

Beth-like definability results, proof-theoretically

Cécilia Pradic (Swansea University)

j.w.w. Michael Benedikt (Oxford University)
Christoph Wernhard (Postdam University)

September 13th 2024, Proof society workshop

Beth definability theorem

Beth definability

Let $\varphi(R)$ be an FO formula over $\Sigma \uplus \{R\}$

If $\varphi(R)$ implicitly defines R , that is

$$\varphi(R) \wedge \varphi(R') \implies \forall \vec{x}. R(\vec{x}) \iff R'(\vec{x})$$

then there is a corresponding explicit FO definition for R .

That is, we have an FO formula $\psi(\vec{x})$ over Σ such that

$$\varphi(R) \implies \forall \vec{x}. R(\vec{x}) \iff \psi(\vec{x})$$

Beth definability theorem

Beth definability

Let $\varphi(R)$ be an FO formula over $\Sigma \uplus \{R\}$

If $\varphi(R)$ implicitly defines R , that is

$$\varphi(R) \wedge \varphi(R') \implies \forall \vec{x}. R(\vec{x}) \iff R'(\vec{x})$$

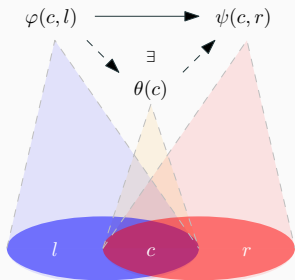
then there is a corresponding explicit FO definition for R .

That is, we have an FO formula $\psi(\vec{x})$ over Σ such that

$$\varphi(R) \implies \forall \vec{x}. R(\vec{x}) \iff \psi(\vec{x})$$

- Model-theoretic proof using amalgamation
- **Proof-theoretic effective proof using interpolation**

Craig interpolation



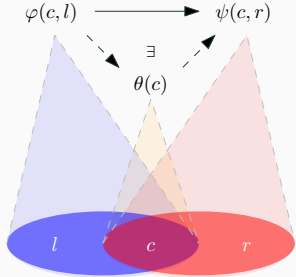
Craig interpolation

If $\varphi \Rightarrow \psi$, there exists θ such that

$$\varphi \Rightarrow \theta \quad \text{and} \quad \theta \Rightarrow \psi$$

Further, θ mentions *only* variables/relation symbols common to φ and ψ .

Craig interpolation



Craig interpolation

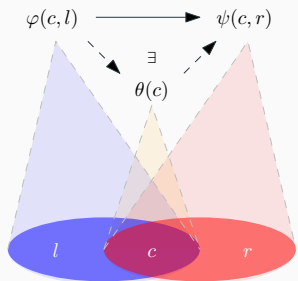
If $\varphi \Rightarrow \psi$, there exists θ such that

$$\varphi \Rightarrow \theta \quad \text{and} \quad \theta \Rightarrow \psi$$

Further, θ mentions *only* variables/relation symbols common to φ and ψ .

- Robust result Δ_0 , intuitionistic/linear logic...

Craig interpolation



Craig interpolation

If $\varphi \Rightarrow \psi$, there exists θ such that

$$\varphi \Rightarrow \theta \quad \text{and} \quad \theta \Rightarrow \psi$$

Further, θ mentions *only* variables/relation symbols common to φ and ψ .

- Robust result
 Δ_0 , intuitionistic/linear logic...
- An actual *factorization* of proofs
[Čubrić 94, Saurin 24]
- computable in $\mathcal{O}(n)$ from cut-free proofs

Beth definability from interpolation

Fix an implicit definition $\varphi(R)$.

Since $\varphi(R)$ determines R , we have

$$\varphi(R) \wedge \varphi(R') \implies \forall \vec{x}. R(\vec{x}) \iff R'(\vec{x})$$

which implies that we have a proof of

$$\varphi(R) \wedge R(\vec{x}) \vdash \varphi(R') \implies R'(\vec{x})$$

Applying interpolation we get $\theta(\vec{x})$ that defines R since we have

$$\varphi(R) \wedge R(\vec{x}) \vdash \theta(\vec{x}) \quad \text{and} \quad \theta(\vec{x}) \wedge \varphi(R) \vdash R(\vec{x})$$

Application to database theory

FO formulas $\psi(\vec{x})$ can be regarded as *queries*

- regard a relation as a table of elements
- $\psi(\vec{x})$ returns a new table

Specification	Implementation
Implicit definition	Explicit definition

Correspondence with (a fragment of) SQL

Restrict that to bounded quantifications.

Plan of the talk

Our work

Extending this to the **nested relational model**.

- Led us to known definability results...
- ...but what about effectivity (and efficiency)?

Plan

1. The setting of nested set
2. Extraction of nested definition
3. Effective relative rigid categoricity?

The nested relational model, logic and NRC

The nested relational model

We work with **typed objects**

Types for nested collections

$$T, U ::= \mathcal{U} \mid \text{Set}(T) \mid 1 \mid T \times U$$

Obvious semantics $\llbracket T \rrbracket$ determined inductively by $\llbracket \mathcal{U} \rrbracket$.

Examples

Taking $\llbracket \mathcal{U} \rrbracket = \text{string}$, we have

$$\{(\text{"seagull"}, \text{"gwyllan"}), (\text{"goats"}, \text{"geifr"}), \dots\} \in \llbracket \text{Set}(\mathcal{U} \times \mathcal{U}) \rrbracket$$

$$\{(\{\text{"shwmae"}, \text{"helo"}\}, \{\text{"hi"}\}), \dots\} \in \llbracket \text{Set}(\text{Set}(\mathcal{U}) \times \text{Set}(\mathcal{U})) \rrbracket$$

$$((), \emptyset, \text{"4"}, \{\text{"1"}, \text{"3"}\}) \in \llbracket 1 \times \text{Set}(\text{Set}(1)) \times \mathcal{U} \times \text{Set}(\mathcal{U}) \rrbracket$$

The nested relational model

We work with **typed objects**

Types for nested collections

$$T, U ::= \mathcal{U} \mid \text{Set}(T) \mid 1 \mid T \times U$$

Obvious semantics $\llbracket T \rrbracket$ determined inductively by $\llbracket \mathcal{U} \rrbracket$.

Examples

Taking $\llbracket \mathcal{U} \rrbracket = \text{string}$, we have

$$\{(\text{"seagull"}, \text{"gwyllan"}), (\text{"goats"}, \text{"geifr"}), \dots\} \in \llbracket \text{Set}(\mathcal{U} \times \mathcal{U}) \rrbracket$$

$$\{(\{\text{"shwmae"}, \text{"helo"}\}, \{\text{"hi"}\}), \dots\} \in \llbracket \text{Set}(\text{Set}(\mathcal{U}) \times \text{Set}(\mathcal{U})) \rrbracket$$

$$((\text{"4"}, \emptyset), \{\text{"1"}, \text{"3"}\}) \in \llbracket 1 \times \text{Set}(\text{Set}(1)) \times \mathcal{U} \times \text{Set}(\mathcal{U}) \rrbracket$$

Nested collection transformations

Types for nested collections

$$T, U ::= \mathfrak{U} \mid \text{Set}(T) \mid 1 \mid T \times U$$

A **transformation** of nested sets is a function $T \rightarrow U$
 \rightarrow is **not** part of the type system

A transformation of **flat** relations

Pre-image of a relation R

$$\begin{aligned} \text{fib} : \quad \text{Set}(\mathfrak{U}) \times \text{Set}(\mathfrak{U} \times \mathfrak{U}) &\rightarrow \text{Set}(\mathfrak{U}) \\ (A, R) &\mapsto R^{-1}(A) = \{x \mid \exists y \in A. (x, y) \in R\} \end{aligned}$$

Nested collection transformations

Types for nested collections

$$T, U ::= \mathcal{U} \mid \text{Set}(T) \mid 1 \mid T \times U$$

Transformations of nested collections

Collect all pre-images of individual elements

$$\begin{aligned} \text{fibs} : \text{Set}(\mathcal{U} \times \mathcal{U}) &\rightarrow \text{Set}(\mathcal{U} \times \text{Set}(\mathcal{U})) \\ R &\mapsto \{(a, \text{fib}(\{a\}, R)) \mid a \in \text{cod}(R)\} \end{aligned}$$

Collect all subsets of the input

$$\begin{aligned} \mathcal{P} : \text{Set}(\mathcal{U}) &\rightarrow \text{Set}(\text{Set}(\mathcal{U})) \\ X &\mapsto \{y \mid y \subseteq X\} \end{aligned}$$

Note: \mathcal{P} is computationally hard on finite instances

Logical specifications

Queries can be specified in *multi-sorted* first-order logic:

- variables explicitly typed $x : T$
- basic predicates $x \in_T z$ and $x =_T y$ $x, y : T$ and $z : \text{Set}(T)$
- terms for tupling and projections

e.g., $\pi_1(((x, z), ()), x) : (T \times \text{Set}(T)) \times 1$

Logical specifications

Queries can be specified in *multi-sorted* first-order logic:

- variables explicitly typed $x : T$
- basic predicates $x \in_T z$ and $x =_T y$ $x, y : T$ and $z : \text{Set}(T)$
- terms for tupling and projections

e.g., $\pi_1(((x, z), ()), x) : (T \times \text{Set}(T)) \times 1$

Consider formulas with only **bounded quantifications**

Δ_0 formulas

$$\varphi, \psi ::= t =_T u \mid t \in_T u \mid \forall x \in t \varphi \mid \varphi \wedge \psi \mid \neg \varphi$$

Examples

Δ_0 formulas

$$\varphi, \psi ::= t =_T u \mid t \in_T u \mid \forall x \in t \varphi \mid \varphi \wedge \psi \mid \neg \varphi$$

Examples

Δ_0 formulas

$$\varphi, \psi ::= t =_T u \mid t \in_T u \mid \forall x \in t \varphi \mid \varphi \wedge \psi \mid \neg \varphi$$

$\varphi_{\text{fib}}(A, R, X)$ for $X = R^{-1}(A)$

- Every $x \in X$ is related to some $a \in A$

$$\forall x \in X. \exists a \in A. (x, a) \in R$$

- For every $(x, y) \in R$, if $y \in A$, then $x \in X$

$$\forall p \in R. \pi_2(p) \in A \Rightarrow \pi_1(p) \in X$$

Examples

Δ_0 formulas

$$\varphi, \psi ::= t =_T u \mid t \in_T u \mid \forall x \in t \varphi \mid \varphi \wedge \psi \mid \neg \varphi$$

$\varphi_{\text{fibs}}(R, O)$ for $O = \{(a, R^{-1}(\{a\})) \mid a \in \text{cod}(R)\}$

- For every $(x, a) \in R$, there is some $(a, X) \in O$ s.t. $x \in X$

$$\forall p \in R. \exists q \in O. \pi_1(p) \in \pi_2(O)$$

- Every element of $(a, X) \in O$ satisfies $\varphi_{\text{fibs}}(\{a\}, R, X)$

$$\forall q \in O. \begin{array}{l} \forall x \in \pi_2(q). (x, \pi_1(q)) \in R \wedge \\ \forall p \in R. \pi_2(p) = \pi_1(q) \Rightarrow \pi_1(p) \in \pi_2(q) \end{array}$$

Explicit definitions: the nested relational calculus (NRC)

Usual terms and rules for variables, tupling, projections plus the following set operators:

$$\frac{\Gamma \vdash e : T}{\Gamma \vdash \{e\} : \text{Set}(T)}$$

$$\frac{\Gamma \vdash e_1 : \text{Set}(T_1) \quad \Gamma, x : T_1 \vdash e_2 : \text{Set}(T_2)}{\Gamma \vdash \bigcup \{e_2 \mid x \in e_1\} : \text{Set}(T_2)}$$

$$\frac{}{\Gamma \vdash \emptyset_T : \text{Set}(T)}$$

$$\frac{\Gamma \vdash e_1 : \text{Set}(T) \quad \Gamma \vdash e_2 : \text{Set}(T)}{\Gamma \vdash e_1 \cup e_2 : \text{Set}(T)}$$

$$\frac{\Gamma \vdash e_1 : \text{Set}(T) \quad \Gamma \vdash e_2 : \text{Set}(T)}{\Gamma \vdash e_1 \setminus e_2 : \text{Set}(T)}$$

Explicit definitions: the nested relational calculus (NRC)

+ Get

Usual terms and rules for variables, tupling, projections plus the following set operators:

$$\frac{\Gamma \vdash e : T}{\Gamma \vdash \{e\} : \text{Set}(T)}$$

$$\frac{\Gamma \vdash e_1 : \text{Set}(T_1) \quad \Gamma, x : T_1 \vdash e_2 : \text{Set}(T_2)}{\Gamma \vdash \bigcup \{e_2 \mid x \in e_1\} : \text{Set}(T_2)}$$

$$\frac{}{\Gamma \vdash \emptyset_T : \text{Set}(T)}$$

$$\frac{\Gamma \vdash e_1 : \text{Set}(T) \quad \Gamma \vdash e_2 : \text{Set}(T)}{\Gamma \vdash e_1 \cup e_2 : \text{Set}(T)}$$

$$\frac{\Gamma \vdash e_1 : \text{Set}(T) \quad \Gamma \vdash e_2 : \text{Set}(T)}{\Gamma \vdash e_1 \setminus e_2 : \text{Set}(T)}$$

$$\frac{\Gamma \vdash e : \text{Set}(T)}{\Gamma \vdash \text{Get}(e) : T}$$

Our running examples

- $(A, R) \mapsto \bigcup \{ \text{case}(\pi_2(p) \in_{\mathcal{U}} A, \{\pi_1(p)\}, \emptyset) \mid p \in R \}$
- $R \mapsto \bigcup \{ \{ \text{fib}(x, R) \} \mid x \in \{ \pi_1(p) \mid p \in R \} \}$

Derivable constructs:

- maps $\{e_1(x) \mid x \in e_2\}$
- at type-level, $\text{Bool} := \text{Set}(1)$
- basic predicates $=_T: T \times T \rightarrow \text{Bool}$, $\in_T: T \times \text{Set}(T) \rightarrow \text{Bool}$
- case analyses

Expressiveness of NRC

Our running examples

- $(A, R) \mapsto \bigcup \{ \text{case}(\pi_2(p) \in_{\mathcal{U}} A, \{\pi_1(p)\}, \emptyset) \mid p \in R \}$
- $R \mapsto \bigcup \{ \{ \text{fib}(x, R) \} \mid x \in \{ \pi_1(p) \mid p \in R \} \}$

Derivable constructs:

- maps $\{e_1(x) \mid x \in e_2\}$
- at type-level, $\text{Bool} := \text{Set}(1)$
- basic predicates $=_T: T \times T \rightarrow \text{Bool}$, $\in_T: T \times \text{Set}(T) \rightarrow \text{Bool}$
- case analyses
- Δ_0 -separation $\{x \in e \mid \varphi(x)\}$

Proposition

NRC terms $e : T \rightarrow \text{Bool}$ correspond to Δ_0 formulas $\varphi(x^T)$

Limits to the expressiveness of NRC

For practical purposes, NRC is not too expressive

- NRC is *conservative* over idealized SQL i.e., for flat queries
- queries *polytime* computable (over finite inputs)

Consequences

- rules out $x \mapsto \mathcal{P}(x)$
- cannot represent function spaces with $\text{Set}(-)$

Consider $(x, y) \mapsto \text{tt}$

$$[T \rightarrow \text{Set}(U)] \not\approx [T \times U \rightarrow \text{Bool}]$$

$$[T \rightarrow \text{Set}(U)] \leftrightarrow [T \times U \rightarrow \text{Bool}]$$

(For the rest of the talk: no finiteness assumptions)

Extraction from Δ_0 specifications

Effective extraction

Recall that $\varphi(i, o)$ is an implicit definition when it is functional:

$$\varphi(i, o) \wedge \varphi(i, o') \implies o = o'$$

Extraction from Δ_0 implicit definitions

For every such $\varphi(i, o)$, there is a compatible NRC term $e(i)$

$$\varphi(i, o) \implies o = e(i)$$

Effective extraction

Recall that $\varphi(i, o)$ is an implicit definition when it is functional:

$$\varphi(i, o) \wedge \varphi(i, o') \implies o = o'$$

Extraction from Δ_0 implicit definitions

For every such $\varphi(i, o)$, there is a compatible NRC term $e(i)$

$$\varphi(i, o) \implies o = e(i)$$

$e(i)$ polytime computable from a **focused** proof

Effective extraction

Recall that $\varphi(i, o)$ is an implicit definition when it is functional:

$$\varphi(i, o) \wedge \varphi(i, o') \implies o = o'$$

Extraction from Δ_0 implicit definitions

For every such $\varphi(i, o)$, there is a compatible NRC term $e(i)$

$$\varphi(i, o) \implies o = e(i)$$

$e(i)$ polytime computable from a **focused** proof

- Extension of Beth definability for flat queries

$$\text{Set}(\mathcal{U}^k) \times \dots \times \text{Set}(\mathcal{U}^m) \rightarrow \text{Set}(\mathcal{U}^n)$$

Effective extraction

Recall that $\varphi(i, o)$ is an implicit definition when it is functional:

$$\varphi(i, o) \wedge \varphi(i, o') \implies o = o'$$

Extraction from Δ_0 implicit definitions

For every such $\varphi(i, o)$, there is a compatible NRC term $e(i)$

$$\varphi(i, o) \implies o = e(i)$$

$e(i)$ polytime computable from a **focused** proof

- Extension of Beth definability for flat queries
 $\text{Set}(\mathcal{U}^k) \times \dots \times \text{Set}(\mathcal{U}^m) \rightarrow \text{Set}(\mathcal{U}^n)$
- Remark: a non-effective proof would still yield an algorithm
 - Can be proven elementarily
 - Alternatively, this can be reduced to a Π_2^0 statement

Brief aside on focusing

A normal form for proofs refining cut-freeness (Andreoli 90s)

Rough idea

Decompose proofs by forcing saturations by certain rules in *positive* and *negative* phases. Roughly:

- Negative: apply invertible rules as much as possible
- Positive: focus on a single formula until it turns negative.

Brief aside on focusing

A normal form for proofs refining cut-freeness (Andreoli 90s)

Rough idea

Decompose proofs by forcing saturations by certain rules in *positive* and *negative* phases. Roughly:

- Negative: apply invertible rules as much as possible
- Positive: focus on a single formula until it turns negative.

Complexity-wise (to the best of my knowledge)

A cut-free proof can be turned into a focused cut-free proof in exponential time.

Toward a proof system for nested sets

Wlog, we restrict to the following syntax

$$t, u ::= x \mid (t, u) \mid \pi_1(t) \mid \pi_2(t) \mid ()$$

$$\varphi, \psi ::= t =_{\mathcal{U}} u \mid t \neq_{\mathcal{U}} u \mid \exists x \in_T t \varphi \mid \forall x \in_T t \varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi$$

Toward a proof system for nested sets

Wlog, we restrict to the following syntax

$$t, u ::= x \mid (t, u) \mid \pi_1(t) \mid \pi_2(t) \mid ()$$

$$\varphi, \psi ::= t =_{\mathcal{U}} u \mid t \neq_{\mathcal{U}} u \mid \exists x \in_T t \varphi \mid \forall x \in_T t \varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi$$

Macros

$$t \in_T u \quad := \quad \exists x \in u. t =_T u$$

$$t \subseteq_T u \quad := \quad \forall x \in t. x \in_T u$$

$$t =_{\text{Set}(T)} u \quad := \quad t \subseteq_T u \wedge u \subseteq_T t$$

$$t =_{T \times U} u \quad := \quad \pi_1(t) =_T \pi_1(u) \wedge \pi_2(t) =_U \pi_2(u)$$

$$t =_1 u \quad := \quad \top$$

Toward a proof system for nested sets

Wlog, we restrict to the following syntax

$$t, u ::= x \mid (t, u) \mid \pi_1(t) \mid \pi_2(t) \mid ()$$

$$\varphi, \psi ::= t =_{\mathcal{U}} u \mid t \neq_{\mathcal{U}} u \mid \exists x \in_T t \varphi \mid \forall x \in_T t \varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi$$

Macros

$$t \in_T u \quad := \quad \exists x \in u. t =_T u$$

$$t \subseteq_T u \quad := \quad \forall x \in t. x \in_T u$$

$$t =_{\text{Set}(T)} u \quad := \quad t \subseteq_T u \quad \wedge \quad u \subseteq_T t$$

$$t =_{T \times U} u \quad := \quad \pi_1(t) =_T \pi_1(u) \quad \wedge \quad \pi_2(t) =_U \pi_2(u)$$

$$t =_1 u \quad := \quad \top$$

- Bakes the axiom of extensionality in the definition of $=_T$
- **No further non-equational axioms**
- For the rest of the talk: we use sequent calculus

Formal proofs of functionality

Certificate that $\varphi(i, o)$ is an implicit definition: a derivation

$$\cdot; \varphi(i, o), \varphi(i, o') \vdash o = o'$$

$$\begin{array}{c}
 \text{AX} \frac{}{z \in o, x \in X, z \in x; z \in o' \vdash z \in o'} \text{ (7)} \\
 \Rightarrow\text{-L} \frac{}{z \in o, x \in X, z \in x; \chi(X, x, z), \chi(X, x, z) \Rightarrow z \in o' \vdash z \in o'} \text{ (6)} \\
 \forall\text{-L} \frac{}{z \in o, x \in X, z \in x; \chi(X, x, z), \forall a \in x (\chi(X, x, a) \Rightarrow a \in o') \vdash z \in o'} \text{ (5)} \\
 =\text{-SUBST} \frac{}{z \in o, x \in X, z' \in x; z =_U z', \chi(X, x, z), \forall a \in x (\chi(X, x, a) \Rightarrow a \in o') \vdash z \in o'} \\
 \exists\text{-L} \frac{}{z \in o, x \in X; z \in x, \chi(X, x, z), \forall a \in x (\chi(X, x, a) \Rightarrow a \in o') \vdash z \in o'} \\
 \wedge\text{-L} \frac{}{z \in o, x \in X; \psi(X, x, z), \forall a \in x (\chi(X, x, a) \Rightarrow a \in o') \vdash z \in o'} \\
 \forall\text{-L} \frac{}{z \in o, x \in X; \psi(X, x, z), \forall y \in X \forall a \in y (\chi(X, y, a) \Rightarrow a \in o') \vdash z \in o'} \text{ (4)} \\
 \wedge\text{-L} \frac{}{z \in o, x \in X; \psi(X, x, z), \Sigma(X, o') \vdash z \in o'} \\
 \exists\text{-L} \frac{}{z \in o; \exists x \in X \psi(X, x, z), \Sigma(X, o') \vdash z \in o'} \text{ (3)} \\
 \forall\text{-L} \frac{}{z \in o; \forall a \in o \exists x \in X \psi(X, x, a), \Sigma(X, o') \vdash z \in o'} \\
 \wedge\text{-L} \frac{}{z \in o; \Sigma(X, o), \Sigma(X, o') \vdash z \in o'} \\
 \subseteq\text{-R} \frac{}{\cdot; \Sigma(X, o), \Sigma(X, o') \vdash o \subseteq o'} \text{ (2)} \\
 =\text{-Set-R} \frac{}{\cdot; \Sigma(X, o), \Sigma(X, o') \vdash o = o'} \text{ (1)}
 \end{array}$$

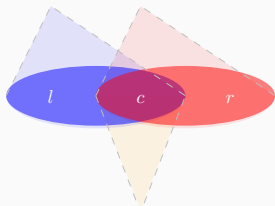
Problem: what inductive invariant?

The intuitionistic case

Input: $\Gamma(c, l), \Delta(c, r) \vdash \psi$, cut-free

Output: a NRC term $e(c)$ s.t.

$\Gamma(c, l), \Delta(c, r) \vdash l \subseteq r$



$\rightsquigarrow \exists e. l \subseteq e(c) \subseteq r$

Key lemma

- if ψ is $l = r$, then

$$\Gamma, \Delta \models l = e \wedge r = e$$

- if ψ is $l \subseteq r$, then

$$\Gamma, \Delta \models l \subseteq e \wedge e \subseteq r$$

- if ψ is $l \in r$, then

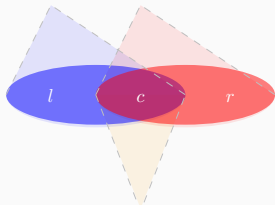
$$\Gamma, \Delta \models l \in e$$

The intuitionistic case

Input: $\Gamma(c, l), \Delta(c, r) \vdash \psi$, cut-free

Output: a NRC term $e(c)$ s.t.

$\Gamma(c, l), \Delta(c, r) \vdash l \subseteq r$



$\rightsquigarrow \exists e. l \subseteq e(c) \subseteq r$

Key lemma

- if ψ is $l = r$, then

$$\Gamma, \Delta \models l = e \wedge r = e$$

- if ψ is $l \subseteq r$, then

$$\Gamma, \Delta \models l \subseteq e \wedge e \subseteq r$$

- if ψ is $l \in r$, then

$$\Gamma, \Delta \models l \in e$$

Why this is easy: single RHS formula, subformula of $=_T$.

Proof idea

Let us look at the key step that involves interpolation.

$$\frac{\Gamma, z \in_T l, \Delta \vdash z \in_T r}{\Gamma, \Delta \vdash l \subseteq_T r} \quad \mapsto \quad \frac{\Gamma, z \in_T l, \Delta \models z \in_T e^{\text{IH}}}{\Gamma, \Delta \models l \subseteq_T e \wedge e \subseteq_T r}$$

Assuming a Δ_0 interpolant $\theta(z)$ such that

$$\Gamma \wedge z \in_T l \models \theta(z) \quad \text{and} \quad \theta(z) \models \Delta \Rightarrow z \in_T r$$

Proof idea

Let us look at the key step that involves interpolation.

$$\frac{\Gamma, z \in_T l, \Delta \vdash z \in_T r}{\Gamma, \Delta \vdash l \subseteq_T r} \quad \mapsto \quad \frac{\Gamma, z \in_T l, \Delta \models z \in_T e^{\text{IH}}}{\Gamma, \Delta \models l \subseteq_T e \wedge e \subseteq_T r}$$

Assuming a Δ_0 interpolant $\theta(z)$ such that

$$\Gamma \wedge z \in_T l \models \theta(z) \quad \text{and} \quad \theta(z) \models \Delta \Rightarrow z \in_T r$$

we can set $e := \{x \in e^{\text{IH}} \mid \theta(z)\}$

Proof idea

Let us look at the key step that involves interpolation.

$$\frac{\Gamma, z \in_T l, \Delta \vdash z \in_T r}{\Gamma, \Delta \vdash l \subseteq_T r} \quad \mapsto \quad \frac{\Gamma, z \in_T l, \Delta \models z \in_T e^{\text{IH}}}{\Gamma, \Delta \models l \subseteq_T e \wedge e \subseteq_T r}$$

Assuming a Δ_0 interpolant $\theta(z)$ such that

$$\Gamma \wedge z \in_T l \models \theta(z) \quad \text{and} \quad \theta(z) \models \Delta \Rightarrow z \in_T r$$

we can set $e := \{x \in e^{\text{IH}} \mid \theta(z)\}$

Other key cases in a hurry

\cup for \forall -L, \cup for \forall -L, $\{-\}$ for \exists -R

The classical case

Difficulty: with classical logic, we can contract the goal formula

Reduction to the \in case

From a proof $\Gamma, \Delta \vdash \exists x \in^+ r. l = x$, compute a NRC term e such that $l \in e$

- Weaker than definability: **set** of possible solutions

The classical case

Difficulty: with classical logic, we can contract the goal formula

Reduction to the \in case

From a proof $\Gamma, \Delta \vdash \exists x \in^+ r. l = x$, compute a NRC term e such that $l \in e$

- Weaker than definability: **set** of possible solutions
- Proven by outer induction on the type structure

The classical case

Difficulty: with classical logic, we can contract the goal formula

Reduction to the \in case

From a proof $\Gamma, \Delta \vdash \exists x \in^+ r. l = x$, compute a NRC term e such that $l \in e$

- Weaker than definability: **set** of possible solutions
- Proven by outer induction on the type structure
- Set-case by modifying the input proof to

$$\exists x \in_{\text{Set}(T)} . \forall c \in a. c \in l \Leftrightarrow c \in x \quad (\text{fresh } a)$$

and then applying interpolation

Parameterized definability, FO(=) version

Working with multisorted FO, no function symbols, signatures L, R and $C \subseteq L \cap R$. (and C “has a sort with ≥ 2 elements”)

Theorem

If we have a **focused** derivation of

$$\Gamma, \Delta \vdash \exists r. \forall c. \lambda(c) \iff \rho(r, c)$$

then we can compute in linear time a $\psi(\vec{p}, c)$ over C such that

$$\Gamma, \Delta \models \exists \vec{p}. \forall c. \lambda(c) \iff \psi(\vec{p}, c)$$

Call ψ a **parameterized definition** for λ .

Reminder of what is the theorem

$$\frac{\Gamma, \Delta \vdash \exists r. \forall c. \lambda(c) \iff \rho(r, c)}{\psi \text{ such that } \Gamma, \Delta \models \exists \vec{p}. \forall c. \lambda(c) \iff \psi(\vec{p}, c)}$$

(Not useful) ways to instantiate the hypotheses:

- Trivial $\exists r$.: same premise as in Beth definability
- One generalization $\varphi(R)$ defines finitely many R s, then we have a parameterized definition of R
 - Throw in a finite sort n to index finitely many distinct R_i with $\varphi(R_i)$ for all i and $\varphi(R)$
 - (Kueker already gave a proof-theoretic method)

Reminder of what is the theorem

$$\frac{\Gamma, \Delta \vdash \exists r. \forall c. \lambda(c) \iff \rho(r, c)}{\psi \text{ such that } \Gamma, \Delta \models \exists \vec{p}. \forall c. \lambda(c) \iff \psi(\vec{p}, c)}$$

An instantiation (of questionable utility?):

- Fix a FO formula $\varphi(P)$ over a signature $\Sigma \uplus \{P\}$
- Call CA the theory of comprehension over Σ
- If we have

$$\varphi(P), \text{CA} \vdash \exists X. \forall x. P(x) \iff x \in X$$

then P is parametrically FO-definable over Σ .

Reminder of what is the theorem

$$\frac{\Gamma, \Delta \vdash \exists r. \forall c. \lambda(c) \iff \rho(r, c)}{\psi \text{ such that } \Gamma, \Delta \models \exists \vec{p}. \forall c. \lambda(c) \iff \psi(\vec{p}, c)}$$

An instantiation (of questionable utility?):

- Fix a FO formula $\varphi(P)$ over a signature $\Sigma \uplus \{P\}$
- Call CA the theory of comprehension over Σ
- If we have

$$\varphi(P), \text{CA} \vdash \exists X. \forall x. P(x) \iff x \in X$$

then P is parametrically FO-definable over Σ .

- Can be derived from a theorem of Chang & Makkai

Parametrized definability from countably many options

Extends the cardinality condition of Kueker as far as possible.

Theorem (Chang-Makkai, 64)

Let \mathcal{T} be a theory over $\Sigma \uplus \{P\}$. TFAE:

- P is parametrically FO-definable over Σ (in \mathcal{T})
- for every model (M, \dots, P) of \mathcal{T} there are at most $|M|$ many valid alternatives for P

(i.e. $P' \subseteq M$ such that (M, \dots, P') is also a model of \mathcal{T})

Parametrized definability from countably many options

Extends the cardinality condition of Kueker as far as possible.

Theorem (Chang-Makkai, 64)

Let \mathcal{T} be a theory over $\Sigma \uplus \{P\}$. TFAE:

- P is parametrically FO-definable over Σ (in \mathcal{T})
- for every model (M, \dots, P) of \mathcal{T} there are at most $|M|$ many valid alternatives for P
(i.e. $P' \subseteq M$ such that (M, \dots, P') is also a model of \mathcal{T})

Proof: via saturated models

(equivalent condition: there exists a saturated model...)

Parametrized definability from countably many options

Extends the cardinality condition of Kueker as far as possible.

Theorem (Chang-Makkai, 64)

Let \mathcal{T} be a theory over $\Sigma \uplus \{P\}$. TFAE:

- P is parametrically FO-definable over Σ (in \mathcal{T})
- for every model (M, \dots, P) of \mathcal{T} there are at most $|M|$ many valid alternatives for P

(i.e. $P' \subseteq M$ such that (M, \dots, P') is also a model of \mathcal{T})

Proof: via saturated models

(equivalent condition: there exists a saturated model...)

Question

Can we give a satisfactory effective counterpart to this?

Back to our nested relations business

Adaptation in our setting with Δ_0 set-theoretic formulas

(sadly not derived as a corollary)

Theorem

If we have a **focused** derivation

$$\Gamma(\vec{i}, \dots), \Delta(\vec{i}, r, \dots) \vdash \exists r' \in r. \forall x \in a. \lambda(x, \dots) \iff \rho(r', \dots, x)$$

then we have in linear time a NRC term $E(\vec{i})$ such that

$$\Gamma(\vec{i}, \dots), \Delta(\vec{i}, r, \dots) \models a \cap \lambda \in E(\vec{i})$$

Proof: induction; we need also to compute a $\theta(\vec{i})$ such that

$$\Delta(\vec{i}, r, \dots) \models \theta(\vec{i}) \quad \text{and} \quad \Gamma(\vec{i}, \dots), \theta(\vec{i}) \models a \cap \lambda \in E(\vec{i})$$

Key step: existential rule introducing the “main” formula

With $\mathcal{G} = \exists r' \in^+ r. \forall z \in c. \lambda(z) \iff \rho(z, r')$

$$\frac{\frac{\vee \frac{\Theta_L, \Theta_R, x \in c \vdash \Delta_L, \Delta_R, \neg \rho(x, w), \lambda(x), \mathcal{G}}{\Theta_L, \Theta_R, x \in c \vdash \Delta_L, \Delta_R, \rho(x, w) \Rightarrow \lambda(x), \mathcal{G}} \quad \vdots}{\wedge \frac{\Theta_L, \Theta_R, x \in c \vdash \Delta_L, \Delta_R, \lambda(x) \Leftrightarrow \rho(x, w), \mathcal{G}}{\vee \frac{\Theta_L, \Theta_R \vdash \Delta_L, \Delta_R, \forall z \in c. (\lambda(z) \Leftrightarrow \rho(z, w)), \mathcal{G}}{\exists^+ \frac{\Theta_L, \Theta_R \vdash \Delta_L, \Delta_R, \mathcal{G}}{\Theta_L, \Theta_R \vdash \Delta_L, \Delta_R, \mathcal{G}}}}}}{\Theta_L, \Theta_R \vdash \Delta_L, \Delta_R, \mathcal{G}}$$

- Shape around the root of the tree guaranteed by focusing
- Applying the induction hypothesis we have

$$\begin{aligned} \Theta_L, x \in c \models \lambda(x), \Delta_L, \theta_1^{\text{IH}} \vee \Lambda \in E_1^{\text{IH}} \\ \Theta_R \models \neg \rho(x, w), \Delta_R, \neg \theta_1^{\text{IH}} \\ \dots \end{aligned}$$

- Take $\theta := \exists x \in c. \theta_1^{\text{IH}} \wedge \theta_2^{\text{IH}}$ and

$$E := \left\{ \left\{ x \in c \mid \theta_2^{\text{IH}} \right\} \right\} \cup \bigcup \left\{ E_1^{\text{IH}} \cup E_2^{\text{IH}} \mid x \in c \right\}$$

What have we not learned?

Extraction from Δ_0 implicit definitions

For every such $\varphi(i, o)$, there is a compatible NRC term $e(i)$

$$\varphi(i, o) \implies o = e(i)$$

What have we not learned?

Extraction from Δ_0 implicit definitions

For every such $\varphi(i, o)$, there is a compatible NRC term $e(i)$

$$\varphi(i, o) \implies o = e(i)$$

$e(i)$ polytime computable from a **focused** proof

- The intuitionistic case is much easier

Conservativity for implicit definitions

If $\phi(i, o)$ is functional, then there is a formula $\chi(\vec{x})$ such that the conjoined formula $\phi^{\neg\neg}(i, o) \wedge \forall \vec{x}. \chi(\vec{x}) \vee \neg\chi(\vec{x})$ can be proved to be functional in intuitionistic logic.

- (but finding χ has no reason being easy!)
- W/o the complexity bound: easier proof via model theory

Model-theory route

Nested collections can be regarded as multi-sorted structures

An object X of sort $\text{Set}(\mathfrak{U} \times \text{Set}(\mathfrak{U}))$

Sorts: $\mathfrak{U}, \text{Set}(\mathfrak{U}), \mathfrak{U} \times \text{Set}(\mathfrak{U})$

Function symbols: $\pi_1, \pi_2, \langle -, - \rangle$

Relation symbol: $\in_{\mathfrak{U}}$

Semantics: subobjects of X

Model-theory route

Nested collections can be regarded as multi-sorted structures

An object X of sort $\text{Set}(\mathfrak{U} \times \text{Set}(\mathfrak{U}))$

Sorts: $\mathfrak{U}, \text{Set}(\mathfrak{U}), \mathfrak{U} \times \text{Set}(\mathfrak{U})$

Function symbols: $\pi_1, \pi_2, \langle -, - \rangle$

Relation symbol: $\in_{\mathfrak{U}}$

Semantics: subobjects of X

Interpretations: maps Models \rightarrow Models via FO formulas

Model-theory route

Nested collections can be regarded as multi-sorted structures

An object X of sort $\text{Set}(\mathfrak{U} \times \text{Set}(\mathfrak{U}))$

Sorts: $\mathfrak{U}, \text{Set}(\mathfrak{U}), \mathfrak{U} \times \text{Set}(\mathfrak{U})$

Function symbols: $\pi_1, \pi_2, \langle -, - \rangle$

Relation symbol: $\in_{\mathfrak{U}}$

Semantics: subobjects of X

Interpretations: maps Models \rightarrow Models via FO formulas

Can express

- product, disjoint union of structures $\mathfrak{M}, \mathfrak{N} \mapsto \mathfrak{M} \times \mathfrak{N}, \mathfrak{M} + \mathfrak{N}$
- definable substructures and quotients

Model-theory route

Nested collections can be regarded as multi-sorted structures

An object X of sort $\text{Set}(\mathfrak{U} \times \text{Set}(\mathfrak{U}))$

Sorts: $\mathfrak{U}, \text{Set}(\mathfrak{U}), \mathfrak{U} \times \text{Set}(\mathfrak{U})$

Function symbols: $\pi_1, \pi_2, \langle -, - \rangle$

Relation symbol: $\in_{\mathfrak{U}}$

Semantics: subobjects of X

Interpretations: maps Models \rightarrow Models via FO formulas

Can express

- product, disjoint union of structures $\mathfrak{M}, \mathfrak{N} \mapsto \mathfrak{M} \times \mathfrak{N}, \mathfrak{M} + \mathfrak{N}$
- definable substructures and quotients

NRC and interpretations

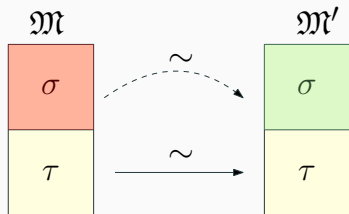
For structures corresponding to nested collections,
NRC and Δ_0 -interpretations coincide

The key model-theoretic lemma

Consider models of a theory \mathcal{T} over two sorts τ and σ

Multi-sorted implicit definability

σ is **implicitly definable from τ** when every $f : \mathfrak{M}|_{\tau} \cong \mathfrak{M}'|_{\tau}$ has a unique extension $\hat{f} : \mathfrak{M} \cong \mathfrak{M}'$

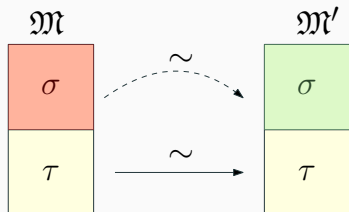


The key model-theoretic lemma

Consider models of a theory \mathcal{T} over two sorts τ and σ

Multi-sorted implicit definability

σ is **implicitly definable from τ** when every $f : \mathfrak{M}|_{\tau} \cong \mathfrak{M}'|_{\tau}$ has a unique extension $\hat{f} : \mathfrak{M} \cong \mathfrak{M}'$



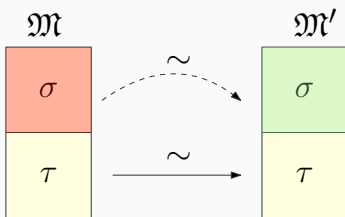
Lemma (\sim relative rigid categoricity, dcl)

σ implicitly definable from $\tau \Rightarrow \exists$ interpretation of \mathcal{T} into $\mathcal{T}|_{\tau}$.

The key model-theoretic lemma : example

Multi-sorted implicit definability

σ is **implicitly definable from** τ when every $f : \mathfrak{M}|_{\tau} \cong \mathfrak{M}'|_{\tau}$ has a unique extension $\hat{f} : \mathfrak{M} \cong \mathfrak{M}'$



Example: \mathcal{T} says that τ is a real closed field and that σ is a dimension two field extension with a distinguished i with $i^2 = -1$.

Lemma (\sim relative rigid categoricity, dcl)

σ implicitly definable from $\tau \Rightarrow \exists$ interpretation of \mathcal{T} into $\mathcal{T}|_{\tau}$.

Example (cont): exercise :)

Coordinizability

Lemma (\sim relative rigid categoricity, dcl)

σ implicitly definable from $\tau \Rightarrow \exists$ interpretation of \mathcal{T} into $\mathcal{T}|_{\tau}$.

Proof idea: use Beth definability after computing coordinates.

Coordinizability (\sim Gaifman's coordinizability)

\exists a FO-definable partial surjection $\tau^n \rightarrow \sigma$.

That is, there exists $\chi(\vec{p}^{\tau}, x^{\sigma})$ such that

$$\forall x. \exists \vec{p}. \chi(\vec{p}, x)$$

$$\forall \vec{p} x y. \chi(\vec{p}, x) \wedge \chi(\vec{p}, y) \implies x = y$$

Coordinizability

Lemma (\sim relative rigid categoricity, dcl)

σ implicitly definable from $\tau \Rightarrow \exists$ interpretation of \mathcal{T} into $\mathcal{T}|_{\tau}$.

Proof idea: use Beth definability after computing coordinates.

Coordinizability (\sim Gaifman's coordinizability)

\exists a FO-definable partial surjection $\tau^n \dashrightarrow \sigma$.

That is, there exists $\chi(\vec{p}^{\tau}, x^{\sigma})$ such that

$$\forall x. \exists \vec{p}. \chi(\vec{p}, x)$$

$$\forall \vec{p} x y. \chi(\vec{p}, x) \wedge \chi(\vec{p}, y) \implies x = y$$

From implicit definability: via the omitting type theorem

Coordinizability

Lemma (\sim relative rigid categoricity, dcl)

σ implicitly definable from $\tau \Rightarrow \exists$ interpretation of \mathcal{T} into $\mathcal{T}|_{\tau}$.

Proof idea: use Beth definability after computing coordinates.

Coordinizability (\sim Gaifman's coordinizability)

\exists a FO-definable partial surjection $\tau^n \rightarrow \sigma$.

That is, there exists $\chi(\vec{p}^{\tau}, x^{\sigma})$ such that

$$\forall x. \exists \vec{p}. \chi(\vec{p}, x)$$

$$\forall \vec{p} x y. \chi(\vec{p}, x) \wedge \chi(\vec{p}, y) \implies x = y$$

From implicit definability: via the omitting type theorem

Question

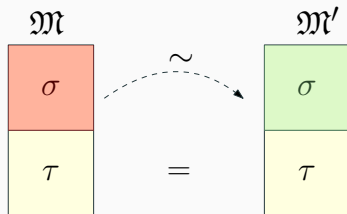
Can we make coordinizability effective/efficient?

Effective implicit definability for sorts (1/2)

Problem #1: witness of implicit definability?

Multi-sorted implicit definability (mild alteration)

σ is **implicitly definable from** τ if whenever $\mathfrak{M}|_{\tau} = \mathfrak{M}'|_{\tau}$, there is a unique isomorphism $f : \mathfrak{M} \cong \mathfrak{M}'$ with $f|_{\tau} = \text{id}$.

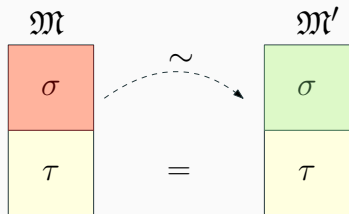


Effective implicit definability for sorts (1/2)

Problem #1: witness of implicit definability?

Multi-sorted implicit definability (mild alteration)

σ is **implicitly definable from** τ if whenever $\mathfrak{M}|_{\tau} = \mathfrak{M}'|_{\tau}$, there is a unique isomorphism $f : \mathfrak{M} \cong \mathfrak{M}'$ with $f|_{\tau} = \text{id}$.



Idea to reduce to provability

Consider the theory $\mathcal{T} \cup \mathcal{T}'$ over τ, σ, σ' talking about the join of a pair of models $\mathfrak{M} \cup \mathfrak{M}'$

Effective implicit definability for sorts (2/2)

Lemma

Implicit definability is equivalent to the existence of a FO formula $\psi(x^\tau, y^{\tau'})$ such that

$\mathcal{T} \cup \mathcal{T}' \vdash \psi$ is an embedding extending the identity

(\Rightarrow) use Beth definability to compute ψ !

(\Leftarrow) FO definable embedding = isomorphism (crucial thing $\tau \neq \tau'$)
(requires comprehension for the trivial part (ACA'₀))

We have a nice Π_2^0 statement

Effective implicit definable \Rightarrow coordinizable, effectively

Question: how efficient can we make that?

Some very tentative results

For intuitionistic logic

From a cut-free proof of totality of $\varphi(x^\tau, y^{\tau'})$ in LJ

$$\vdash \forall x^\tau \exists y^{\tau'}. \varphi(x, y)$$

we can compute in polynomial time coordinates for τ .

Proof idea: induction until we hit the \exists -R rule

Restricted to Σ_1 formulas

From a cut-free proof of totality of $\varphi(x^\tau, y^{\tau'})$, if φ is Σ_1 and functional, we can compute coordinates for τ .

Proof: via Herbrand's theorem

(and some fiddly steps to get rid of function symbols)

Why these results are no good

For intuitionistic logic

From a cut-free proof of totality of $\varphi(x^\tau, y^{\tau'})$ in LJ we can compute in polynomial time coordinates for τ .

While all φ can be “made intuitionistically total”, I don’t have a proof of this that does not presuppose classical coordinizability.

Restricted to Σ_1 formulas

From a cut-free proof of totality of $\varphi(x^\tau, y^{\tau'})$, if φ is Σ_1 and functional, we can compute coordinates for τ .

For NRC, φ exhausts the quantifier hierarchy.

Why that won’t generalize too well

There exists a silly Π_2 φ functional and total (but which is not an embedding) such that τ is not coordinizable.

Takeaways/further work that could be done

Nice theorem about NRC and implicit definitions but:

- Proof-theoretic take on definability results?
 - (one excuse: complexity)
- Chang-Makkai
 - (\sim countably many predicates)
- Definable closures
 - (what I just discussed)
- Algebraic closures

Takeaways/further work that could be done

Nice theorem about NRC and implicit definitions but:

- Proof-theoretic take on definability results?
 - (one excuse: complexity)
 - Chang-Makkai (\sim countably many predicates)
 - Definable closures (what I just discussed)
 - Algebraic closures
- NRC with multisets/list operators
 - Basic unanswered question: specification logic?
 - Rough idea for the model-theoretic route: theory of families indexed by $\mathbf{FinSet}/\mathbf{FinOrd}$, and look at coarser equivalences

Takeaways/further work that could be done

Nice theorem about NRC and implicit definitions but:

- Proof-theoretic take on definability results?
 - (one excuse: complexity)
 - Chang-Makkai (\sim countably many predicates)
 - Definable closures (what I just discussed)
 - Algebraic closures
- NRC with multisets/list operators
 - Basic unanswered question: specification logic?
 - Rough idea for the model-theoretic route: theory of families indexed by $\mathbf{FinSet}/\mathbf{FinOrd}$, and look at coarser equivalences
- Implementation.

Takeaways/further work that could be done

Nice theorem about NRC and implicit definitions but:

- Proof-theoretic take on definability results?
 - (one excuse: complexity)
 - Chang-Makkai (\sim countably many predicates)
 - Definable closures (what I just discussed)
 - Algebraic closures
- NRC with multisets/list operators
 - Basic unanswered question: specification logic?
 - Rough idea for the model-theoretic route: theory of families indexed by $\mathbf{FinSet}/\mathbf{FinOrd}$, and look at coarser equivalences
- Implementation.

Thanks for listening! :)