Local Proofs and Interpolants

Krystof Hoder

Laura Kovacs

Andrei Voronkov

Interpolants

Craig's Interpolation Theorem
Let R, B be closed formulas and let $R \vdash B$.

Then there exists a formula I such that

1. $R \vdash I$ and $I \vdash B$;
2. every symbol of I occurs both in R and B;

I is called an **interpolant** of R and B.

Motivation

Bounded model-checking

- checks safety property after N unrollings
- good for finding bugs
- not so good for proving correctness
  - showing that bug isn't in the first N iterations is not enough
- correctness can be proved by finding an invariant
  1) implied by initial states
  2) preserved by transition
  3) implies safety property
- *R* formula contains first few unrollings, *B* the rest together with safety property
  - we get (1) and (3), hope to get (2) as well

$R \vdash I$ and $I \vdash B$

$a_0=1$
$b_0=0$
$a_{i+1}= a_i \dashrightarrow b_i$
$b_{i+1}= b_i \dashrightarrow a_i$
S: $a_n \vee b_n$

$a_0=1$
$b_0=0$
$a_1= a_0 \dashrightarrow b_0$
$b_1= b_0 \dashrightarrow a_0$
$a_2= a_1 \dashrightarrow b_1$
$b_2= b_1 \dashrightarrow a_1$
$a_3= a_2 \dashrightarrow b_2$
$b_3= b_2 \dashrightarrow a_2$
$a_4= a_3 \dashrightarrow b_3$
$b_4= b_3 \dashrightarrow a_3$
$\neg a_4$
$\neg b_4$

we may get either
$a_2= 1 \wedge b_2= 0$ (useless)
or
$a_2 \oplus b_2$ (desider invariant)

Interpolation Through Colors

- There are three colors: blue, red and grey.
- Each symbol (function or predicate) is colored in exactly one of these colors.
- We have two formulas: R and B.
- Each symbol in R is either red or grey.
- Each symbol in B is either blue or grey.
- We know that $\vdash R \rightarrow B$.
- Task of interpolation: find a grey formula I such that
    1. $\vdash R \rightarrow I$;
    2. $\vdash I \rightarrow B$.

Local Proofs

Local proofs: No inference mixes blue and red symbols

- $R := \forall x (x = a)$
- $B := c = b$

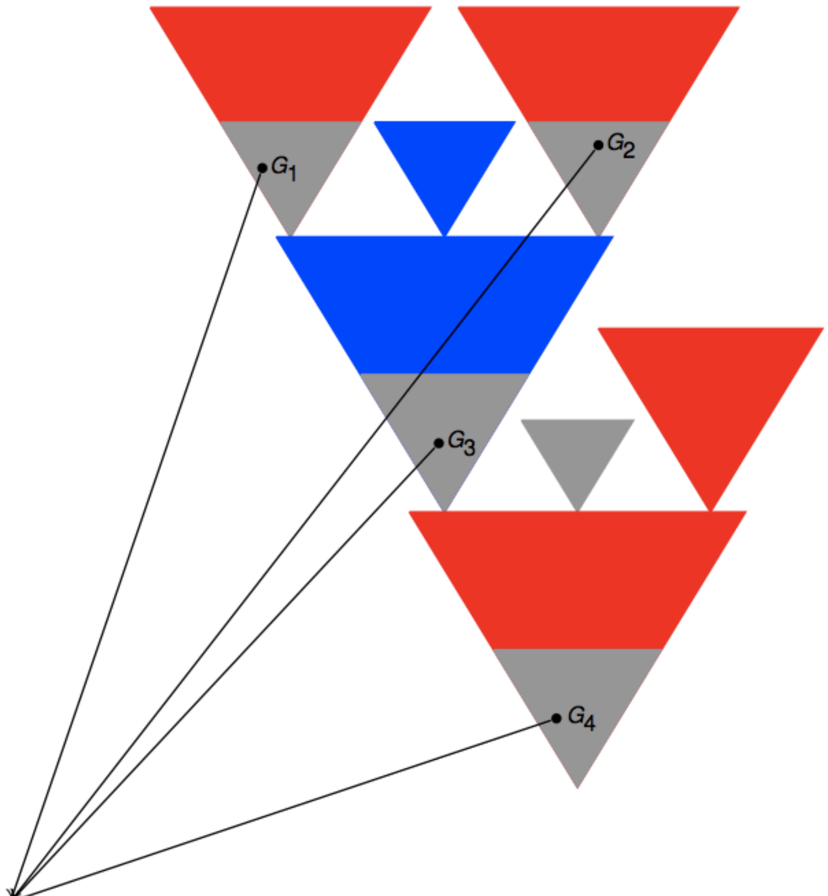
Non-local proof

$$\dfrac{\dfrac{\dfrac{x = a}{c = a} \qquad \dfrac{x = a}{b = a}}{c = b} \qquad c \neq b}{\perp}$$

Local Proof

$$\dfrac{\dfrac{\dfrac{x = a \qquad y = a}{x = y} \qquad c \neq b}{y \neq b}}{\perp}$$

Extracting Interpolants from Local Proofs



Interpolant: boolean combination of $\{G_1, \ldots, G_4\}$

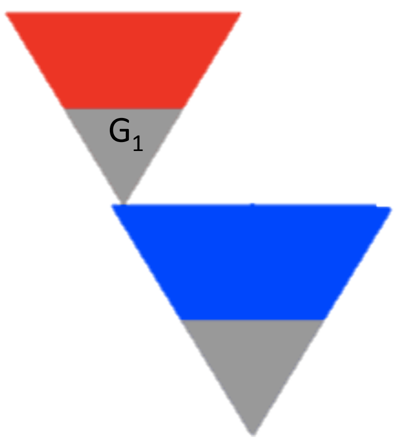
[McMillan05, KV09]

Given an unsatisfiable set $\{R, B\}$.
A **reverse interpolant** I of R and B is a formula such that:

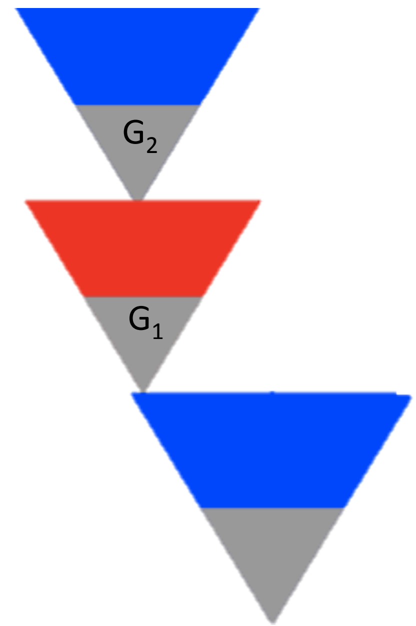
1. $R \vdash I$ and $\{I, B\}$ is unsatisfiable;
2. every symbol of I occurs both in R and B.

Basic Idea

Easy case:
Contradiction follows from R, so interpolant is $\bot$

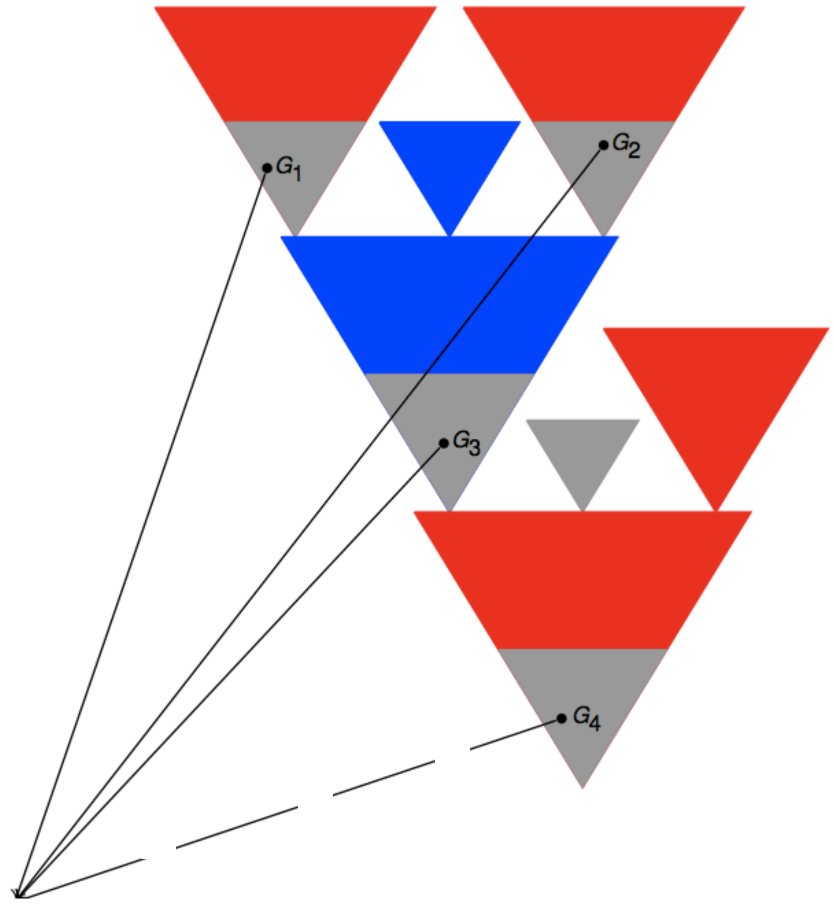


Still quite easy:
G_1 is interpolant as it follows from R and is unsat with B



A bit more subtle:
$\{G_1, B\}$ is unsat, but G_1 but doesn't follow from R alone. However it follows from R $\wedge$ G_2, and G_2 follows from B.
Therefore G_2 --> G_1 is an interpolant.

Extracting Interpolants from Local Proofs



$G_1 \wedge G_2 \wedge \neg G_3$

Proof Localization

- Not many tools generate local proofs
  - most SMT solvers don't output any proofs at all
- Under few reasonable conditions proofs can be localized
  - only constants are colored
  - input formulas do not mix colors
- We can quantify away the colored symbols

Given $R(a) \vdash B$ where a is an uninterpreted constant not occurring in B.

Then, $R(a) \vdash (\exists x) R(x)$ and $(\exists x) R(x) \vdash B$.

Proof Localization

Given $R(a) \vdash B$ where a is an uninterpreted constant not occurring in B.

Then, $R(a) \vdash (\exists x)R(x)$ and $(\exists x)R(x) \vdash B$.

- Naïve approach
  - quantify away all colored symbols in R and get interpolant $(\exists x)R(x)$
  - does not give a "nice" interpolant

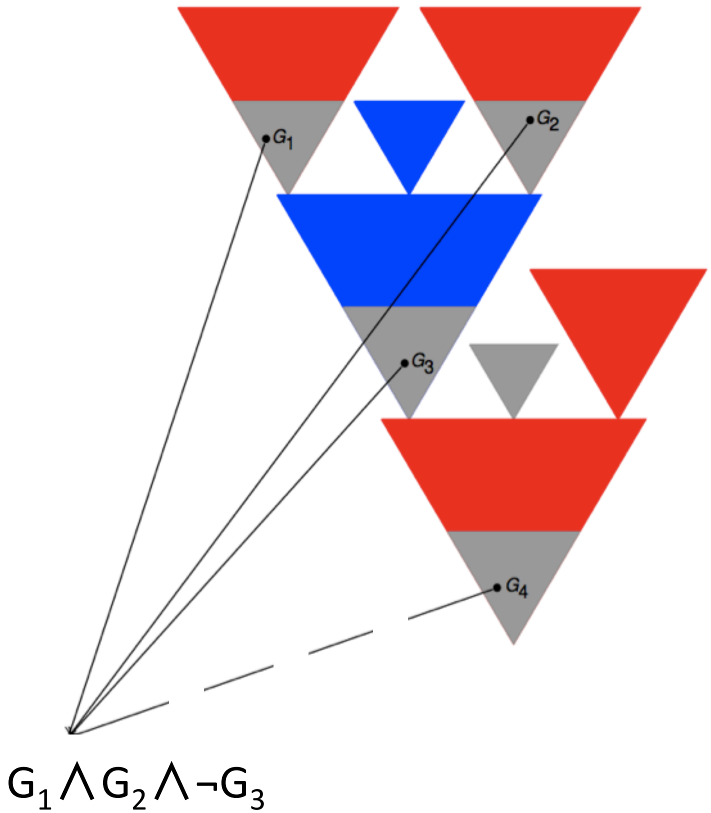
$(\exists x_0, y_0, x_1, y_1)\ (x_0=1 \wedge y_0=0 \wedge x_1=x_0 \dashrightarrow y_0 \wedge y_1=y_0 \dashrightarrow x_0 \wedge a_2=x_1 \dashrightarrow y_1 \wedge b_2=y_1 \dashrightarrow x_1)$

- Detect non-local parts of the proof and try to localize locally

$$\dfrac{\dfrac{R_1(a)}{R_2(a)} \quad B}{A} \quad \Bigg\| \quad \dfrac{\dfrac{R_1(a)}{(\exists x)R_2(x)} \quad B}{A}$$

- May still require non-local transformations

Interpolant Minimization



- $G_1,\ldots,G_4$ are conclusions of *symbol-eliminating inferences*
  - their premises are colored, they themselves not (i.e. they are grey)
- A subset of sym-el formulas forms digest, the set of formulas used in the interpolant
- We try to modify the proof so that different formulas appear in the digest

Proof Transformations

Idea: Change the grey areas of the local proof

Slicing off formulas

$$\dfrac{A_1 \quad \cdots \quad A_n \qquad \dfrac{A_{n+1} \quad \cdots \quad A_m}{A}}{A_0} \quad \overset{\text{slicing off } A}{\longrightarrow} \quad \dfrac{A_1 \quad \cdots \quad A_n \quad A_{n+1} \quad \cdots \quad A_m}{A_0}$$

If A is grey: Grey slicing

Proof Transformations

Idea: Change the grey areas of the local proof, but preserve locality!

Slicing off formulas

$$\dfrac{B_0 \quad \dfrac{R_0}{G_1}}{G_0} \quad \xrightarrow{\text{slicing off } G_1} \quad \dfrac{B_0 \quad R_0}{G_0}$$

Proof Transformations

$$\dfrac{\dfrac{\dfrac{R_3 \qquad \dfrac{\dfrac{\dfrac{R_1 \quad G_1}{G_3} \qquad \dfrac{B_1 \quad G_2}{G_4}}{G_5}}{G_6}}{R_4}}{G_7}}{\bot}$$

Digest: $\{G_4, G_7\}$

Reverse interpolant: $G_4 \rightarrow G_7$

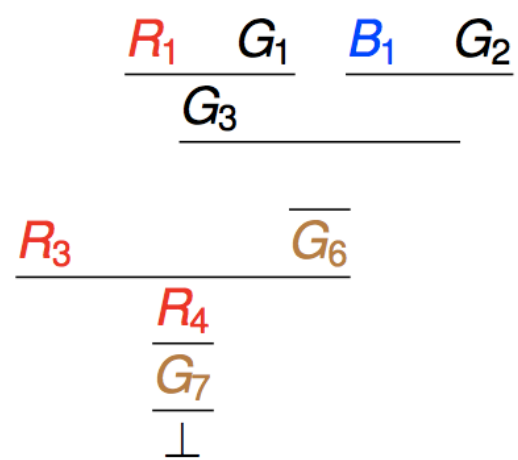
Proof Transformations

$$\dfrac{\dfrac{\dfrac{R_3 \qquad \dfrac{\dfrac{\dfrac{R_1 \quad G_1}{G_3} \qquad \dfrac{B_1 \quad G_2}{}}{G_5}}{G_6}}{R_4}}{G_7}}{\perp}$$

Digest: $\{G_5, G_7\}$

Reverse interpolant: $G_5 \rightarrow G_7$

Proof Transformations



Digest: $\{G_6, G_7\}$

Reverse interpolant: $G_6 \rightarrow G_7$

Proof Transformations

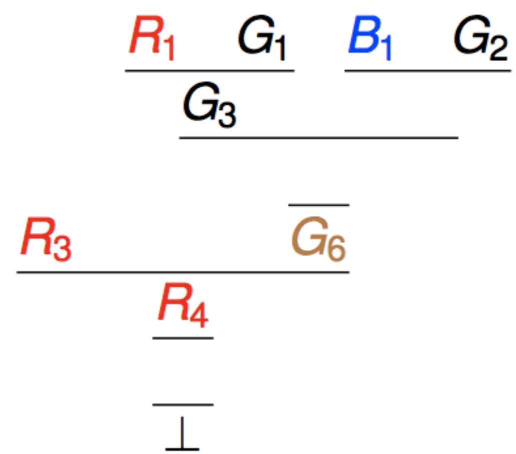
$$\cfrac{\cfrac{R_1 \quad G_1}{G_3} \qquad \cfrac{B_1 \quad G_2}{}}{}$$

$$\cfrac{\cfrac{R_3 \qquad \overline{G_6}}{R_4}}{\overline{\underline{\bot}}}$$

Digest: $\{G_6\}$

Reverse interpolant: $\neg G_6$

Proof Transformations



Note that the interpolant has changed from $G_4 \rightarrow G_7$ to $\neg G_6$.

- There is no obvious logical relation between $G_4 \rightarrow G_7$ and $\neg G_6$, for example none of these formulas implies the other one;
- These formulas may even have no common atoms or no common symbols.

Interpolant Minimization

If grey slicing gives us very different interpolants, we can use it for finding small interpolants.

Problem: if the proof contains n grey formulas, the number of possible different slicing off transformations is 2^n.

Solution:

- encode all sequences of transformations as an instance of SAT

Interpolant Minimization

Solution:

- encode all sequences of transformations as an instance of SAT

$$\frac{\dfrac{R}{G_1} \quad \dfrac{B}{G_2}}{G_3}$$

Some predicates on grey formulas:

- sliced(G): G was sliced off;
- red(G): the trace of G contains a red formula;
- blue(G): the trace of G contains a blue formula;
- grey(G): the trace of G contains only grey formulas;
- digest(G): G belongs to the digest.

Interpolant Minimization

Solution:

- encode all sequences of transformations as an instance of SAT
- solutions encode all slicing off transformations

$$\cfrac{\cfrac{R}{G_1} \quad \cfrac{B}{G_2}}{G_3}$$

Some predicates on grey formulas:

- $\text{sliced}(G)$: G was sliced off;
- $\text{red}(G)$: the trace of G contains a red formula;
- $\text{blue}(G)$: the trace of G contains a blue formula;
- $\text{grey}(G)$: the trace of G contains only grey formulas;
- $\text{digest}(G)$: G belongs to the digest.

$$\neg\text{sliced}(G_1) \rightarrow \text{grey}(G_1)$$
$$\text{sliced}(G_1) \rightarrow \text{red}(G_1)$$
$$\neg\text{sliced}(G_3) \rightarrow \text{grey}(G_3)$$
$$\text{sliced}(G_3) \rightarrow (\text{grey}(G_3) \leftrightarrow \text{grey}(G_1) \wedge \text{grey}(G_2))$$
$$\text{sliced}(G_3) \rightarrow (\text{red}(G_3) \leftrightarrow \text{red}(G_1) \vee \text{red}(G_2))$$
$$\text{sliced}(G_3) \rightarrow (\text{blue}(G_3) \leftrightarrow \text{blue}(G_1) \vee \text{blue}(G_2))$$
$$\text{digest}(G_1) \rightarrow \neg\text{sliced}(G_1)$$
$$\ldots$$

Interpolant Minimization

Solution:

- encode all sequences of transformations as an instance of SAT;
- solutions encode all slicing off transformations;
- compute small interpolants: smallest digest of grey formulas;

$$\mathtt{min}_{\{G_{i_1},\ldots,G_{i_n}\}}\left(\sum_{G_i}\mathrm{digest}(G_i)\right)$$

$$\mathtt{min}_{\{G_{i_1},\ldots,G_{i_n}\}}\left(\sum_{G_i}\mathrm{quantifier\_number}(\mathrm{G_i})\ \mathrm{digest}(G_i)\right)$$

- use a pseudo-boolean optimisation tool or an SMT solver to minimise interpolants;
- minimising interpolants is an NP-complete problem.

Conclusion

- We localise proofs by quantifying away colored constants;
- We minimise interpolants by:
  - expressing constraints on grey formulas;
  - finding a minimal interpolants as a solution to the constraint system;
- Experiments show that interpolants become smaller in size, weight, or number of quantifiers;
  - 9632 first-order examples from the TPTP library:

    for example, for 2000 problems the size of the interpolants became 20-49 times smaller;
  - 4347 SMT examples:
    - we used Z3 for proving SMT examples;
    - Z3 proofs were localised in Vampire;
    - minimal interpolants were generated for 2123 SMT examples.