Interpolation for resolution and superposition [1]

Maria Paola Bonacina

Dipartimento di Informatica
Università degli Studi di Verona
Verona, Italy

November, 2012

[1] Joint work with Moa Johansson

What is interpolation?

- Extracts interpolants from proofs
- *Interpolant*: a formula *in between* formulæ
- Proof of inconsistency of A and B: $A, B \vdash \bot$
- Interpolant I:
  - $A \vdash I$
  - $I, B \vdash \bot$
  - I: only *shared* symbols

Why interpolation?

Theorem proving support of system analysis/synthesis:

- Model Checking:
  Interpolants as *over-approximations* of images or pre-images
- Counter-Example Guided Abstraction Refinement:
  Interpolants to *refine abstractions*
- Automated invariant generation:
  Interpolants as *candidate invariants*:
  formulæ with quantifiers, alternating quantifiers

Why resolution and superposition?

- At the heart of state-of-the-art theorem provers
- *Decision procedures* for decidable theories, fragments
- Integration with SMT-solvers
- *Complete* reasoning on formulæ with quantifiers
- Instantiation heuristics in SMT-solvers:
  incomplete, fragile, demanding on users

Interpolation system

- Given refutation of $A \cup B$ extracts interpolant of (A, B)
- Associates *partial interpolant* to every clause in refutation
- Every inference: Partial interpolant of conclusion defined *inductively* from those of premises
- Partial interpolant of $\Box$ is interpolant of (A, B)
- *Complete* for inference system: for all its refutations gives interpolant
- Proof of completeness: show that partial interpolants are such

What is a partial interpolant?

- C occurs in a refutation of $A \wedge B$
- $A \wedge B \vdash C$
- $A \wedge B \wedge \neg C \vdash \perp$
- Interpolant of $A \wedge \neg C$ and $B \wedge \neg C$
- $(A \wedge \neg C, B \wedge \neg C)$ may share more symbols than (A, B)
- Use *projections*

Colors

Symbol:

- A-colored: in A but not in B
- B-colored: in B but not in A
- *Transparent*: in both

Term/atom/literal/clause:

- *Transparent*: all symbols transparent
- A-colored: all symbols in A and at least one A-colored
- B-colored: all symbols in B and at least one B-colored
- Otherwise: AB-*mixed*

Projections and Partial Interpolants

- $C|_A$: A-colored and transparent literals of C
- $C|_B$: B-colored and transparent literals of C
- $\perp$ if empty

- *Partial interpolant* $PI(C)$:
  interpolant of $A \wedge \neg(C|_A)$ and $B \wedge \neg(C|_B)$:
  - $A \wedge \neg(C|_A) \vdash PI(C)$
  - $B \wedge \neg(C|_B) \wedge PI(C) \vdash \perp$
  - $PI(C)$ transparent

Interpolation for propositional/ground resolution

All literals are input literals:
either A-colored or B-colored or *transparent*:

- $C \in A$: $PI(C) = \bot$
- $C \in B$: $PI(C) = \top$
- $C \vee D$ resolvent of $p_1 : L \vee C$ and $p_2 : \neg L \vee D$:
  - L is A-colored: $PI(C \vee D) = PI(p_1) \vee PI(p_2)$
  - L is B-colored: $PI(C \vee D) = PI(p_1) \wedge PI(p_2)$
  - L is transparent: $PI(C \vee D) = (L \vee PI(p_1)) \wedge (\neg L \vee PI(p_2))$

How about non-ground proofs?

Inferences apply substitutions: mix symbols and therefore colors:

- Even in inferences *with no AB-mixed literals*:
  $\neg P(x, b) \vee C$ and $P(a, y) \vee D$ with $x \notin Var(C)$, $y \notin Var(D)$,
  $\sigma = \{x \leftarrow a, y \leftarrow b\}$ yield $(C \vee D)\sigma$:
  $\neg P(a, b)$ and $P(a, b)$ AB-mixed:
  case analysis of ground interpolation system does not suffice

How about non-ground proofs?

Inferences apply substitutions: mix symbols and therefore colors:

- Even in inferences *with no AB-mixed literals*:
  $\neg P(x, b) \vee C$ and $P(a, y) \vee D$ with $x \notin Var(C)$, $y \notin Var(D)$,
  $\sigma = \{x \leftarrow a, y \leftarrow b\}$ yield $(C \vee D)\sigma$:
  $\neg P(a, b)$ and $P(a, b)$ AB-mixed:
  case analysis of ground interpolation system does not suffice
- Even in inferences *involving only one color*:
  $p_1 : L(x, a) \vee R(x)$ and $p_2 : \neg L(c, y) \vee Q(y)$ with
  $\sigma = \{x \leftarrow c, y \leftarrow a\}$ yield $R(c) \vee Q(a)$:
  $(PI(p_1) \vee PI(p_2))\sigma$ not necessarily transparent!

A two-stage approach

Idea: separate entailment and transparency requirements.

- **I stage:** provisional interpolant $\hat{I}$
  such that $A \vdash \hat{I}$ and $B, \hat{I} \vdash \perp$
  may contain colored symbols
- **II stage:** apply *lifting* to turn $\hat{I}$ in interpolant
  by replacing colored terms by quantified variables

Labels track where literals come from

- $c\colon C \in A$: for all $L \in C$, $label(L, c) = \mathbf{A}$
- $c\colon C \in B$: for all $L \in C$, $label(L, c) = \mathbf{B}$
- $c\colon (C \vee D)\sigma$ resolvent of $p_1\colon L \vee C$ and $p_2\colon \neg L' \vee D$:
  $m \in C$: $label(m\sigma, c) = label(m, p_1)$
  $m \in D$: $label(m\sigma, c) = label(m, p_2)$
- $c\colon (C \vee L[r] \vee D)\sigma$ paramodulant of $p_1\colon s \simeq r \vee C$ into $p_2\colon L[s'] \vee D$:
  $label(L[r]\sigma, c) = label(L[s'], p_2)$

Factoring, Superposition, Simplification, Equational Factoring

Labeled Projections and Provisional Partial Interpolants

- $C|_{\mathbf{A}}$: literals of C labeled **A**
- $C|_{\mathbf{B}}$: literals of C labeled **B**
- $\bot$ if empty

Commute with substitutions: resolvent $(C \vee D)\sigma$
$(C \vee D)\sigma|_{\mathbf{A}} = (C|_{\mathbf{A}} \vee D|_{\mathbf{A}})\sigma$

- *Provisional partial interpolant* $\widehat{PI}(C)$: provisional interpolant of $A \wedge \neg(C|_{\mathbf{A}})$ and $B \wedge \neg(C|_{\mathbf{B}})$:
  - $A \wedge \neg(C|_A) \vdash \widehat{PI}(C)$
  - $B \wedge \neg(C|_B) \wedge \widehat{PI}(C) \vdash \bot$
- $\widehat{PI}(\Box)$ provisional interpolant of (A, B)

Provisional interpolation system

- $C \in A$: $\widehat{PI}(C) = \bot$
- $C \in B$: $\widehat{PI}(C) = \top$
- $c: (C \vee D)\sigma$ resolvent of $p_1: L \vee C$ and $p_2: \neg L' \vee D$:
  - $label(L, p_1) = label(\neg L', p_2) = \mathbf{A}$: $\widehat{PI}(c) = (\widehat{PI}(p_1) \vee \widehat{PI}(p_2))\sigma$
  - $label(L, p_1) = label(\neg L', p_2) = \mathbf{B}$: $\widehat{PI}(c) = (\widehat{PI}(p_1) \wedge \widehat{PI}(p_2))\sigma$
  - $label(L, p_1) = \mathbf{A}$ and $label(\neg L', p_2) = \mathbf{B}$: $\widehat{PI}(c) = [(L \vee \widehat{PI}(p_1)) \wedge \widehat{PI}(p_2)]\sigma$
  - $label(L, p_1) = \mathbf{B}$ and $label(\neg L', p_2) = \mathbf{A}$: $\widehat{PI}(c) = [\widehat{PI}(p_1) \wedge (\neg L' \vee \widehat{PI}(p_2))]\sigma$

Provisional interpolation system

- $c\colon (C \vee L[r] \vee D)\sigma$ paramodulant of $p_1\colon s \simeq r \vee C$ and $p_2\colon L[s'] \vee D$:
  - $label(s \simeq r, p_1) = label(L[s'], p_2) = \mathbf{A}$: $\widehat{PI}(c) = (\widehat{PI}(p_1) \vee \widehat{PI}(p_2))\sigma$
  - $label(s \simeq r, p_1) = label(L[s'], p_2) = \mathbf{B}$: $\widehat{PI}(c) = (\widehat{PI}(p_1) \wedge \widehat{PI}(p_2))\sigma$
  - $label(s \simeq r, p_1) = \mathbf{A}$ and $label(L[s'], p_2) = \mathbf{B}$: $\widehat{PI}(c) = [(s \simeq r \vee \widehat{PI}(p_1)) \wedge \widehat{PI}(p_2)]\sigma$
  - $label(s \simeq r, p_1) = \mathbf{B}$ and $label(L[s'], p_2) = \mathbf{A}$: $\widehat{PI}(c) = [\widehat{PI}(p_1) \wedge (s \not\simeq r \vee \widehat{PI}(p_2))]\sigma$

A complete provisional interpolation system

- Build provisional interpolant mostly by adding instances of **A**-labeled literals resolved or paramodulated with **B**-labeled ones: *communication interface*
- **Theorem:** Provisional interpolation system is complete
- **Lemma:** Provisional interpolants are in negation normal form with $\forall$-quantified variables and transparent predicate symbols

II stage: lifting

- Replace term with *A*-colored head by $\exists$-quantified variable
- Replace term with *B*-colored head by $\forall$-quantified variable

- Vars z and w replace t and $s[t]$: w depends on z
- $Q_1 z$ *before* $Q_2 w$: $Q_2 w$ in scope of $Q_1 z$
- $Q_1 = Q_2 = \forall$ or $Q_1 = Q_2 = \exists$: order immaterial
- $Q_1 = \forall$ and $Q_2 = \exists$: use $\forall z . \exists w$ as w depends on z
- $Q_1 = \exists$ and $Q_2 = \forall$: use $\exists z . \forall w$ as z does not depend on w
- **Solution:** look for $\triangleright$-*maximal colorful subterms*

Colorful: *A*-colored or *B*-colored or *AB*-mixed

Lifting

- Given provisional interpolant $F = \forall \bar{v}.G$
- Take list of $\triangleright$-maximal colorful subterms with their positions $MCS(G) = [(t_1, [p^1_1, \ldots, p^1_{n_1}]), \ldots, (t_k, [p^k_1, \ldots, p^k_{n_k}])]$ in *decreasing order* by $\triangleright$
- $Lift(F) = \forall \bar{v}.lifting(G, MCS(G))$
- $lifting(G, [\,]) = G$
  $lifting(G, (t_1, [p^1_1, \ldots, p^1_{n_1}]) :: T) =$
  $lifting(Q_1 x_1.G[x_1]_{p^1_1} \ldots [x_1]_{p^1_{n_1}}, T)$
- Head symbol of t_i A-colored: $Q_i = \exists$
  Head symbol of t_i B-colored: $Q_i = \forall$

A complete interpolation system

- **Theorem:** Lifting of provisional interpolant $\hat{I}$ of (A, B) is interpolant:
  - **Lemma:** $B, \hat{I} \vdash \perp$ implies $B, Lift(\hat{I}) \vdash \perp$
    bwoc: assume $B, Lift(\hat{I})$ has model ...
    A-colored symbols new, interpreted as $\exists$-vars
  - **Lemma:** $A, \neg\hat{I} \vdash \perp$ implies $A, \neg Lift(\hat{I}) \vdash \perp$
    bwoc: assume $A, \neg Lift(\hat{I})$ has model ...
    B-colored symbols new, interpreted as $\exists$-vars
  - $Lift(\hat{I})$ is transparent
- **Corollary:** Provisional interpolation system + lifting = complete interpolation system

Discussion

- Color-based approaches do not extend beyond ground
- First *complete* interpolation system for *general* refutations by standard inference system for FOL+=
- Current work: interpolation for DPLL($\Gamma + \mathcal{T}$)
- Future work:
  - Quality of interpolants (length? strength? min quantifiers?)
  - Implementation and experiments