Computable Information from Ultraproducts

Henry Towsner

University of Pennsylvania

October 3, 2015

The ultraproduct construction takes an sequence of first-order structures $\mathfrak{M}_i$ and a non-principal ultrafilter $\mathcal{U}$ and constructs a limiting object

$$\lim_{i \to \mathcal{U}} \mathfrak{M}_i = \mathfrak{M}^{\mathcal{U}}.$$

Ultraproducts have been used to prove new theorems in various fields of math:

- Functional analysis,
- Commutative algebra,
- Ergodic theory,
- Combinatorics.

Ultraproducts have been used to prove new theorems in various fields of math:

- Functional analysis,
- Commutative algebra,
- Ergodic theory,
- Combinatorics.

In various settings, proofs using ultraproducts have been criticized for:

- Foundational concerns

Ultraproducts have been used to prove new theorems in various fields of math:

- Functional analysis,
- Commutative algebra,
- Ergodic theory,
- Combinatorics.

In various settings, proofs using ultraproducts have been criticized for:

- Foundational concerns/the use of a non-canonical construction,

Ultraproducts have been used to prove new theorems in various fields of math:

- Functional analysis,
- Commutative algebra,
- Ergodic theory,
- Combinatorics.

In various settings, proofs using ultraproducts have been criticized for:

- Foundational concerns/the use of a non-canonical construction,
- Being non-constructive,

Ultraproducts have been used to prove new theorems in various fields of math:

- Functional analysis,
- Commutative algebra,
- Ergodic theory,
- Combinatorics.

In various settings, proofs using ultraproducts have been criticized for:

- Foundational concerns/the use of a non-canonical construction,
- Being non-constructive,
- Obscuring the real content of the proof.

Usually the power of the ultraproduct technique comes from the fact that the ultraproduct is a (maybe very large) object of the kind we're studying.

Question

What determines the properties of the ultraproduct?

Usually the power of the ultraproduct technique comes from the fact that the ultraproduct is a (maybe very large) object of the kind we're studying.

Question

What determines the properties of the ultraproduct?

There is a great deal of work by Keisler and others on how different choices of ultrafilter affects the saturation properties of ultraproduct.

Usually the power of the ultraproduct technique comes from the fact that the ultraproduct is a (maybe very large) object of the kind we're studying.

Question

What determines the properties of the ultraproduct?

There is a great deal of work by Keisler and others on how different choices of ultrafilter affects the saturation properties of ultraproduct.

But every proof in the literture uses an essentially arbitrary ultrafilter. So the choice of ultrafilter doesn't matter—everything that matters about the ultraproduct must be a consequence of the properties of the original structures.

The first order properties of an ultraproduct are determined by

Theorem (Łoś's Theorem)

If σ is a sentence of first-order logic, $\mathfrak{M}$ satisfies σ iff for almost every i, $\mathfrak{M}_i$ satisfies σ.

The first order properties of an ultraproduct are determined by

Theorem (Łoś's Theorem)

If σ is a sentence of first-order logic, $\mathfrak{M}$ satisfies σ iff for almost every i, $\mathfrak{M}_i$ satisfies σ.

Any proof making non-trivial use of ultraproducts must make use of properties beyond those controlled by Łoś's Theorem.

To express the properties we are actually interested in, we need a multi-sorted language. We need to distinguish between *internal* and *external* sorts.

Internal sorts get interpreted based on the model we are discussing.

To express the properties we are actually interested in, we need a multi-sorted language. We need to distinguish between *internal* and *external* sorts.

Internal sorts get interpreted based on the model we are discussing.

External sorts have a *fixed* interpretation. In particular, we are allowed to quantify over natural numbers, functions on natural numbers, and so on.

To express the properties we are actually interested in, we need a multi-sorted language. We need to distinguish between *internal* and *external* sorts.

Internal sorts get interpreted based on the model we are discussing.

External sorts have a *fixed* interpretation. In particular, we are allowed to quantify over natural numbers, functions on natural numbers, and so on.

We will write $\exists^{ex}$ or $\forall^{ex}$ to indicate external quantifiers, and $\exists, \forall$ to indicate internal quantifiers.

Definition

A formula ϕ is *internal* if it contains only internal quantifiers.

A Σ_1^{ex} statement is a statement of the form $\exists^{ex} n\, \phi(n)$ where ϕ is internal.

Theorem

$\mathfrak{M}^{\mathcal{U}}$ satisfies a Σ_1^{ex} statement $\exists^{ex} n\, \phi(n)$ if any only if almost every $\mathfrak{M}_i$ satisfies $\exists^{ex} n\, \phi(n)$ uniformly.

That is, there is an n so that almost every $\mathfrak{M}_i$ satisfies $\phi(n)$.

Definition

A Π_2^{ex} statement is a statement of the form $\forall^{ex} m \exists^{ex} n \phi(n, m)$ where ϕ is internal.

Theorem (Transfer Theorem)

$\mathfrak{M}^{\mathcal{U}}$ satisfies a Π_2^{ex} statement $\forall^{ex} m \exists^{ex} n\, \phi(m, n)$ if and only if for each m there is an n so that almost every $\mathfrak{M}_i$ satisfies $\phi(m, n)$.

Definition

A Π_2^{ex} statement is a statement of the form $\forall^{ex} m \exists^{ex} n \phi(n, m)$ where ϕ is internal.

Theorem (Transfer Theorem)

$\mathfrak{M}^{\mathcal{U}}$ satisfies a Π_2^{ex} statement $\forall^{ex} m \exists^{ex} n \, \phi(m, n)$ if and only if for each m there is an n so that almost every $\mathfrak{M}_i$ satisfies $\phi(m, n)$.

Most applications of ultraproducts in the literature (to problems not involving ultraproducts) conclude by showing that the ultraproduct satisfies a Π_2^{ex} statement. Note that when $\mathfrak{M}_i = \mathfrak{M}$ for a fixed structure, this implies that Π_2^{ex} statements hold in $\mathfrak{M}$ iff they hold in $\mathfrak{M}^{\mathcal{U}}$.

If a proof using an ultraproduct consists entirely of Π_2^{ex} statements, we could replace each statement with the corresponding uniform statement about the original structures.

We make the obvious analogy between Π_2^{ex} statements and Π_2 statements. If a proof consists entirely of Π_2 statements, it has computable bounds.

We make the obvious analogy between Π_2^{ex} statements and Π_2 statements. If a proof consists entirely of Π_2 statements, it has computable bounds.

If a proof has intermediate steps which are not Π_2, the functional interpretation lets us extract the computable information from these intermediate steps.

Various results have generalized this to other "Π_2-like" statements:

- Hernest (distinction between internal and external quantifiers),
- Avigad-Towsner (ordinal bounds),
- van den Berg-Briseid-Safarik (standard data from nonstandard proofs),
- Sanders (standard data from nonstandard proofs).

A modification of the functional interpretation gives us a transformation on statements with the following properties:

- σ is Π_2^{ex} then σ^T is equivalent to σ,
- if σ implies τ then σ^T implies τ^T,
- σ^T is Π_2^{ex} (where the quantifiers may be over higher-order functionals).

A modification of the functional interpretation gives us a transformation on statements with the following properties:

- σ is Π_2^{ex} then σ^T is equivalent to σ,
- if σ implies τ then σ^T implies τ^T,
- σ^T is Π_2^{ex} (where the quantifiers may be over higher-order functionals).

Therefore, given a proof using ultraproducts, we can do the following:

- list the lemmas in the proof, $\sigma_1, \sigma_2, \ldots, \sigma_k$,
- replace each lemma σ_i with the corresponding σ_i^T,
- conclude that each σ_i^T holds uniformly in the ground models,
- translate the proof that σ_i implies σ_{i+1} into a proof that uniform bounds on σ_i^T imply uniform bounds on σ_{i+1}^T,
- the conclusion is a constructive, ultraproduct-free proof of σ_k.

The best known example is convergence: the statement that a sequence a_n converges is Π_3^{ex}:

$$\forall^{ex}\epsilon > 0 \exists^{ex} n \forall^{ex} m \geq n \ |a_n - a_m| < \epsilon.$$

The best known example is convergence: the statement that a sequence a_n converges is Π_3^{ex}:

$$\forall^{ex}\epsilon > 0 \exists^{ex} n \forall^{ex} m \geq n \; |a_n - a_m| < \epsilon.$$

The functional interpretation tells us that the sequence converges in the ultraproduct if it converges *uniformly metastably* in the original models:

> *For every* $\epsilon > 0$ *and every function* $F : \mathbb{N} \to \mathbb{N}$*, there is an* n *so that, for almost every* i*,* $\mathfrak{M}_i \vDash |a_n - a_{F(n)}| < \epsilon$*.*

The Gilmore-Robinson characterization of Hilbertian fields says that a field k with characteristic 0 is Hilbertian if

There is a $t \in k^{\mathcal{U}}$ *such that* $\overline{k(t)} \cap k^{\mathcal{U}} = k(t)$.

This has the form

$$\exists t \forall^{ex} p \exists^{ex} U \cdots .$$

The functional interpretation tells us a characteristic 0 field is Hilbertian iff

There are functions

- $U : \mathcal{P}_{\text{fin}}(k) \times \mathbb{N} \to \mathcal{P}_{\text{fin}}(k)$, *and*
- $D : \mathcal{P}_{\text{fin}}(k) \times \mathbb{N} \to \mathbb{N}$

such that for any finite sets $S, T \subseteq k$ *and any natural number* b, *there is a* $t \in k \setminus T$ *so that for each* $S_0 \subseteq S$ *and* $b_0 \leq b$, *whenever* $p \in k[x]$ *such that*

- *the degree of* p *is at most* b_0
- *each coefficient in* p *has the forum* $\sum_{i \leq b_0} a_i t^{c_i}$ *where* $a_i \in S_0$ *and* $|c_i| \leq b_0$,

then if p *has a root in* k, p *has a root of the form* $\sum_{i \leq D(S_0, b_0)} a_i t^{c_i}$ *where* $a_i \in U(S_0, b_0)$ *and* $c_i \leq D(S_0, b_0)$.

A theorem from functional analysis depends on applying the following theorem in an ultraproduct:

Theorem

Let $(f_n)_n$ and $(g_p)_p$ be sequences of L^1 functions such that

- *The sequences $(f_n)_n$ and $(g_p)_p$ converge weakly,*
- *For each p, $(f_n g_p)_n$ converges weakly,*
- *For each n, $(f_n g_p)_p$ converges weakly.*

Then $\lim_n \lim_p (f_n g_p)$ and $\lim_p \lim_n (f_n g_p)$ converge weakly to the same function.

The functional interpretation tells us that the corresponding finite structures must uniformly satisfy, for certain sequences of functions f_n, g_p:

For every $\epsilon > 0$, p, n, K, R and any set A, there are $m \geq n$, $q \geq p$, L, and S so that

$$|(f_m g_{S(K(m,q,L,S),R(m,q,L,S))}(A)) - (f_{L(K(m,q,L,S),R(m,q,L,S))} g_q)(A)| < \epsilon.$$

- Ultraproducts are non-canonical, but proofs that use ultraproducts do not make real use of this,

- Ultraproducts are non-canonical, but proofs that use ultraproducts do not make real use of this,
- Ultraproduct proofs contain all the information needed to calculate constructive bounds,

- Ultraproducts are non-canonical, but proofs that use ultraproducts do not make real use of this,
- Ultraproduct proofs contain all the information needed to calculate constructive bounds,
- We know exactly what ultraproducts do in a proof: they describe the uniformity of bounds in a convenient way.

- The analogy between computable information and ultraproduct information is strong from a proof-theoretic perspective. What about from a computability theoretic perspective?

- The analogy between computable information and ultraproduct information is strong from a proof-theoretic perspective. What about from a computability theoretic perspective?
- Besides computable, hyperarithmetic, and standard/ground-model information, are there other kinds of information which have the same behavior?