

Automated reasoning 2016

Lecture 5: Interpolants

Instructor: Ashutosh Gupta

TIFR, India

Compile date: 2016-04-26

Where are we and where are we going?

We have seen

- ▶ proof generation

We will see

- ▶ computing interpolants using proofs

Topic 5.1

Interpolants

Interpolants

Definition 5.1

For mutually unsat formulas A and B , a formula I is *interpolant* between A and B if

1. $A \Rightarrow I$,
2. $B \wedge I \Rightarrow \perp$, and
3. $\text{vars}(I) \subseteq (\text{vars}(A) \cap \text{vars}(B))$

Example 5.1

Consider:

$$A = x_1 + x_2 \leq 2 \wedge x_3 - x_2 \leq 0$$

$$B = 6x_4 - 2x_1 \leq -8 \wedge -3x_4 - x_3 \leq 0$$

$$\text{vars}(A) = \{x_1, x_2, x_3\} \quad \text{vars}(B) = \{x_1, x_3, x_4\} \quad \text{vars}(I) \subseteq \{x_1, x_3\}$$

$$I = x_1 + x_3 \leq 2$$

A-local, B-local, global symbols

Definition 5.2

*A symbol is **A-local** if it only occurs in A.*

Definition 5.3

*A symbol is **B-local** if it only occurs in B.*

Definition 5.4

*A symbol is **global** if it occurs in both A and B.*

The set of global symbols is written $\text{Globals} = \text{vars}(A) \cap \text{vars}(B)$

Interpolants - equivalent definition

Definition 5.5

Let $A \Rightarrow B$. A formula I is *interpolant* between A and B if

1. $A \Rightarrow I$,
2. $I \Rightarrow B$, and
3. $\text{vars}(I) \subseteq \text{Globals}$

Exercise 5.1

Show both the definitions are equivalent.

In this form, interpolant captures a general idea behind every human inquiry.

Example 5.2

$\text{Human activity} \Rightarrow \underbrace{\text{Increase in CO}_2 \text{ level}}_{\text{interpolant}} \Rightarrow \text{Climate change}$

A concise explanation of cause and effect

Interpolants from proofs

Back to the original definition of interpolants.

We seek interpolants for mutually unsat A and B .

Interpolant captures the interactions of A and B that makes them mutually unsat.

If we have unsat proof for $A \wedge B$, we may use the reasoning that proves $A \wedge B$ to find a interpolant.

As proofs have intermediate results, we should have equivalent concept of intermediate interpolants.

Partial interpolants

Definition 5.6

Let A, B and C be formulas such that $A \wedge B \Rightarrow C$.

A *partial interpolant* I_C between A and B for C is a formula such that

- ▶ $A \Rightarrow I_C$
- ▶ $B \wedge I_C \Rightarrow C$
- ▶ $\text{vars}(I_C) \subseteq \text{Globals} \cup \text{vars}(C)$

Interpolation via proofs

We present here a proof based method for computing interpolants.

Consider A and B are mutually unsat formulas in some given theory.

1. produce proof of unsatisfiability of $A \wedge B$
2. annotate each intermediate derived formulas with **partial interpolants** inductively
3. Since $\perp$ is the final node of the proof, the annotation of $\perp$ is the interpolant between A and B .

We will write annotations in proof rules within square brackets after derived formulas.

$$\cdots \frac{\cdots}{C[I_C]} \cdots$$

Topic 5.2

Interpolation in $\mathcal{T}_{LRA}$

Annotation rules for conjunction of linear arithmetic formulas

Linear arithmetic proof system

$$\text{hyp} \frac{}{aX \leq c} aX \leq c \in A, B \quad \text{comb} \frac{a_1X \leq c_1 \quad a_2X \leq c_2}{(\lambda_1 a_1 + \lambda_2 a_2)X \leq (\lambda_1 c_1 + \lambda_2 c_2)} \lambda_1, \lambda_2 \geq 0$$

Proof rules with partial interpolant annotations

$$\text{HYP-A} \frac{}{aX \leq c[aX \leq c]} aX \leq c \in A \quad \text{HYP-B} \frac{}{aX \leq c[0 \leq 0]} aX \leq c \in B$$

$$\text{COMB} \frac{a_1X \leq c_1[a'_1X \leq c'_1] \quad a_2X \leq c_2[a'_2X \leq c'_2]}{(\lambda_1 a_1 + \lambda_2 a_2)X \leq (\lambda_1 c_1 + \lambda_2 c_2)[(\lambda_1 a'_1 + \lambda_2 a'_2)X \leq (\lambda_1 c'_1 + \lambda_2 c'_2)]} \lambda_1 \geq 0, \lambda_2 \geq 0$$

Example: LRA annotations

Example 5.3

Consider:

$$A = x_1 + x_2 \leq 2 \wedge x_3 - x_2 \leq 0$$

$$B = 6x_4 - 2x_1 \leq -8 \wedge -3x_4 - x_3 \leq 0$$

$$\frac{x_1 + x_2 \leq 2[x_1 + x_2 \leq 2] \quad 6x_4 - 2x_1 \leq -8[0 \leq 0] \quad x_3 - x_2 \leq 0[x_3 - x_2 \leq 0] \quad -3x_4 - x_3 \leq 0[0 \leq 0]}{\frac{3x_4 + x_2 \leq -2[x_1 + x_2 \leq 2] \quad -3x_4 - x_2 \leq 0[x_3 - x_2 \leq 0]}{0 \leq -2[x_1 + x_3 \leq 2]}}$$

Annotation correctness

Theorem 5.1

The annotations in the proof rules are partial interpolants

Proof.

We show the following facts for annotations that imply the conditions of partial interpolants.

Let $a'X \leq c'$ be the annotation for $aX \leq c$.

1. $A \Rightarrow a'X \leq c'$, //1st cond. of partial interpolants
2. $B \Rightarrow (a - a')X \leq (c - c')$, //implies 2nd cond. (why?)
3. B-locals do not occur in $a'X \leq c'$, and
4. A-locals have same coefficient in $aX \leq c$ and $a'X \leq c'$. //3-4 imply 3rd cond.

Annotation correctness(contd.)

Proof(contd.)



base case:

$$\text{HYP-A} \frac{}{aX \leq c[aX \leq c]} aX \leq c \in A$$

1. $A \Rightarrow aX \leq c$
2. $B \Rightarrow (a-a)X \leq (c-c)$
3. B-locals do not occur in $aX \leq c$, and
4. A-locals have same coefficient in $aX \leq c$ and $a'X \leq c'$

$$\text{HYP-B} \frac{}{aX \leq c[0 \leq 0]} aX \leq c \in B$$

1. $A \Rightarrow 0 \leq 0$
2. $B \Rightarrow (a-0)X \leq (c-0)$
3. B-locals do not occur in $0 \leq 0$, and
4. A-locals have same coefficient in $0 \leq 0$ and $aX \leq c$

Annotation correctness(contd.)

Proof(contd.)

induction step:

$$\text{COMB} \frac{a_1X \leq c_1[a'_1X \leq c'_1] \quad a_2X \leq c_2[a'_2X \leq c'_2] \quad \lambda_1 \geq 0 \quad \lambda_2 \geq 0}{(\lambda_1 a_1 + \lambda_2 a_2)X \leq (\lambda_1 c_1 + \lambda_2 c_2)[(\lambda_1 a'_1 + \lambda_2 a'_2)X \leq (\lambda_1 c'_1 + \lambda_2 c'_2)]}$$

Due to ind. hyp., the conditions of the partial interpolants of the antecedents holds.

- | | |
|--|--|
| 1. $A \Rightarrow a'_1X \leq c'_1$ | 1. $A \Rightarrow a'_2X \leq c'_2$ |
| 2. $B \Rightarrow (a_1 - a'_1)X \leq (c_1 - c'_1)$ | 2. $B \Rightarrow (a_2 - a'_2)X \leq (c_2 - c'_2)$ |
| 3. No B-locals in $a'_1X \leq c'_1$ | 3. No B-locals in $a'_2X \leq c'_2$ |
| 4. A-locals have same coefficient in $a_1X \leq c_1$ and $a'_1X \leq c'_1$. | 4. A-locals have same coefficient in $a_2X \leq c_2$ and $a'_2X \leq c'_2$. |

Rest is exercises



Exercise 5.2

Show the above four properties hold for the annotation of the consequent (Only proving 4th condition is non-trivial.)

Topic 5.3

Interpolation in Propositional logic

Some notation

For a clause C and a formula B , let

$$C|_B = \{\ell \in C \mid \text{vars}(\ell) \subseteq \text{vars}(B)\}$$

and

$$C/_B = \{\ell \in C \mid \text{vars}(\ell) \not\subseteq \text{vars}(B)\}.$$

Theorem 5.2

$$C = C|_B \vee C/_B$$

Example 5.4

Consider formula $B = (p \vee q) \wedge (q \wedge r)$.

Let $C = p \vee \neg r \vee \neg x \vee z$.

- ▶ $C|_B = p \vee \neg r$
- ▶ $C/_B = \neg x \vee z$

Annotation rules for propositional formulas

Resolution proof system

$$\text{HYP} \frac{}{C \in A, B} \quad \text{RES} \frac{C \vee x \quad D \vee \neg x}{C \vee D}$$

Proof rules with partial interpolant annotations

$$\begin{aligned} \text{HYP-A} & \frac{}{C[C|_B \vee C/B]} C \in A & \text{HYP-B} & \frac{}{C[\top \vee C/B]} C \in B \\ \text{RES-B} & \frac{C \vee x[I_1 \vee C/B] \quad D \vee \neg x[I_2 \vee D/B]}{C \vee D[(I_1 \wedge I_2) \vee (C \vee D)/B]} x \in \text{vars}(B) \\ \text{RES-A} & \frac{C \vee x[I_1 \vee (C/B \vee x)] \quad D \vee \neg x[I_2 \vee (D/B \vee \neg x)]}{C \vee D[(I_1 \vee I_2) \vee (C \vee D)/B]} x \notin \text{vars}(B) \end{aligned}$$

Simplified annotations

Sometimes C/B part of a partial interpolant is not mentioned and implicitly assumed to be around. The **shorthanded** annotation rules are as follows.

$$\text{HYP-A} \frac{}{C[C|_B]} C \in A \quad \text{HYP-B} \frac{}{C[\top]} C \in B$$

$$\text{RES-B} \frac{C \vee x[I_1] \quad D \vee \neg x[I_2]}{C \vee D[(I_1 \wedge I_2)]} x \in \text{vars}(B)$$

$$\text{RES-A} \frac{C \vee x[I_1] \quad D \vee \neg x[I_2]}{C \vee D[(I_1 \vee I_2)]} x \notin \text{vars}(B)$$

Example: interpolation for propositional formulas

Example 5.5

$$A = (p \vee \neg q) \wedge (\neg p \vee \neg r) \quad B = (\neg q \vee r) \wedge (q \vee s) \wedge \neg s$$

$$\frac{\frac{\frac{\neg p \vee \neg r[\neg r] \quad \neg q \vee r[\top]}{\neg p \vee \neg q[\neg r]} \quad p \vee \neg q[\neg q]}{\neg q[\neg r \vee \neg q]} \quad q \vee s[\top]}{s[\neg r \vee \neg q] \quad \neg s[\top]}{\perp[\neg r \vee \neg q]}$$

The interpolant between A and B is $\neg r \vee \neg q$

Exercise 5.3

Compute interpolants between the following pairs of formulas

1. $(q \vee p \vee t) \wedge \neg q \wedge (q \vee p \vee \neg u)$ and $(u \vee \neg t \vee \neg s) \wedge s \wedge \neg p$
2. $(q \vee p \vee t) \wedge \neg q \wedge (q \vee p \vee \neg u) \wedge \neg p$ and $(u \vee \neg t \vee \neg s) \wedge s$
3. $(q \vee p \vee t) \wedge \neg q \wedge (q \vee p \vee \neg u) \wedge \neg p \wedge (u \vee \neg t \vee \neg s)$ and s

Resolution Annotation correctness

Theorem 5.3

The annotations in the above proof rules are partial interpolants

Proof.

We will prove stronger conditions than the partial interpolation conditions.

Assume $(I \vee C/B)$ is an annotation of C (non-shorthanded annotations).

1. $A \Rightarrow (I \vee C/B)$
2. $B \wedge I \Rightarrow C|_B$, which implies the second condition $B \wedge (I \vee C/B) \Rightarrow C_{(\text{why?})}$
3. $I \subseteq \text{Globals}$, which implies the 3rd condition

base case:

$$\text{HYP-A} \frac{}{C[C|_B \vee C/B]} C \in A$$

1. $A \Rightarrow C|_B \vee C/B$
2. $B \wedge C|_B \Rightarrow C|_B$
3. $\text{vars}(C|_B) \subseteq \text{Globals}_{(\text{why?})}$

$$\text{HYP-B} \frac{}{C[\top \vee C/B]} C \in B$$

1. $A \Rightarrow \top \vee C/B$
2. $B \wedge \top \Rightarrow C|_B_{(\text{why?})}$
3. $\text{vars}(\top) \subseteq \text{Globals}$

Resolution Annotation correctness(contd.)

Proof(contd.)

induction step:

$$\text{RES-B} \frac{C \vee x[I_1 \vee C/B] \quad D \vee \neg x[I_2 \vee D/B]}{C \vee D[(I_1 \wedge I_2) \vee (C \vee D)/B]} x \in \text{vars}(B)$$

1. Since $A \Rightarrow I_1 \vee C/B$ and $A \Rightarrow I_2 \vee D/B$, $A \Rightarrow (I_1 \wedge I_2) \vee (C \vee D)/B$. (why?)
2. Since $B \wedge I_1 \Rightarrow (C \vee x)|_B$ and $B \wedge I_2 \Rightarrow (D \vee \neg x)|_B$,

$$B \wedge I_1 \wedge I_2 \Rightarrow (C|_B \vee x) \wedge (D|_B \vee \neg x) \Rightarrow (C \vee D)|_B.$$

3. $\text{vars}(I_1 \wedge I_2) \subseteq \text{Globals}$

Resolution Annotation correctness(contd.)

Proof(contd.)

$$\text{RES-A} \frac{C \vee x[I_1 \vee (C/B \vee x)] \quad D \vee \neg x[I_2 \vee (D/B \vee \neg x)]}{C \vee D[(I_1 \vee I_2) \vee (C \vee D)/B]} x \notin \text{vars}(B)$$

1. Since $A \Rightarrow I_1 \vee C/B \vee x$ and $A \Rightarrow I_2 \vee D/B \vee \neg x$,

$$A \Rightarrow (I_1 \vee I_2) \vee (C \vee D)/B. (\text{why?})$$

2. Since $B \wedge I_1 \Rightarrow C|_B$ and $B \wedge I_2 \Rightarrow D|_{B(\text{why?})}$,

$$B \wedge (I_1 \vee I_2) \Rightarrow C|_B \vee D|_B.$$

3. $\text{vars}(I_1 \vee I_2) \subseteq \text{Globals}$.

Topic 5.4

Problems

Asymmetric annotation rules

Exercise 5.4

Consider the following equivalent definition of interpolants

1. $A \Rightarrow I$,
2. $B \Rightarrow \neg I$, and
3. $\text{vars}(I) \subseteq (\text{vars}(A) \cap \text{vars}(B))$

The above definition is symmetric upto negation, but the following annotation rules are not.

$$\text{HYP-A} \frac{}{C[C|_B]} C \in A \quad \text{HYP-B} \frac{}{C[\top]} C \in B$$

$$\text{RES-B} \frac{C \vee x[I_1] \quad D \vee \neg x[I_2]}{C \vee D[(I_1 \wedge I_2)]} x \in \text{vars}(B)$$

$$\text{RES-A} \frac{C \vee x[I_1] \quad D \vee \neg x[I_2]}{C \vee D[(I_1 \vee I_2)]} x \notin \text{vars}(B)$$

Find annotation rules that are mirror to the above rules and symmetry is broken in a similar way. Prove that your rules are also correct.

Symmetric annotation rules

Exercise 5.5

Prove that the following annotation rules will always annotate $\perp$ with the interpolant between A and B .

$$\text{HYP-A} \frac{}{C[\perp]} C \in A \quad \text{HYP-B} \frac{}{C[\top]} C \in B$$

$$\text{RES-B-LOCAL} \frac{C \vee x[I_1] \quad D \vee \neg x[I_2]}{C \vee D[I_1 \wedge I_2]} x \in \text{vars}(B) - \text{Globals}$$

$$\text{RES-A-LOCAL} \frac{C \vee x[I_1] \quad D \vee \neg x[I_2]}{C \vee D[I_1 \vee I_2]} x \in \text{vars}(A) - \text{Globals}$$

$$\text{RES-GLOBAL} \frac{C \vee x[I_1] \quad D \vee \neg x[I_2]}{C \vee D[(I_1 \vee x) \wedge (\neg x \vee I_2)]} x \in \text{Globals}$$

End of Lecture 5