Interpolation Properties for Array Theories: Positive and Negative Results

Silvio Ghilardi[1]

Dipartimento di Matematica
Università degli Studi di Milano, Italy

Workshop on Craig Interpolation and Beth Definability

Amsterdam, April 22, 2024

[1] Based on joint work with various (recent and less recent) collaborators: R. Bruttomesso, S. Ranise, A. Gianola, D. Calvanese, M. Montali, D. Kapur, C. Naso

Outline

Motivation

A first-order theory T has quantifier-free interpolation iff for every quantifier free formulae ϕ, ψ such that $T \vdash \phi \to \psi$, there exists a quantifier free formula θ such that:

(i) $T \vdash \phi \to \theta$;

(ii) $T \vdash \theta \to \psi$;

(iii) only variables occurring both in ψ and in ϕ occur in θ.

Quantifier-free interpolants are commonly used in formal verification during abstraction-refinement cycles (since [McMillan CAV 03], [McMillan TACAS 04], ...).

Motivation

- In infinite-state model checking, the search of formulae is **not** *finitely bounded*.

Motivation

- In infinite-state model checking, the search of formulae is **not** *finitely bounded*.
- Analyzing spurious error traces:

$$In(\underline{x}_0) \wedge Tr(\underline{x}_0, \underline{x}_1) \wedge \cdots \wedge Tr(\underline{x}_{n-1}, \underline{x}_n) \wedge U(\underline{x}_n)$$

one can produce (via interpolation) formulae ϕ such that

$$In(\underline{x}_0) \wedge \bigwedge_{j=0}^{i} Tr(\underline{x}_{j-1}, \underline{x}_j) \models \phi(\underline{x}_i) \quad \text{and} \quad \phi(\underline{x}_i) \wedge \bigwedge_{j=i+1}^{n} Tr(\underline{x}_{j-1}, \underline{x}_j) \wedge U(\underline{x}_n) \models \bot.$$

Motivation

- In infinite-state model checking, the search of formulae is **not** *finitely bounded*.
- Analyzing spurious error traces:

$$In(\underline{x}_0) \land Tr(\underline{x}_0, \underline{x}_1) \land \cdots \land Tr(\underline{x}_{n-1}, \underline{x}_n) \land U(\underline{x}_n)$$

  one can produce (via interpolation) formulae ϕ such that

$$In(\underline{x}_0) \land \bigwedge_{j=0}^{i} Tr(\underline{x}_{j-1}, \underline{x}_j) \models \phi(\underline{x}_i) \text{ and } \phi(\underline{x}_i) \land \bigwedge_{j=i+1}^{n} Tr(\underline{x}_{j-1}, \underline{x}_j) \land U(\underline{x}_n) \models \bot.$$

- These formulae (and the atoms they contain) can contribute to the **refinement** of the candidate loop invariant guaranteeing safety.

General Interpolation Property

In verification theory, people uses the following stronger property for a theory T:

General Interpolation Property

In verification theory, people uses the following stronger property for a theory T:

Definition

Let T be a theory in a signature Σ; we say that T has the general quantifier-free interpolation property iff for every signature Σ' (disjoint from Σ) and for every ground $\Sigma \cup \Sigma'$-formulæ ϕ, ψ such that $T \vdash \phi \to \psi$ is T-unsatisfiable, there is a ground formula θ such that:

(i) $T \vdash \phi \to \theta$;

(ii) $T \vdash \theta \to \psi$;

(iii) all predicate, constants and function symbols from Σ' occurring in θ occur also in ϕ and in ψ.

Uniform Interpolation Property

A considerable strengthening of plain interpolation is uniform interpolation:

Uniform Interpolation Property

A considerable strengthening of plain interpolation is uniform interpolation:

Definition

We say that a theory T has uniform quantifier-free interpolation iff every tuple of variables $\underline{x}$ and every quantifier-free formula ϕ there is a quantifier-free formula θ not containing the $\underline{x}$ such that:

(i) $T \vdash \phi \to \theta$;

(ii) for every quantifier-free formula ψ not containing the $\underline{x}$

$$T \vdash \phi \to \psi \quad \Rightarrow \quad T \vdash \theta \to \psi \quad .$$

Semantic Reformulations

Theorem

Let T be a universal theory. Then

(i) *T has quantifier-free interpolation iff T has the amalgamation property [B 75];*

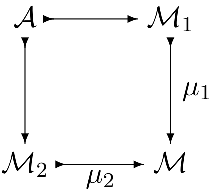
(ii) *T has the general quantifier-free interpolation iff T has the strong amalgamation property [BGR 14];*

(iii) *T has the uniform interpolation property iff T has a model completion [M 95, CGGMR 20].*

Amalgamation

Definition

A universal theory T has the amalgamation property iff whenever we are given models $\mathcal{M}_1$ and $\mathcal{M}_2$ of T and a common submodel $\mathcal{A}$ of them, there exists a further model $\mathcal{M}$ of T endowed with embeddings $\mu_1 : \mathcal{M}_1 \longrightarrow \mathcal{M}$ and $\mu_2 : \mathcal{M}_2 \longrightarrow \mathcal{M}$ whose restrictions to $|\mathcal{A}|$ coincide. The amalgamation property is strong iff in addition we require that $\mu_1(a_1) = \mu_2(a_2)$ implies that $a_1 = a_2 \in \mathcal{A}$.



Equality Interpolating Property

Definition

A theory T is equality interpolating [YM 05, BGR 14] iff it has the quantifier-free interpolation property and satisfies the following condition:

- for every quintuple $\underline{x}, \underline{y}_1, \underline{z}_1, \underline{y}_2, \underline{z}_2$ of tuples of variables and pair of quantifier-free formulae $\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1)$ and $\delta_2(\underline{x}, \underline{z}_2, \underline{y}_2)$ such that

$$\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T \underline{y}_1 \cap \underline{y}_2 \neq \emptyset \qquad (1)$$

  there exists a tuple $\underline{v}(\underline{x})$ of terms (called interpolating terms) such that

$$\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T (\underline{y}_1 \cup \underline{y}_2) \cap \underline{v} \neq \emptyset \ . \qquad (2)$$

Equality Interpolating Property

Theorem (BGR 14)

A universal theory T has the strong amalgamation property (i.e. the general interpolation property) iff it is equality interpolating. Equality interpolating is a modular property (under signature disjointness and stably-infiniteness assumptions).

Recall that T is stably infinite iff every model of T embeds into an infinite model (this is equivalent, via compactness, to the standard definition).

Equality Interpolating Property

Interpolating terms play an essential role in combined interpolation algorithms (see below).

Example

$\mathcal{EUF}$ is equality interpolating: interpolating terms can be computed by ground Knuth-Bendix completion (giving higher precedence to symbols to be eliminated).

Example

Universal Theories with QE (like linear real/integer arithmetics, under careful choice of the language) are equality interpolating: interpolating terms come from 'testing points' lemmas.

Equality Interpolating Property

Theorem (BGR 14)

Let T be a universal theory admitting quantifier-free interpolation and Σ be a signature disjoint from the signature of T containing at least a unary predicate symbol. Then, $T \cup EUF(\Sigma)$ has quantifier-free interpolation iff T has the strong amalgamation property.

Equality Interpolating Property

Theorem (BGR 14)

Let T be a universal theory admitting quantifier-free interpolation and Σ be a signature disjoint from the signature of T containing at least a unary predicate symbol. Then, $T \cup EUF(\Sigma)$ has quantifier-free interpolation iff T has the strong amalgamation property.

Here you are the relevant modularity result:

Equality Interpolating Property

Theorem (BGR 14)

Let T be a universal theory admitting quantifier-free interpolation and Σ be a signature disjoint from the signature of T containing at least a unary predicate symbol. Then, $T \cup EUF(\Sigma)$ has quantifier-free interpolation iff T has the strong amalgamation property.

Here you are the relevant modularity result:

Theorem (BGR 14)

Let T_1 and T_2 be two universal, stably infinite theories over disjoint signatures Σ_1 and Σ_2. If both T_1 and T_2 have the strong amalgamation property, then so does $T_1 \cup T_2$. In particular, $T_1 \cup T_2$ admits quantifier-free interpolation.

Equality Interpolating Property

Equality interpolating plays a role also for combined uniform interpolation, but only in presence of convexity:

Equality Interpolating Property

Equality interpolating plays a role also for combined uniform interpolation, but only in presence of convexity:

Theorem (CGGMR IJCAR '12)

Let T_1 and T_2 be two universal, stably infinite, strongly amalgamating convex theories over disjoint signatures Σ_1 and Σ_2. If both T_1 and T_2 have uniform interpolation, then so does $T_1 \cup T_2$.

Equality Interpolating Property

Equality interpolating plays a role also for combined uniform interpolation, but only in presence of convexity:

Theorem (CGGMR IJCAR '12)

Let T_1 and T_2 be two universal, stably infinite, strongly amalgamating convex theories over disjoint signatures Σ_1 and Σ_2. If both T_1 and T_2 have uniform interpolation, then so does $T_1 \cup T_2$.

Recall that a theory T is said to be *convex* iff every finite set of literals entailing (modulo T) a disjunction of $n > 0$ equalities entails one of them.

Outline

The theory $\mathcal{AR}_{\text{ext}}$ of arrays with extensionality

This is an important theory in verification:

- we have three sorts INDEX, ELEM, ARRAY;

The theory $\mathcal{AR}_{\mathrm{ext}}$ of arrays with extensionality

This is an important theory in verification:

- we have three sorts INDEX, ELEM, ARRAY;
- besides equality, we have function symbols

$$rd : \mathtt{ARRAY} \times \mathtt{INDEX} \longrightarrow \mathtt{ELEM},$$

$$wr : \mathtt{ARRAY} \times \mathtt{INDEX} \times \mathtt{ELEM} \longrightarrow \mathtt{ARRAY}$$

The theory $\mathcal{AR}_{\mathrm{ext}}$ of arrays with extensionality

This is an important theory in verification:

- we have three sorts INDEX, ELEM, ARRAY;
- besides equality, we have function symbols

$$rd : \mathtt{ARRAY} \times \mathtt{INDEX} \longrightarrow \mathtt{ELEM},$$
$$wr : \mathtt{ARRAY} \times \mathtt{INDEX} \times \mathtt{ELEM} \longrightarrow \mathtt{ARRAY}$$

- as axioms, we have

$$\forall y, i, e. \qquad rd(wr(y, i, e), i) = e \tag{3}$$
$$\forall y, i, j, e. \qquad i \neq j \to rd(wr(y, i, e), j) = rd(y, j) \tag{4}$$
$$\forall x, y. \qquad x \neq y \to (\exists i.\ rd(x, i) \neq rd(y, i)) \tag{5}$$

The theory $\mathcal{AR}_{\mathrm{ext}}$ of arrays with extensionality

Unfortunately, $\mathcal{AR}_{\mathrm{ext}}$ does not have interpolation, witness the following well-known counterexample (due to Ranjit Jhala).

The theory $\mathcal{AR}_{\text{ext}}$ of arrays with extensionality

Unfortunately, $\mathcal{AR}_{\text{ext}}$ does not have interpolation, witness the following well-known counterexample (due to Ranjit Jhala).

$$
\begin{aligned}
A :=& \quad \{a = wr(b, i, e)\} \\
B :=& \quad \{rd(a, j_1) \neq rd(b, j_1), rd(a, j_2) \neq rd(b, j_2), j_1 \neq j_2\}
\end{aligned}
$$

The theory $\mathcal{AR}_{\rm ext}$ of arrays with extensionality

Unfortunately, $\mathcal{AR}_{\rm ext}$ does not have interpolation, witness the following well-known counterexample (due to Ranjit Jhala).

$$\begin{array}{rl} A := & \{a = wr(b, i, e)\} \\ B := & \{rd(a, j_1) \neq rd(b, j_1), rd(a, j_2) \neq rd(b, j_2), j_1 \neq j_2\} \end{array}$$

Take ψ, ϕ to be the conjunctions of the literals from A, B, respectively. Then $\psi \wedge \phi$ is $\mathcal{AR}_{\rm ext}$-unsatisfiable, but no quantifier-free interpolant exists (notice that it should mention only a, b).

The theory $\mathcal{AX}_{\texttt{diff}}$ of arrays with diff

Since $\mathcal{AR}_{\mathrm{ext}}$ does not have quantifier-free interpolants, we consider the following variant, which we call $\mathcal{AX}_{\texttt{diff}}$. We add a further symbol in the signature

$$\texttt{diff} : \texttt{ARRAY} \times \texttt{ARRAY} \longrightarrow \texttt{INDEX}$$

The theory $\mathcal{AX}_{\texttt{diff}}$ of arrays with diff

Since $\mathcal{AR}_{\text{ext}}$ does not have quantifier-free interpolants, we consider the following variant, which we call $\mathcal{AX}_{\texttt{diff}}$. We add a further symbol in the signature

$$\texttt{diff} : \texttt{ARRAY} \times \texttt{ARRAY} \longrightarrow \texttt{INDEX}$$

We replace the extensionality axiom (9) by its skolemization

$$\forall x, y. \qquad x \neq y \rightarrow \; rd(x, \texttt{diff}(x,y)) \neq rd(y, \texttt{diff}(x,y))$$

The theory $\mathcal{AX}_{\texttt{diff}}$ of arrays with diff

Since $\mathcal{AR}_{\mathrm{ext}}$ does not have quantifier-free interpolants, we consider the following variant, which we call $\mathcal{AX}_{\texttt{diff}}$. We add a further symbol in the signature

$$\texttt{diff} : \texttt{ARRAY} \times \texttt{ARRAY} \longrightarrow \texttt{INDEX}$$

We replace the extensionality axiom (9) by its skolemization

$$\forall x, y. \qquad x \neq y \to \; rd(x, \texttt{diff}(x, y)) \neq rd(y, \texttt{diff}(x, y))$$

Theorem

The (universal) theory $\mathcal{AX}_{\texttt{diff}}$ has quantifier-free interpolation.

Arrays with `diff`: amalgamation

The above theorem can be proved in various independent ways:

- *semantically* [BGT 12]: by showing amalgamation property;
- *syntactically* [BGT 12]: by rewriting techniques, via a specific adaptation of Knuth-Bendix completion (called 'Gaussian completion');
- *syntactically* [TW 16]: by hierarchical reduction to $\mathcal{EUF}$ (this is the best method from the complexity viewpoint).

Arrays with diff: strong amalgamation

We can strenghten the above result

Arrays with `diff`: strong amalgamation

We can strenghten the above result

Theorem

The (universal) theory $\mathcal{AX}_{\mathtt{diff}}$ has general quantifier-free interpolation.

Arrays with `diff`: strong amalgamation

Again this theorem can be proved:

- *semantically* [BGT 12]: by showing strong amalgamation property;
- *syntactically* [implicit in BGT 12]: by rewriting techniques.

Arrays with `diff`: strong amalgamation

Again this theorem can be proved:

- *semantically* [BGT 12]: by showing strong amalgamation property;
- *syntactically* [implicit in BGT 12]: by rewriting techniques.

The reason why rewriting techniques work is because they allow to compute equality interpolating terms in the following way.

Arrays with `diff`: strong amalgamation

Again this theorem can be proved:

- *semantically* [BGT 12]: by showing strong amalgamation property;
- *syntactically* [implicit in BGT 12]: by rewriting techniques.

The reason why rewriting techniques work is because they allow to compute equality interpolating terms in the following way.

The completion of a pair of constraints $\delta(\underline{x}, \underline{y}) \wedge \theta(\underline{x}, \underline{z})$ produces a finite disjunction $\bigvee_i (\delta_i(\underline{x}, \underline{y}) \wedge \theta_i(\underline{x}, \underline{z}))$ of constraints without mixed terms. So whenever a disjunction of equalities is entailed, each disjunct entails a single equality whose normal form is an equality of the kind $t = t$, with shared t. Such t's are the equality interpolating terms.

Arrays with `diff`: uniform interpolation?

Theorem

The (universal) theory $\mathcal{AX}_{\texttt{diff}}$ does not have uniform quantifier-free interpolation.

Arrays with `diff`: uniform interpolation?

Theorem

The (universal) theory $\mathcal{AX}_{\texttt{diff}}$ does not have uniform quantifier-free interpolation.

A counterexample is the formula

$$rd(c_1, i) \neq rd(c_2, i) \wedge rd(d_1, i) = rd(d_2, i) \tag{6}$$

An argument based on ultraproducts show that we cannot eliminate uniformly the index variable i from it.

Arrays with diff: uniform interpolation?

This is the schema of the argument. A uniform interpolant (supposing it exists) is a formula $UI(c_1, c_2, d_1, d_2)$ implied by (6) and having the property that it implies all formulas - not containing i - implied by (6). Consider the infinitely many formulae

$$\phi_n \ \equiv \ c_1 \sim_n c_2 \to \bigvee_{j=1}^{n} rd(d_1, \mathtt{diff}_n(c_1, c_2)) = rd(d_2, \mathtt{diff}_n(c_1, c_2))$$

where $c_1 \sim_n c_2$ says that c_1 and c_2 differ in at most n indices and $\mathtt{diff}_n$ is the iterated `diff` operator (both such constructs are quantifier-free definable in $\mathcal{AX}_{\mathtt{diff}}$).

Arrays with `diff`: uniform interpolation?

One now builds models M_n such that $M_n \not\models \phi_n$. Hence $M_n \models \neg UI$.

Arrays with `diff`: uniform interpolation?

One now builds models M_n such that $M_n \not\models \phi_n$. Hence $M_n \models \neg UI$.

Taking an untraproduct $\Pi_D M_n$ modulo a non principal ultrafilter, by Łos theorem, we get $\Pi_D M_n \models \neg UI$.

Arrays with `diff`: uniform interpolation?

One now builds models M_n such that $M_n \not\models \phi_n$. Hence $M_n \models \neg UI$.

Taking an untraproduct $\Pi_D M_n$ modulo a non principal ultrafilter, by Łos theorem, we get $\Pi_D M_n \models \neg UI$.

However, since it is possible to build an extension $N \supseteq \Pi_D M_n$ satisfying (6), we get $N \models UI$ (because (6) implies UI) and also $\Pi_D M_n \models UI$, because UI is quantifier-free and hence preserved by substructures. Contradiction.

Outline

A more expressive theory

- Plain `diff` operation is semantically undetermined, we want to replace it with a more informative operation.

A more expressive theory

- Plain `diff` operation is semantically undetermined, we want to replace it with a more informative operation.
- To this aim we introduce [GGKN 23] the **Theory of Arrays with MaxDiff**, where **MaxDiff** returns the biggest index where two arrays differ (it returns the conventional value 0 is they are equal).

A more expressive theory

- Plain `diff` operation is semantically undetermined, we want to replace it with a more informative operation.
- To this aim we introduce [GGKN 23] the **Theory of Arrays with MaxDiff**, where **MaxDiff** returns the biggest index where two arrays differ (it returns the conventional value 0 is they are equal).
- This theory is parameterized on different 'index theories'; the typical index theory is Presburger arithmetic (with 'division by n' for all n in the language).

A more expressive theory

- Plain `diff` operation is semantically undetermined, we want to replace it with a more informative operation.
- To this aim we introduce [GGKN 23] the **Theory of Arrays with MaxDiff**, where **MaxDiff** returns the biggest index where two arrays differ (it returns the conventional value 0 is they are equal).
- This theory is parameterized on different 'index theories'; the typical index theory is Presburger arithmetic (with 'division by n' for all n in the language).
- The theory has also a length operation $| - |$: now an array a has the undefined value '$\perp$' outside the interval $[0, |a|]$.

A more expressive theory

- Plain `diff` operation is semantically undetermined, we want to replace it with a more informative operation.
- To this aim we introduce [GGKN 23] the **Theory of Arrays with MaxDiff**, where **MaxDiff** returns the biggest index where two arrays differ (it returns the conventional value 0 is they are equal).
- This theory is parameterized on different 'index theories'; the typical index theory is Presburger arithmetic (with 'division by n' for all n in the language).
- The theory has also a length operation $|-|$: now an array a has the undefined value '$\perp$' outside the interval $[0, |a|]$.
- There is a remarkable gain in expressiveness, we show that interpolation properties can be maintained.

Index Theory

To locate our contribution, we need the notion of *index theory*.

Definition

An *index theory* T_I is a mono-sorted theory (let `INDEX` be its sort) satisfying the following conditions:

- T_I is **universal, stably infinite** and has the **general quantifier-free interpolation property**;
- T_I has **decidable** quantifier-free fragment;
- T_I **extends** the theory TO of linear orderings with a distinguished element 0.

Index Theory

To locate our contribution, we need the notion of *index theory*.

Definition

An *index theory* T_I is a mono-sorted theory (let `INDEX` be its sort) satisfying the following conditions:

- T_I is **universal, stably infinite** and has the **general quantifier-free interpolation property**;
- T_I has **decidable** quantifier-free fragment;
- T_I **extends** the theory TO of linear orderings with a distinguished element 0.

Examples of index theories T_I are TO itself, integer difference logic integer linear arithmetic, and real linear arithmetics.

$\mathcal{ARD}(T_I)$: the Theory of Arrays with MaxDiff

Axioms: the axioms of T_I, and

$$\forall y, i, e,\ |wr(y, i, e)| = |y| \tag{7}$$

$$\forall y, i,\ wr(y, i, \bot) = y \tag{8}$$

$$\forall y, i, e,\ (e \neq \bot \wedge\ 0 \leq i \leq |y|) \to rd(wr(y, i, e), i) = e \tag{9}$$

$$\forall y, i, j, e,\ i \neq j \to rd(wr(y, i, e), j) = rd(y, j) \tag{10}$$

$$\forall y, i,\ rd(y, i) \neq \bot \leftrightarrow 0 \leq i \leq |y| \tag{11}$$

$$\forall y,\ |y| \geq 0 \tag{12}$$

$$\forall y, \texttt{diff}(y, y) = 0 \tag{13}$$

$$\forall x, y,\ x \neq y \to rd(x, \texttt{diff}(x, y)) \neq rd(y, \texttt{diff}(x, y)). \tag{14}$$

$$\forall x, y, i,\ \texttt{diff}(x, y) < i \to rd(x, i) = rd(y, i). \tag{15}$$

$$\bot \neq el. \tag{16}$$

$\mathcal{ARD}(T_I)$: the Theory of Arrays with MaxDiff

The quantifier-free fragment of theis theory is decidable, because it can be embedded into Bradley's 'array property fragment'.

$\mathcal{ARD}(T_I)$: the Theory of Arrays with MaxDiff

The quantifier-free fragment of theis theory is decidable, because it can be embedded into Bradley's 'array property fragment'.
In fact atoms of the kind

$$a = b, \quad |a| = k, \quad \texttt{diff}(a, b) = j, \quad wr(a, i, e) = b \tag{17}$$

can be translated into universal formulae of $T_I \cup \mathcal{EUF}$ in Bradley's fragment (we call such formulae their B-translations).

$\mathcal{ARD}(T_I)$: the Theory of Arrays with MaxDiff

Theorem

The (universal) theory $\mathcal{ARD}(T_I)$ has quantifier-free interpolation.

The theorem can be proved [GGKN 23]:

$\mathcal{ARD}(T_I)$: the Theory of Arrays with MaxDiff

Theorem

The (universal) theory $\mathcal{ARD}(T_I)$ has quantifier-free interpolation.

The theorem can be proved [GGKN 23]:

- *semantically*: by showing amalgamation property;

$\mathcal{ARD}(T_I)$: the Theory of Arrays with MaxDiff

Theorem

The (universal) theory $\mathcal{ARD}(T_I)$ has quantifier-free interpolation.

The theorem can be proved [GGKN 23]:

- *semantically*: by showing amalgamation property;
- *syntactically*: by hierarchical reduction to $T_I \cup \mathcal{EUF}$.

$\mathcal{ARD}(T_I)$: the Theory of Arrays with MaxDiff

Theorem

The (universal) theory $\mathcal{ARD}(T_I)$ has quantifier-free interpolation.

The theorem can be proved [GGKN 23]:

- *semantically*: by showing amalgamation property;
- *syntactically*: by hierarchical reduction to $T_I \cup \mathcal{EUF}$.

In both cases, the proof follows the same schema as in the case of in the case of $\mathcal{AX}_{\texttt{diff}}$, but details are much more challenging. We give some qualitative account of the second proof.

Interpolation for $\mathcal{ARD}(T_I)$

- **Our problem**: given two qf formulae A and B s.t. $A \wedge B$ is not satisfiable (modulo $\mathcal{ARD}(T_I)$), to compute a qf formula C s.t. $\mathcal{ARD}(T_I) \models A \to C$, $\mathcal{ARD}(T_I) \models C \wedge B \to \bot$ and s.t. C contains only the free constants (called *common constants*) occurring both in A and in B.

Interpolation for $\mathcal{ARD}(T_I)$

- **Our problem**: given two qf formulae A and B s.t. $A \wedge B$ is not satisfiable (modulo $\mathcal{ARD}(T_I)$), to compute a qf formula C s.t. $\mathcal{ARD}(T_I) \models A \to C$, $\mathcal{ARD}(T_I) \models C \wedge B \to \bot$ and s.t. C contains only the free constants (called *common constants*) occurring both in A and in B.
- There are infinitely many common terms out of finitely many common constants: **iterated diff operations** $\texttt{diff}_k$ are needed in our algorithm to discover 'implicit' common facts.

Interpolation for $\mathcal{ARD}(T_I)$

- **Our problem**: given two qf formulae A and B s.t. $A \wedge B$ is not satisfiable (modulo $\mathcal{ARD}(T_I)$), to compute a qf formula C s.t. $\mathcal{ARD}(T_I) \models A \to C$, $\mathcal{ARD}(T_I) \models C \wedge B \to \bot$ and s.t. C contains only the free constants (called *common constants*) occurring both in A and in B.
- There are infinitely many common terms out of finitely many common constants: **iterated diff operations** $\mathtt{diff}_k$ are needed in our algorithm to discover 'implicit' common facts.
- E.g., $\mathtt{diff}_2$ returns the last-but-one index where a, b differ (0 if a, b differ in at most one index), $\mathtt{diff}_3$ the last-but-two index etc.

Interpolation for $\mathcal{ARD}(T_I)$

- **Our problem**: given two qf formulae A and B s.t. $A \wedge B$ is not satisfiable (modulo $\mathcal{ARD}(T_I)$), to compute a qf formula C s.t. $\mathcal{ARD}(T_I) \models A \to C$, $\mathcal{ARD}(T_I) \models C \wedge B \to \bot$ and s.t. C contains only the free constants (called *common constants*) occurring both in A and in B.
- There are infinitely many common terms out of finitely many common constants: **iterated diff operations** $\mathtt{diff}_k$ are needed in our algorithm to discover 'implicit' common facts.
- E.g., $\mathtt{diff}_2$ returns the last-but-one index where a, b differ (0 if a, b differ in at most one index), $\mathtt{diff}_3$ the last-but-two index etc.
- Those iterated operators are definable in our language.

Interpolation for $\mathcal{ARD}(T_I)$

Step 0. Write both A and B in the form $\Phi_1 \wedge \Phi_2$, where Φ_2 is a pure $T_I \cup \mathcal{EUF}$-formula and Φ_1 is a conjunction of atoms of the form (17); add also missing atoms of the kind $|d| = k_d$ to both A and B (extra free constants are employed here).

Step 1. Let N be equal to the number of index constants occurring in A, B (plus one); for every pair of common `ARRAY`-constants c_1, c_2, pick fresh `INDEX` constants $k_1, \dots, k_N$ and add the atoms $\texttt{diff}_n(c_1, c_2) = k_n$ (for all $n = 1, \dots, N$) to both A and B.

Step 2. B-instantiate formulae (17) with index constants (both inside A and inside B).

Step 3. Now (this is the delicate fact to be proved) the $T_I \cup \mathcal{EUF}$-part of $A \cup B$ become inconsistent. Since T_I has general quantifier-free interpolation, we can compute the related interpolant. To get our desired $\mathcal{ARD}(T_I)$-interpolant, we only have to replace back in it the fresh constants introduced in Step 1 by the common terms they name.

General Interpolation for $\mathcal{ARD}(T_I)$?

Strong amalgamation however fails [GGKN 23]:

Theorem

The (universal) theory $\mathcal{ARD}(T_I)$ does not have general quantifier-free interpolation.

In fact, this is a counterexample to general interpolation:

$$\begin{array}{ll}(A) & |a| = 0 \wedge rd(a, 0) = e \wedge P(a) \\ (B) & |b| = 0 \wedge rd(b, 0) = e \wedge \neg P(b).\end{array}$$

General Interpolation

To restore superamalgamation, one needs a use of constant arrays. We add a unary function $\texttt{Const} : \texttt{INDEX} \to \texttt{ARRAY}$, constrained by the following axioms:

$$\forall i, |\texttt{Const}(i)| = \max(i, 0). \tag{18}$$

$$\forall i, j,\ (0 \le j \land j \le |\texttt{Const}(i)| \to rd(\texttt{Const}(i), j) = el). \tag{19}$$

Thus $\texttt{Const}(i)$ is the constant array of length i and value the distinguished element el (the atom $P(wr(\texttt{Const}(0), 0, e))$ works now as interpolant in the above counterexample).

General Interpolation

General interpolation for this theory has been proved in [GGKN 23] only semantically (via strong amalgamation).

We conjecture hierarchical reduction works too, but it is not clear whether the reduction to $T_I \cup \mathcal{EUF}$ can be kept to be polynomial.

Conclusions

As we saw, it is possible to design array theories which are significantly expressive, while still enjoying quantifier-free and general quantifier-free interpolation properties.

Conclusions

As we saw, it is possible to design array theories which are significantly expressive, while still enjoying quantifier-free and general quantifier-free interpolation properties.

This is remarkable, because array theories are not decidable at the elementary level (only a limited use of quantifiers can guarantee decidability).

Conclusions

As we saw, it is possible to design array theories which are significantly expressive, while still enjoying quantifier-free and general quantifier-free interpolation properties.

This is remarkable, because array theories are not decidable at the elementary level (only a limited use of quantifiers can guarantee decidability).

Further enrichments still need to be adequately investigated.