predicate for each first order definable set, Chang adding a predicate for each $L_{\omega_1,\omega}$ definable set, T^{eq}, [CHL85] adding predicates $R_n(\mathbf{x}, y)$ for closure (in an ambient geometry) of $\mathbf{x}$, and the expansion by naming the orbits in Fraîssè model[4].

An important point in both [Vasa] and our relational presentation is that the process does not just reduce the complexity of already definable sets (as Morley, Chang) but adds new definable sets. But the crucial distinction here is that the expansion in Shelah's presentation theorem is not 'functorial' in the sense here: each model has several expansions, rather than a single expansion. That is why there is an extended proof for amalgamation transfer in Section 3.1, while the transfer in Section 3.2 follows from the following result which is easily proved by chasing arrows.

Proposition 3.0.7. *Let $\boldsymbol{K}$ to $\hat{\boldsymbol{K}}$ be a functorial expansion. $(K, \prec)$ has λ-amalgamation [joint embedding, etc.] iff $\hat{\boldsymbol{K}}$ has λ-amalgamation [joint embedding, etc.].*

[4] This has been done for years but there is a slight wrinkle in e.g. [BKL15] where the orbits are not first order definable.

3.1. **Shelah's Presentation Theorem.** In this section, we provide syntactic characterizations of the various amalgamation properties in a finitary language. The results depend directly (or with minor variations) on Shelah's Presentation Theorem and illustrate its advantages (finitary language) and disadvantage (lack of canonicity).

Fact 3.1.1 (Shelah's presentation theorem)**.** *If $\boldsymbol{K}$ is an AEC (in a vocabulary τ with $|\tau| \leq \mathrm{LS}(\boldsymbol{K})$) with Löwenheim-Skolem number $\mathrm{LS}(\boldsymbol{K})$, there is a vocabulary $\tau_1 \supseteq \tau$ with cardinality $|\mathrm{LS}(\boldsymbol{K})|$, a first order τ_1-theory T_1 and a set Γ of at most $2^{\mathrm{LS}(\boldsymbol{K})}$ partial types such that*

1. $\boldsymbol{K} = \{M'|\tau : M' \models T_1 \text{ and } M' \text{ omits } \Gamma\}$*;*
2. *if M' is a τ_1-substructure of N' where M', N' satisfy T_1 and omit Γ then $M'|\tau \prec_{\boldsymbol{K}} N'|\tau$; and*
3. *if $M \prec N \in \boldsymbol{K}$ and $M' \in EC(T_1, \Gamma)$ such that $M'|\tau = M$, then there is $N' \in EC(T_1, \Gamma)$ such that $M' \subset N'$ and $N'|\tau = N$.*

The exact assertion for part 3 is new in this paper; we don't include the slight modification in the standard proofs (e.g. [Bal09, Theorem 4.15]). Note that we have a weakening of Definition 3.0.6 caused by the possibility of multiple 'good' expansion of a model M.

Here are the syntactic conditions equivalent to DAP and DJEP.

Definition 3.1.2.

- Ψ *has* $< \lambda$-DAP satisfiability *iff for any expansion by constants* $\mathbf{c}$ *and all sets of atomic and negated atomic formulas (in $\tau(\Psi) \cup \{\mathbf{c}\}$) $\delta_1(\mathbf{x}, \mathbf{c})$ and $\delta_2(\mathbf{y}, \mathbf{c})$ of size $< \lambda$, if $\Psi \wedge \exists \mathbf{x} \left(\bigwedge \delta_1(\mathbf{x}, \mathbf{c}) \wedge \bigwedge x_i \neq c_j\right)$ and $\Psi \wedge \exists \mathbf{y} \left(\bigwedge \delta_2(\mathbf{y}, \mathbf{c}) \wedge \bigwedge y_i \neq c_j\right)$ are separately satisfiable, then so is*
$$\Psi \wedge \exists \mathbf{x}, \mathbf{y} \left(\bigwedge \delta_1(\mathbf{x}, \mathbf{c}) \wedge \bigwedge \delta_2(\mathbf{y}, \mathbf{c}) \wedge \bigwedge_{i,j} x_i \neq y_j\right)$$
- Ψ *has* $< \lambda$-DJEP satisfiability *iff for all sets of atomic and negated atomic formulas (in $\tau(\Psi)$) $\delta_1(\mathbf{x})$ and $\delta_2(\mathbf{y})$ of size $< \lambda$, if $\Psi \wedge \exists \mathbf{x} \bigwedge \delta_1(\mathbf{x})$ and $\Psi \wedge \exists \mathbf{y} \bigwedge \delta_2(\mathbf{y})$ are separately satisfiable, then so is*
$$\Psi \wedge \exists \mathbf{x}, \mathbf{y} \left(\bigwedge \delta_1(\mathbf{x}) \wedge \bigwedge \delta_2(\mathbf{y}) \wedge \bigwedge_{i,j} x_i \neq y_j\right)$$

We now outline the argument for $DJEP$; the others are similar. Note that $(2) \rightarrow (1)$ for the analogous result with DAP replacing DJEP has been shown by Hyttinen and Kesälä [HK06, 2.16].

Lemma 3.1.3. *Suppose that $\boldsymbol{K}$ is an AEC, $\lambda > LS(\boldsymbol{K})$, and T_1 and Γ are from Shelah's Presentation Theorem. Let Φ be the $L_{LS(\boldsymbol{K})^+, \omega}$ theory that asserts the satisfaction of T_1 and omission of each type in Γ. Then the following are equivalent:*

1. $\boldsymbol{K}_{<\lambda}$ *has* $DJEP$*.*
2. $(EC(T_1, \Gamma), \subset)_{<\lambda}$ *has DJEP.*
3. Φ *has* $< \lambda$-$DJEP$*-satisfiability.*