Calculating Hanf Numbers

John T. Baldwin

July 7, 2011

Universal Domains

The 1970's witnessed two attempts to organize large areas of mathematics by introducing a structure in which all of the mathematics took place.

Grothendieck Universes

Grothendieck explicitly developed cohomology theory using the existence of (a proper class of) universes.
Grothendieck proved that the existence of a single universe is equivalent over ZFC to the existence of a strongly inaccessible cardinal.

Shelah's Universal domain - the monster

$\mathfrak{C}$ is a saturated model of T with cardinality $\overline{\kappa}$ a strongly inaccessible cardinal.
He writes, 'The assumption on $\overline{\kappa}$ does not, in fact, add any axiom of set theory as a hypothesis to our theorems.'

Which theorems?

Grothendieck Universes

summarizing McLarty:
The great applications of cohomology theory (e.g. Wiles and Faltings) as written rely on universes.
This reliance greatly simplifies the organization of the subject.
McLarty (in preparation) reduces

1. ' all of SGA' to Bounded Zermelo plus a Universe.
2. '"the currently existing applications" to Bounded Zermelo itself, thus the consistency strength of simple type theory.'
   The Grothendieck duality theorem and others like it become theorem schema.

What is the strength of Wiles proof?

McLarty writes: ‘Wiles’s proof uses hard arithmetic some of which is on its face one or two orders above PA, and it uses functorial organizing tools some of which are on their face stronger than ZFC.’

to ZFC

Reducing to ZFC

McLarty's current work aims to reduce the 'on their face' strength of the results in cohomology.

Most number theorists regard the applications as requiring much less than their 'on their face' strength and in particular believe the large cardinal appeals are 'easily eliminable'.

Reducing to PA

MacIntyre aims to reduce the 'on their face' strength of results in hard arithmetic.

These programs may be complementary or a full implementation of Macintyre's might avoid the first.

Model theory papers begin

‘We work in a big saturated model.’

Model theory papers begin

‘We work in a big saturated model.’

or slightly more formally

‘We are working in a saturated model of cardinality kappa for sufficiently large kappa (a monster model).’

Model theory papers begin

'We work in a big saturated model.'

or slightly more formally

'We are working in a saturated model of cardinality kappa for sufficiently large kappa (a monster model).'

What does sufficiently mean?

Example

Let M be a big saturated model.

Theorem

For every set $I \subset M$, there exists an infinite sequence of order indiscernibles $J \subset M$ such that for every finite $\mathbf{b} \in J$ there is an $\boldsymbol{a} \in I$ with $\mathrm{tp}(\mathbf{b}/\emptyset) = \mathrm{tp}(\boldsymbol{a}/\emptyset)$.

Scow: modeling property

Example

Let M be a big saturated model.

Theorem

For every set $I \subset M$, there exists an infinite sequence of order indiscernibles $J \subset M$ such that for every finite $\mathbf{b} \in J$ there is an $\boldsymbol{a} \in I$ with $\mathrm{tp}(\mathbf{b}/\emptyset) = \mathrm{tp}(\boldsymbol{a}/\emptyset)$.

Scow: modeling property

We can guarantee this only if $|M| \geq \beth_{\omega_1}$?

Set Theoretic Assumption

To avoid difficulties about finding saturated models, we assume here:

Assumption

Assume the collection of λ with $\lambda^{<\lambda} = \lambda$ is a proper class.

This assumption follows from GCH.
or from

There are a proper class of strongly inaccessible cardinals.

How big is the monster model?

Consider the following situation:

T_1 is a first order theory in a vocabulary L_1.
L is a sublanguage of L_1 and p is a type involving symbols from L_1.

How big is the monster model?

Consider the following situation:

T_1 is a first order theory in a vocabulary L_1.
L is a sublanguage of L_1 and p is a type involving symbols from L_1.

There is a submodel M of the monster model $\mathbb{M}$ such that $|M| = |\mathbb{M}|$ and

1. M omits p
2. $M \restriction L$ is saturated

How big is the monster model?

Consider the following situation:

T_1 is a first order theory in a vocabulary L_1.
L is a sublanguage of L_1 and p is a type involving symbols from L_1.

There is a submodel M of the monster model $\mathbb{M}$ such that $|M| = |\mathbb{M}|$ and

1. M omits p
2. $M \restriction L$ is saturated

How big must $\mathbb{M}$ be to guarantee that every larger saturated model of T satisfies the condition?

Hanf's argument

Ask for each $\boldsymbol{K}$ in a set $\mathcal{K}$ of classes of structures, 'Does $\boldsymbol{K}$ have arbitrarily large members?'

There is a cardinal $\kappa = H(\mathcal{K})$, the Hanf number of $\mathcal{K}$, such that any class with a member larger than κ has arbitrarily large models.

$H(\mathcal{K})$ is the sup of the sups of the bounded $\boldsymbol{K} \in \mathcal{K}$.

We call a Hanf number for a set $\mathcal{K}$ of classes calculable if it is less than the first inaccessible.

Question

What is a better formulation of this concept?

Examples

1. The Hanf number for first order theories with vocabulary of size κ is κ.
2. The Hanf number for $L_{\omega_1,\omega}$ sentences in a countable vocabulary is $\beth_{\omega_1}$.
3. The Hanf number for $L_{\kappa^+,\omega}$ sentences in a vocabulary of cardinality $\leq \kappa$ is at most $\beth_{(2^\kappa)^+}$.

Löwenheim number

The Löwenheim number $\ell(\mathcal{K})$ of a set $\mathcal{K}$ of classes is the least cardinal μ such that any class $\boldsymbol{K} \in \mathcal{K}$ that has a model has one of cardinality $\leq \mu$.

Saturation and Omission

Definition

We say $M_1 \models \boldsymbol{t}$ where $\boldsymbol{t} = (T, T_1, p)$ is a triple of theories in vocabularies $\tau \subset \tau_1$ with $|\tau_1| \leq \lambda$, $T \subseteq T_1$ and p is a τ_1-type over the empty set if M_1 is a model of T_1 which omits p, but $M_1 \restriction \tau$ is saturated.

Let $\boldsymbol{N}_\lambda$ denote the set of $\boldsymbol{t}$ with $|\tau_1| = \lambda$. Then $H(\boldsymbol{N}_\lambda)$ denotes the Hanf number of $\boldsymbol{N}_\lambda$:

$H(\boldsymbol{N}_\lambda)$ is least so that if some $\boldsymbol{t} \in \boldsymbol{N}_\lambda$ has a model of cardinality $H(\boldsymbol{N}_\lambda)$ it has arbitrarily large models.

$H(\boldsymbol{N}_\lambda)$ is not calculable

Theorem

Assume the collection of λ with $\lambda^{<\lambda} = \lambda$ is a proper class.

$$H(\boldsymbol{N}_\lambda) = \ell(L^{II})$$

where L^{II} denotes the collection of sentences of second order logic.

How big is $\ell(L^{II})$?

Jouko Vaananen:
$\ell(L^{II})$ is bigger than the first (second, third, etc) fixed point of any normal function on cardinals that itself can be described in second order logic.

Bigger than the first κ such that $\kappa = \beth_\kappa$, bigger than the first κ such that there are κ cardinals λ below κ such that $\lambda = \beth_\lambda$, etc.

If there are measurable (inaccessible, Mahlo, weakly compact, Ramsey, huge) cardinals, then the Lowenheim number of second order logic exceeds the first of them (respectively, the first inaccessible, Mahlo, weakly compact, Ramsey, huge) (and second, third, etc).

A result of Magidor shows the Lowenheim number of second order logic is always below the first supercompact.

Second order Logic

(Pure) second order logic, L^{II}, means the logic with individual variables and variables for relations of all arities but no non-logical constants.

The atomic formulas are equalities between variables and expressions $X(\mathbf{x})$ where X is an n-ary relation variable and $\mathbf{x}$ is an n-tuple of individual variables.

A structure A for this logic is simply a set so is determined entirely by its cardinality.

But we use the full semantics: the n-ary relation variables range over all n-ary relations on A.

Second order spectra

Definition

Let ψ be a sentence of second order logic.

1. $\mathrm{spec}^1(\psi) = \{\lambda : \lambda \models \psi\}$.
2. $\mathrm{spec}^2(\psi) = \{\lambda : \lambda = \lambda^{<\lambda} \wedge \lambda \models \psi\}$.

Definition

Define H^2 and ℓ^2 to be the Hanf and Lowenheim numbers with respect to spec^2.

There is a sentence χ in second order logic which has a model of size λ if and only if $\lambda^{<\lambda} = \lambda$.
$\lambda^{<\lambda} = \lambda$ denotes this sentence.

Immediate Consequences of Set Theoretic Assumption

Fact

1. $H^1(L^{II}), H^2(L^{II}), \ell^1(L^{II}), \ell^2(L^{II})$ are strong limit cardinals.
2. There is no sentence attaining any of these values exactly. (E.g., there is no $\phi \in L^{II}$ with $\sup(\mathrm{spec}(\phi)) = H^1(L^{II})$.)
3. For either spectrum, $\ell^i(L^{II}) = \sup\{\min\{\mathrm{spec}^i(\phi)\} : \phi \in L^{II}$ has a model$\}$ and similarly $H^i(L^{II}) = \sup\{\sup\{\mathrm{spec}^i(\phi)\} : \phi \in L^{II}$ is bounded$\}$.
4. $H(L^{II}) = H^2(L^{II})$, $\ell(L^{II}) = \ell^2(L^{II})$

Return

Some restrictions

Notation

N_0^f denotes the set of triples $\boldsymbol{t}$ such that T is a finitely axiomatizable in a countable language and p is second order definable.

Dealing with T that are not finitely axiomatizable leads to more complicated arguments for the upper bound.

Upper bound

Lemma

For every $\boldsymbol{t} \in N_0^f$ there is a second order $\phi_{\boldsymbol{t}}$, such that $\phi_{\boldsymbol{t}}$ has a model in λ if and only if $\boldsymbol{t}$ has a model in λ.

Claim

$H(N_0^f) \leq \ell^2(L^{II})$ where L^{II} denotes second order logic.

This argument is slightly more complicated that one would expect because we don't know the smoothness properties of

2nd order smoothness

Key Result

Key Theorem

For every second order sentence ϕ, there is a triple $\boldsymbol{t}_\phi \in N_0^f$ such that if $\lambda^{<\lambda} = \lambda$, then the following are equivalent:

1. $\boldsymbol{t}_\phi$ has a model in λ.
2. ϕ has a model in every cardinal strictly less than λ.

I'll sketch this in a moment. But first:

Lower bound

Corollary to Key Theorem

$H(N_0^f) \geq \ell^2(L^{II})$ where L^{II} denotes second order logic.

Proof of Lower bound from Key Theorem

Proof. Suppose for contradiction that there is a second order sentence ψ such that $\lambda_0 = \min(\mathrm{spec}^2(\psi)) \geq H(N_0^f)$.

By the definition of spec^2, $\lambda_0^{<\lambda_0} = \lambda_0$. Let $\hat{\psi}$ express $(\exists U)(\psi^U \wedge |U|^{<|U|} = |U|)$.

We apply the Key Theorem to $\neg(\hat{\psi})$.
Note that $\hat{\psi}$ is true on all cardinals $\geq \lambda_0$ and false on all $\mu < \lambda_0$. By the Key Theorem,
$\lambda_0 \models \boldsymbol{t}_{\neg(\hat{\psi})}$ and $\lambda_0 \geq H(N_0^f)$.

So $\boldsymbol{t}_{\neg(\hat{\psi})}$ and therefore $\neg(\hat{\psi})$ has arbitrarily large models. But $\neg(\hat{\psi})$ has no models larger than λ_0.

This contradiction yields the theorem.

Sketch of Key Theorem

τ: $Q_1(x), Q_2(x), R(x,y), F_2(x,y)$
$<_0, F(x,y,z), c_0, g$

τ_1: $<_1, c_\omega$

$$p = \{c_0 <_1 x < c_\omega\} \cup \{x \neq g^i(c_0) : i < \omega\}$$

T_1 says Q_2 'codes' subsets of Q_1 and $<_1$ linearly orders Q_1. ($\aleph_1$-Saturation and p omitted guarantees $<_1$ is a well-order; so the coding is correct.

Precise Characterization

Theorem

For any cardinals $\kappa \leq \theta$, the following four cardinals are equal.

1. λ_1 is the Hanf number of N_θ.
2. λ_2 is the Löwenheim number of $L_{\theta^+,\omega}(II) = \ell^2(L_{\theta^+,\kappa}(II))$.
3. λ_3 is the Löwenheim number of $L_{\theta^+,\theta^+}(II) = \ell^2(L_{\theta^+,\theta^+}(II))$.
4. $\lambda_4 = \sup\{\text{spec}(\psi, \theta, \kappa, A_{\tau,\phi}) : \psi \in L(II, \tau_*), \phi \in L_{\theta^+,\theta^+}(II)$, and $A_{\tau,\phi} \subset \theta$ such that $\text{spec}(\psi, \theta, \kappa, A_{\tau,\phi})$ is bounded$\}$.

Superstable Case

$\boldsymbol{N}^{ss}_{\lambda}$ denotes the set of $\boldsymbol{t}$ with $\tau_1 = \lambda$ with the additional requirement that for $M = M_1 \restriction \tau$, $\mathrm{Th}(M)$ is a superstable theory.

$H(\boldsymbol{N}^{ss}_{\lambda})$ is least λ such that:
If $\boldsymbol{t} \in \boldsymbol{N}^{ss}_{\lambda}$ has a model of cardinality $\geq (\boldsymbol{N}^{ss}_{\lambda})$, it has arbitrarily large models.

Theorem (with Shelah, in preparation)

$$H(L_{\lambda^+,\omega}) << H(\boldsymbol{N}^{ss}_{\lambda}) < H(L_{(2^{\lambda})^+,\omega}) \leq \beth_{(2^{(2^{\lambda})^+})^+}.$$

Hanf Function

Definition

Let the Hanf function $H(\mathcal{L}, \theta)$ be the least cardinal μ such that for any vocabulary τ of cardinality $\leq \theta$ and any $\phi \in \mathcal{L}(\tau)$, if ϕ has a model of cardinality μ, it has arbitrarily large models.

For countable vocabularies:

Well-known Theorem

$$\beth_\omega = H(L_{\omega,\omega}(Q), \aleph_0) < H(L_{\omega_1,\omega}, \aleph_0) = \beth_{\omega_1}.$$

Uncountable languages differ

Theorem

If κ is uncountable,

$$H(L_{\omega,\omega}(Q), \kappa) = H(L_{\omega_1,\omega}, \kappa).$$

$|\tau| = |\tau^*| = \kappa$

Chang/Lopez-Escobar: There is a spectrum preserving transformation:

$$\phi \in L_{\omega_1,\omega}(\tau) \to (T, p).$$

$T \in L_{\omega,\omega}(\tau^*)$ and $p = \{P(x)\} \cup \{x \neq c_i : i < \omega\}$

Using uncountable language, replace omitting p by $\neg(Qx)P(x)$.

Transition

The only tool used so far has been ‘coding’ to translate the formal statement of a problem.

Now we consider the model theoretic context where the problems arose. So there are more tools and technical notions which are mostly just alluded to.

Context

Classic Theorem

A theory is unstable iff has the strict order property or the independence property

Chronology

1. 1970-mid-90's: focus on stable theories
2. 95-2005: focus on simple theories
3. post 2005: focus on NIP and generalizing: NIP or simple

Driven by 'generality' and examples: valued fields

Keisler Measure

Definition

A Keisler measure on a model M is finitely additive probability measure on the algebra of all definable subsets of M.

It naturally extends to a measure on the Stone space of M.

Definable Amenability

Context

T is a first order theory extending the theory of groups with countable vocabulary τ. Each model of T acts on itself and therefore on its Stone space by left translation:
$p \mapsto gp = \{\phi(x, g^{-1}\mathbf{a}) : \phi(x, \mathbf{a}) \in p\}$.

Definition

A group G is definably amenable if it admits a left invariant Keisler measure.

Definable Amenability implies

Conversano/Pillay - probably earlier

Let G be a κ-saturated model of an NIP-group where κ is strongly inaccessible:

If G is definably amenable then G has a bounded orbit. That is,
There exists $p \in S(G)$ with $Gp = \{gp : g \in G\}$ such that

$$|Gp| < |G|$$

.

This is straight-forward from Hrushovski-Pillay, NIP and Invariant measures.
Why the inaccessible?

To get Definable Amenability

Conjecture (Petrykowski)

If there is a bounded G orbit in a monster model G then G is definably amenable.

Monster Models

We work officially in Von Neumann, Bernays, Gódel set theory NBG which is a *conservative* extension of ZFC.

Definition

A *monster model* is a class model $\mathbb{M}$ which is a union of saturated models.

Newelski's Context

Assumption

T is a first order theory extending the theory of groups with countable vocabulary τ. Each model of T acts on itself and therefore on its Stone space by left translation: $p \mapsto gp = \{\phi(x, g^{-1}\mathbf{a}) : \phi(x, \mathbf{a}) \in p\}$.

Let $\mathbb{M}$ be the monster model of T. We say $\mathbb{M}$ has a bounded orbit if there is a $p \in S(\mathbb{M})$ such that $\mathbb{M}p = \{ap : a \in \mathbb{M}\}$ is a set.

Stable Case

In the monster model of the countable stable theory of a group there is a type with orbit of cardinality at most $2^{\aleph_0}$.

Newelski's Question

Question

When does $\mathbb{M}$ have a bounded orbit?

That is, can we define a cardinal κ and a property of set models such that $\mathbb{M}$ has a bounded orbit if and only if there is model of cardinality κ that has the property.

Newelski originally reduced the problem to the saturation and omission question. But that doesn't help!

Basic examples

Example

Let T be the theory of $(\mathbb{Q}, +, <)$. Then the type p_+ of an element at $+\infty$ and the type p_- of an element at $-\infty$ both have orbits of size 1.

Example

Consider Z_2^ω with a family H_i of descending definable subgroups so that H_{i+1} has index 2 in H_i. The intersection of the H_i is the connected component, which has an orbit of size $2^{\aleph_0}$.

Some examples

(with Malliaris) By considering equivalence classes with a copy of $(\mathbb{Q}, +, <)$ in each one and functions picking elements out of each class as forming a group under component-wise addition:
For each $\alpha \leq \omega$ there is an unstable example with bounded orbits of size α.

Result (Baldwin-Malliaris)

For any theory T, if there is a model of T which has an invariant type, then arbitrarily large models of T have invariant types.

Is this true if 'invariant' is replaced by 'generic'?
A type $p \in S(M)$ is called *generic* if for each formula $\phi(\mathbf{x}; \boldsymbol{a}) \in p$ there is in an $n_{\phi,\boldsymbol{a}}$ such that $n_{\phi,\boldsymbol{a}}$ conjugates of $\phi(\mathbf{x}; \boldsymbol{a}) \in p$ cover M.

κ-Monsters

1. A structure *M* of infinite cardinality κ is special if *M* is the union of an elementary chain $\langle M_\lambda : \lambda < \kappa, \lambda \text{ a cardinal}\rangle$, where each M_λ is λ^+-saturated.
2. A structure *M* is strongly κ-homogeneous if for every *A* contained in *M* with $|A| < \kappa$, every embedding of *A* into *M* can be extended to an automorphism of *M*.
3. A κ-*monster* model M_κ is a special model of cardinality $\mu = \beth_{\kappa^+}(\kappa)$.

Uniqueness

Fact

A κ-monster is unique up to isomorphism, μ^+-universal and strongly κ^+-homogenous.

This is much tighter that just saying some κ-saturated model.

Extending Orbits

The following remark is implicit in Newelski's paper.

Fact

If T is stable and M is a κ-monster of T with a type $p \in S(M)$ with bounded orbit of cardinality $\lambda < \mu$ then in any κ'-monster $M' \succ M$, there is a type with orbit of cardinality λ.

But in unstable theories this result is false? problematic?
Both the
(algebraic properties true in simple theories, e.g. extension)
and
(the multiplicity properties true in NIP, e.g. stationarity)
are used in the argument.

Hanf function for bounded orbit

Let μ_T be the sup of the cardinalities of κ-monsters for T with bounded orbits if that sup exists; else 0.

$$H(\sigma) = \text{sup}\{\mu_T : |T| = \sigma\}$$

Newelski's problem is: 'Calculate' $H(\sigma)$.

Bounded vrs κ-bounded

Conjecture

The following are equivalent:

1. $\mathbb{M}$ has a bounded orbit.
2. For arbitrarily large κ, the κ-monster M_κ has an orbit of cardinality less than κ.

Proof. Clearly if an orbit of $\mathbb{M}$ is bounded, its restriction to all sufficiently large submodels is also bounded so i) implies ii) is obvious.

For the converse, we have only superstition.

How could one prove that arbitrarily large κ-monsters have a bounded orbit except by a proof that translates to the class monster has bounded orbit?

Shelah's monster models

Newelski's question on the Hanf number of N_λ

A positive resolution of the conjecture would resolve any foundational concerns about Newelski's problem, by giving it a natural formulation, conservative over ZFC via NBG.

Newelski has solved several related problems successfully.

Shelah's monster models

Newelski's question on the Hanf number of N_λ

A positive resolution of the conjecture would resolve any foundational concerns about Newelski's problem, by giving it a natural formulation, conservative over ZFC via NBG.

Newelski has solved several related problems successfully.

Still need a metatheorem

How do we decide on inspection that the N_λ is a 'second order' property?

A Basic Distinction

The vastness of $H(N_\lambda)$ comes from different $\boldsymbol{t}$ having larger maximal models.

The difficulty with Newelski's problem stems from the behavior within individual theories.

Acknowledgements

The main results here are joint with Shelah and appear in the preprints:

A Hanf number for saturation and omission

Calculating Hanf Numbers (not yet posted)
http://www.math.uic.edu/ jbaldwin/model11.html
Newelski, Casanovas, Gismatullin, Conversano & Pillay, Hrushovski-Pillay, McLarty, Macintyre, Vaananen

Mittag-Leffler, Rutgers, Templeton, CIRM-Barcelona