

Interpolation systems for non-ground proofs¹

Maria Paola Bonacina

Dipartimento di Informatica
Università degli Studi di Verona
Verona, Italy

August 31, 2016

¹Joint work with Moa Johansson

Preliminaries

Counter-examples to the color-based approach

A two-stage approach

Discussion

What is interpolation?

- ▶ Formulæ A and B such that $A \vdash B$
- ▶ An **interpolant** I is a formula such that
 - ▶ $A \vdash I$
 - ▶ $I \vdash B$
 - ▶ All uninterpreted symbols in I are **common** to A and B

Assume that at least one of A and B has at least one symbol that does not appear in the other

Proofs by refutation: reverse interpolant

- ▶ A and B inconsistent: $A, B \vdash \perp$
- ▶ Then a **reverse interpolant** I is a formula such that
 - ▶ $A \vdash I$
 - ▶ $B, I \vdash \perp$
 - ▶ All uninterpreted symbols in I are **common** to A and B

Clausal theorem proving: A and B are sets of clauses

Remarks

Reverse interpolant of (A, B) : interpolant of $(A, \neg B)$
because $A, B \vdash \perp$ means $A \vdash \neg B$ and $B, I \vdash \perp$ means $I \vdash \neg B$

I reverse interpolant of (A, B) : $\neg I$ reverse interpolant of (B, A)
because $A \vdash I$ means $A, \neg I \vdash \perp$ and $B, I \vdash \perp$ means $B \vdash \neg I$

In refutational settings we say interpolant for reverse interpolant

Terminology for interpolation: Colors

Uninterpreted symbol:

- ▶ *A*-colored: occurs in *A* and not in *B*
- ▶ *B*-colored: occurs in *B* and not in *A*
- ▶ Transparent: occurs in both

Alternative terminology: *A*-local, *B*-local, global

Terminology for interpolation: Colors

Ground term/literal/clause:

- ▶ All transparent symbols: **transparent**
- ▶ A -colored (at least one) and transparent symbols: **A -colored**
- ▶ B -colored (at least one) and transparent symbols: **B -colored**
- ▶ Otherwise: **AB -mixed**

Interpolation system

- ▶ Given refutation of $A \cup B$ extracts interpolant of (A, B)
- ▶ Associates **partial interpolant** $PI(C)$ to every clause C
- ▶ Defined **inductively** based on those of parents
- ▶ $PI(\square)$ is interpolant of (A, B)

Complete interpolation system

An interpolation system is **complete** for an inference system if

- ▶ For all sets of clauses A and B such that $A \cup B$ is unsatisfiable
- ▶ For all refutations of $A \cup B$ by the inference system

It generates **an** interpolant of (A, B)

There may be more than one

What an interpolation system really does

An interpolation system determines whether a literal L should be added to the interpolant I by:

- ▶ Detecting whether L comes from the A side or the B side of the refutation to ensure $A \vdash I$ and $B, I \vdash \perp$
- ▶ Checking that uninterpreted symbols in L are transparent to ensure that I is transparent

Color-based interpolation systems

- ▶ Achieve both goals by classifying symbols based on signature (the colors) and tracking them in the refutation
- ▶ Cannot handle AB -mixed literals
- ▶ Good for:
 - ▶ Propositional refutations [Krajíček 1997] [Pudlák 1997] [McMillan 2003]
 - ▶ Equality sharing combination of convex equality-interpolating theories [Yorsh, Musuvathi 2005]
 - ▶ Ground first-order refutations under a separating ordering (transparent terms smaller than colored) [MPB, Johansson 2011]

Interpolation of non-ground proofs?

- ▶ Inference system Γ for first-order logic with equality
- ▶ Γ -inferences apply substitutions: most general unifiers, matching substitutions, to instantiate (universally quantified) variables
- ▶ Interpolation in the presence of variables and substitutions?
- ▶ Substitutions easily create AB -mixed literals

Conjecture

Does a separating ordering prevent AB -mixed literals in the general case like in the ground case?

No

Counter-example

f, g : transparent a : A -colored b : B -colored

- ▶ $g(y, b) \simeq y$ and
- ▶ $f(g(a, x), x) \simeq f(x, a)$
- ▶ With $\sigma = \{y \leftarrow a, x \leftarrow b\}$
- ▶ Generate $f(a, b) \simeq f(b, a)$
- ▶ Where both sides are AB -mixed literals
- ▶ And the inference is compatible with a separating ordering

Conjecture

Can the color-based approach work if we give up completeness and restrict the attention to proofs with no AB -mixed literals?

No

Counter-example

P : transparent a : A -colored b : B -colored

- ▶ $\neg P(x, b) \vee C$ and $P(a, y) \vee D$
- ▶ Where C and D contain no AB -mixed literals,
 $x \notin \text{Var}(C)$, $y \notin \text{Var}(D)$
- ▶ With $\sigma = \{x \leftarrow a, y \leftarrow b\}$
- ▶ Generate $(C \vee D)\sigma = C \vee D$: no AB -mixed literals
- ▶ But literals resolved upon $\neg P(a, b)$ and $P(a, b)$ are AB -mixed
so that the A -colored/ B -colored/transparent case analysis of
the colored approach does not suffice

Local or colored proofs

- ▶ **Local** proof: only **local** inferences
- ▶ **Local** inference: involves at most one color
- ▶ Equivalent characterization: no **AB-mixed** clauses
- ▶ Hence the name **colored** proof

[McMillan 2008] [Kovács, Voronkov 2009] [Hoder, Kovács, Voronkov 2012]

Conjecture

Can the color-based approach work if we give up completeness and restrict the attention to **colored** proofs?

No

Counter-example

L, R, Q : transparent a, c : A -colored

- ▶ $p_1: L(x, a) \vee R(x)$ with partial interpolant $PI(p_1)$ and
- ▶ $p_2: \neg L(c, y) \vee Q(y)$ with partial interpolant $PI(p_2)$
- ▶ With $\sigma = \{x \leftarrow c, y \leftarrow a\}$
- ▶ Generate $R(c) \vee Q(a)$
- ▶ Even if $PI(p_1)$ and $PI(p_2)$ are transparent
- ▶ $(PI(p_1) \vee PI(p_2))\sigma$ is not guaranteed to be, because x may appear in $PI(p_1)$ and y may appear in $PI(p_2)$

A two-stage approach

- ▶ Separate entailment and transparency requirements
- ▶ **First stage:** compute **provisional interpolant** \hat{I} such that $A \vdash \hat{I}$ and $B, \hat{I} \vdash \perp$
- ▶ \hat{I} may contain colored symbols
- ▶ **Second stage:** transform \hat{I} into interpolant I

Use labels to track where literals come from

- ▶ **Labeled Γ -proof tree**: attach a label to every literal
- ▶ A literal L may occur in more than one clause; the label depends on **both** literal and clause
- ▶ Labels are independent of signatures
- ▶ Labels are independent of substitutions
- ▶ All literals are labeled, including **AB-mixed** ones

Labeled Γ -proof tree

- ▶ Clause in A : literals get label **A**
- ▶ Clause in B : literals get label **B**
- ▶ Literals in resolvents inherit labels from literals in parents
- ▶ **Resolvent** c : $(C \vee D)\sigma$ of $p_1: L \vee C$ and $p_2: \neg L' \vee D$ with $L\sigma = L'\sigma$: for all $M \in C$, $label(M\sigma, c) = label(M, p_1)$
for all $M \in D$, $label(M\sigma, c) = label(M, p_2)$
- ▶ **Factor** c : $(L \vee C)\sigma$ of $p: L \vee L' \vee C$ with $L\sigma = L'\sigma$:
for all $M \in C$, $label(M\sigma, c) = label(M, p)$, and

$$label(L\sigma, c) = \begin{cases} \mathbf{A} & \text{if } label(L, p) = label(L', p) = \mathbf{A} \\ \mathbf{B} & \text{otherwise} \end{cases}$$

Example

$$L(x_1, c)_A \vee P(x_1)_A \vee Q(x_1, y_1)_A$$

$$\neg L(c, x_2)_B \vee P(x_2)_B \vee R(x_2, y_2)_B$$

$$\sigma = \{x_1 \leftarrow c, x_2 \leftarrow c\}$$

$$\text{Resolvent: } P(c)_A \vee Q(c, y_1)_A \vee P(c)_B \vee R(c, y_2)_B$$

which becomes $Q(c, y_1)_A \vee P(c)_B \vee R(c, y_2)_B$ after factoring

Labeled Γ -proof tree with equality

- ▶ **Paramodulation/Superposition/Simplification**: as for resolution except that **new** literal generated by **equational replacement** inherits label of **para-into** literal
- ▶ $(C \vee L[r] \vee D)\sigma$ generated by paramodulating $p_1: s \simeq r \vee C$ into $p_2: L[s'] \vee D$ with $s\sigma = s'\sigma$:
 - for all $M \in C$, $label(M\sigma, c) = label(M, p_1)$
 - for all $M \in D$, $label(M\sigma, c) = label(M, p_2)$
 - and $label(L[r]\sigma, c) = label(L[s'], p_2)$

Partial interpolant

- ▶ Clause C in refutation of $A \cup B$
- ▶ $A \wedge B \vdash C$
- ▶ $A \wedge B \vdash C \vee \bar{C}$
- ▶ $A \wedge \bar{C} \vdash \bar{B} \vee C$
- ▶ Interpolant of $A \wedge \bar{C}$ and $\bar{B} \vee C$
- ▶ Reverse interpolant of $A \wedge \bar{C}$ and $B \wedge \bar{C}$
- ▶ The literals of $A \wedge \bar{C}$ ($B \wedge \bar{C}$) do not necessarily come from the A (B) side of the proof
- ▶ Use **projections based on labels**

Labeled projections

- ▶ $C|_{\mathbf{A}}$: literals of C labeled \mathbf{A}
- ▶ $C|_{\mathbf{B}}$: literals of C labeled \mathbf{B}
- ▶ \perp if empty
- ▶ **Commute with substitutions:**
for resolvent $(C \vee D)\sigma$
 $(C \vee D)\sigma|_{\mathbf{A}} = (C|_{\mathbf{A}} \vee D|_{\mathbf{A}})\sigma$

Provisional partial interpolants

- ▶ **Provisional partial interpolant** $\widehat{PI}(C)$ of clause C in refutation of $A \cup B$:
provisional interpolant of $A \wedge \neg(C|_A)$ and $B \wedge \neg(C|_B)$
- ▶ $\widehat{PI}(\square)$ is provisional interpolant of (A, B)

Provisional interpolation system $\widehat{\Gamma}$

- ▶ $c: C \in A: \widehat{PI}(c) = \perp$
- ▶ $c: C \in B: \widehat{PI}(c) = \top$
- ▶ **Resolvent** $c: (C \vee D)\sigma$ of $p_1: L \vee C$ and $p_2: \neg L' \vee D$:
 - ▶ Both literals **A**-labeled: $\widehat{PI}(c) = (\widehat{PI}(p_1) \vee \widehat{PI}(p_2))\sigma$
 - ▶ Both literals **B**-labeled: $\widehat{PI}(c) = (\widehat{PI}(p_1) \wedge \widehat{PI}(p_2))\sigma$
 - ▶ Positive **A**-labeled and negative **B**-labeled:
 $\widehat{PI}(c) = [(L \vee \widehat{PI}(p_1)) \wedge \widehat{PI}(p_2)]\sigma$
 - ▶ Positive **B**-labeled and negative **A**-labeled:
 $\widehat{PI}(c) = [\widehat{PI}(p_1) \wedge (\neg L' \vee \widehat{PI}(p_2))]\sigma$

Provisional interpolation system $\widehat{\Gamma}$

- ▶ **Factor** $c: (L \vee C)\sigma$ of $p: L \vee L' \vee C$:

$$\widehat{PI}(c) = \begin{cases} \widehat{PI}(p)\sigma & \text{if } \text{label}(L, p) = \text{label}(L', p) \\ (L \vee \widehat{PI}(p))\sigma & \text{otherwise} \end{cases}$$

Provisional interpolation system $\hat{\Gamma}$

► **Paramodulation/Superposition/Simplification:**

$(C \vee L[r] \vee D)\sigma$ generated by paramodulating $p_1: s \simeq r \vee C$ into $p_2: L[s'] \vee D$:

- Both literals **A**-labeled: $\widehat{PI}(c) = (\widehat{PI}(p_1) \vee \widehat{PI}(p_2))\sigma$
- Both literals **B**-labeled: $\widehat{PI}(c) = (\widehat{PI}(p_1) \wedge \widehat{PI}(p_2))\sigma$
- Para-from **A**-labeled and para-into **B**-labeled:
 $\widehat{PI}(c) = [(s \simeq r \vee \widehat{PI}(p_1)) \wedge \widehat{PI}(p_2)]\sigma$
- Para-from **B**-labeled and para-into **A**-labeled:
 $\widehat{PI}(c) = [\widehat{PI}(p_1) \wedge (s \not\simeq r \vee \widehat{PI}(p_2))]\sigma$

Example

$$A = \{f(x) \simeq g(a, x)\} \quad B = \{P(f(b)), \neg P(g(y, b))\}$$

\succ : recursive path ordering based on precedence $f > g > a$

1. $f(x) \simeq g(a, x)_{(\mathbf{A})} [\perp]$ paramodulates into $P(f(b))_{(\mathbf{B})} [\top]$ to yield $P(g(a, b))_{(\mathbf{B})} [f(b) \simeq g(a, b)]$

$$\widehat{PI}(P(g(a, b))) = (f(b) \simeq g(a, b) \vee \perp) \wedge \top = f(b) \simeq g(a, b)$$

2. $P(g(a, b))_{(\mathbf{B})} [f(b) \simeq g(a, b)]$ and $\neg P(g(y, b))_{(\mathbf{B})} [\top]$ resolve to yield $\square [f(b) \simeq g(a, b)]$

$$\hat{\Gamma} = \widehat{PI}(\square) = f(b) \simeq g(a, b) \wedge \top = f(b) \simeq g(a, b)$$

A complete provisional interpolation system

- ▶ $\Gamma\hat{\Gamma}$ builds provisional interpolant mostly by adding instances of **A**-labeled literals resolved, factorized, or paramodulated with **B**-labeled ones: [communication interface](#)
- ▶ **Theorem:** The provisional interpolation system $\Gamma\hat{\Gamma}$ is complete
- ▶ **Lemma:** The provisional interpolants generated by $\Gamma\hat{\Gamma}$ are in negation normal form with \forall -quantified variables and all predicate symbols are either transparent or interpreted (e.g., equality)

Second stage: lifting

- ▶ A closed formula is **color-flat** if its only colored symbols are constant symbols
- ▶ Equivalently: all function symbols are interpreted or transparent
- ▶ **Lifting** replaces **A-colored** constants by \exists -quantified variables and **B-colored** constants by \forall -quantified variables
- ▶ If \hat{I} is color-flat, $Lift(\hat{I})$ is **transparent**
- ▶ Since only constants are replaced the order of introduced quantifiers is immaterial: different orders yield different interpolants

Example (continued)

$$A = \{f(x) \simeq g(a, x)\} \quad B = \{P(f(b)), \neg P(g(y, b))\}$$

a is A -colored, P and b are B -colored, f and g are transparent

1. Provisional interpolant:

$$\hat{I} = f(b) \simeq g(a, b) \wedge \top = f(b) \simeq g(a, b)$$

The only colored symbols are constants

2. Two interpolants:

$$I_1 = \text{Lift}(\hat{I}) = \forall v. \exists w. f(v) \simeq g(w, v)$$

$$I_2 = \text{Lift}(\hat{I}) = \exists w. \forall v. f(v) \simeq g(w, v)$$

From provisional interpolants to interpolants

- ▶ **Lemma:** If \hat{I} is a color-flat, $B \wedge \hat{I} \vdash \perp$ implies $B \wedge \text{Lift}(\hat{I}) \vdash \perp$
BWOC: assume $B \wedge \text{Lift}(\hat{I})$ has model \mathcal{M} ;
 \mathcal{M} satisfies also the instance of $\text{Lift}(\hat{I})$ where the \forall -quantified vars are replaced by the **B-colored** constants originally in \hat{I} ;
we build model \mathcal{M}' of $B \wedge \hat{I}$;
 \mathcal{M}' interprets **B-colored** and **transparent** symbols like \mathcal{M} ;
the only difference is given by the **A-colored** constants in \hat{I} that are **new** for \mathcal{M} :
let \mathcal{M}' interpret them with the individuals picked by \mathcal{M} for the \exists -quantified vars in $\text{Lift}(\hat{I})$.

From provisional interpolants to interpolants

- ▶ **Lemma:** If \hat{I} is a color-flat, $A \vdash \hat{I}$ implies $A \vdash \text{Lift}(\hat{I})$
 $A \wedge \neg \hat{I} \vdash \perp$ implies $A \wedge \neg \text{Lift}(\hat{I}) \vdash \perp$
BWOC: assume $A \wedge \neg \text{Lift}(\hat{I})$ has model \mathcal{M} ;
 \mathcal{M} satisfies also the instance of $\text{Lift}(\hat{I})$ where the \forall -quantified vars (after negation!) are replaced by the **A-colored** constants originally in \hat{I} ; we build model \mathcal{M}' of $A \wedge \neg \hat{I}$;
 \mathcal{M}' interprets **A-colored** and **transparent** symbols like \mathcal{M} ;
the only difference is given by the **B-colored** constants in $\neg \hat{I}$ that are **new** for \mathcal{M} :
let \mathcal{M}' interpret them with the individuals picked by \mathcal{M} for the \exists -quantified vars (after negation!) in $\neg \text{Lift}(\hat{I})$.

A complete interpolation system

- ▶ **Theorem:** If \hat{I} is a color-flat provisional interpolant of (A, B) , then $Lift(\hat{I})$ is an interpolant of (A, B)
- ▶ **Corollary:** Complete provisional interpolation system + lifting = complete interpolation system

Summary

- ▶ Interpolation systems for non-ground proofs
- ▶ The color-based approach does not work
- ▶ The two-stage approach does
- ▶ Other approaches: transform the proof; but none works for non-ground proofs with colored uninterpreted function symbols
- ▶ The two-stage approach covers also $DPLL(\Gamma + \mathcal{T})$

DPLL($\Gamma + \mathcal{T}$)

- ▶ Integrates SMT-solver DPLL(\mathcal{T}) and first-order inference system Γ
- ▶ Combines built-in and axiomatized theories
- ▶ Makes first-order inferences model-driven by the candidate model built by the SMT-solver
- ▶ Yields some decision procedures for satisfiability of first-order formulæ

DPLL($\Gamma + \mathcal{T}$)

- ▶ Works with **hypothetical clauses** $H \triangleright C$, where C is a clause, and H a set of ground literals from the trail used to infer C
- ▶ When $H \triangleright C$, with C ground, is in conflict, it generates the ground conflict clause $\neg H \vee C$
- ▶ $\neg H \vee C$ may enter a DPLL($\Gamma + \mathcal{T}$)-refutation, with its Γ -proof tree as subproof
- ▶ The Γ -proof tree is **not necessarily ground**

Refutation by $\text{DPLL}(\Gamma + \mathcal{T})$

- ▶ DPLL-CDCL -refutation: propositional resolution
- ▶ $\text{DPLL}(\mathcal{T})$ -refutation: propositional resolution + \mathcal{T} -lemmas (\mathcal{T} -conflict clauses are \mathcal{T} -lemmas)
- ▶ $\text{DPLL}(\Gamma + \mathcal{T})$ -refutation: $\text{DPLL}(\mathcal{T})$ -refutation + Γ -proof trees as subtrees

Model-based theory combination in $DPLL(\Gamma+\mathcal{T})$

- ▶ Each \mathcal{T}_i -solver builds a candidate \mathcal{T}_i -model M_i
- ▶ Generate and propagate ground equalities $t \simeq s$ true in M_i
- ▶ If inconsistent, backtrack
- ▶ $t \simeq s$ may end up in \mathcal{T} -lemmas or hypothetical clauses, hence in the $DPLL(\Gamma+\mathcal{T})$ -refutation
- ▶ No guarantee that $t \simeq s$ is not **AB-mixed**

Interpolation for $DPLL(\Gamma + \mathcal{T})$

- ▶ $\Gamma \hat{+}$ (provisional) interpolation system for $DPLL(\mathcal{T}) =$
provisional interpolation system for $DPLL(\Gamma + \mathcal{T})$
- ▶ Color-flat provisional interpolants: interpolants via lifting
- ▶ Provisional interpolants do not need to be transparent: no
need to restrict \mathcal{T} to convex equality-interpolating theories to
avoid **AB-mixed** literals
- ▶ Model-based theory combination also allowed

References

- ▶ Maria Paola Bonacina and Moa Johansson. On interpolation in automated theorem proving. *Journal of Automated Reasoning*, 54(1):69-97, 2015 [providing 61 references]
- ▶ Maria Paola Bonacina. Two-stage interpolation systems (Abstract). Notes of the First International Workshop on Interpolation: from Proofs to Applications (IPrA), St. Petersburg, Russia, July 2013; TR TU-Wien 2013