

# 1 Overbinding in one step

**Conjecture 1.** *Suppose every variable occurs only once in  $\Gamma \cup \Delta$ . Then the order of the quantifiers for  $\text{PI}(\Box)^*$  does not matter.*

**Proposition 2.** *Let  $A(x_1, \dots, x_n)$  be an atom in a relative interpolant. A variable occurs in one of the  $x_i$  if and only if there are atoms  $A(y_1, \dots, y_n)$  and  $A(z_1, \dots, z_n)$  in  $\Gamma$  and  $\Delta$  respectively, where  $x_i$  can be unified with  $z_i$  and  $y_i$  such that there is still a variable at that location.*

*This means that either the term structure above the variable is the same in the original clauses or there are some variables. Intended meaning: the original clauses prove at least the  $x_i$ , i.e. are at least as or more general.*

*Special case for outermost variables:*

*Let  $A(x_1, \dots, x_n)$  be an atom in a relative interpolant. An  $x_i$  is a variable if and only if there are atoms  $A(y_1, \dots, y_n)$  and  $A(z_1, \dots, z_n)$  in  $\Gamma$  and  $\Delta$  respectively, where  $y_i$  and  $z_i$  are variables.*

need more narrow version: clauses do appear in parent clauses in derivation.

**Proposition 3.** *Suppose in a partial interpolant, there are two maximal terms  $t_1$  and  $t_2$  such that w.l.o.g.  $t_1$  is smaller (as defined in 5) than  $t_2$ . Then in the final interpolant, an overbinding can be defined where the variable corresponding to  $t_1$  is quantified over before the variable corresponding to  $t_2$  is.*

The subterm-relation is reflexive.

**Definition 4.** (OLD) Let  $s$  be a term that is in  $\text{PI}(C)$  but not in any predecessor  $\text{PI}(C_i)$ ,  $i \in \{1, 2\}$ .  $s$  is smaller than a term  $t$  in  $\text{PI}(C)$  if  $s$  is of strictly smaller length than  $t$  and there is a subterm in  $s$  which also occurs in  $t$ . △

**Definition 5.** (NEW)

OUTDATED! DOES NOT WORK LIKE THIS

Let  $C$  be a clause.

A maximal term  $s$  of  $C$  is smaller than a maximal term  $t$  of  $C$  if  $s$  is a variable and occurs in  $t$ , but  $s \neq t$ .

OUTDATED! DOES NOT WORK LIKE THIS △

## 2 Half-baked approaches

**Definition 6.** Direct interpolation extraction.

This version of overline and star does NOT overbind variables! If they happen to be in the final interpolant, just overbind them somehow, but not earlier. This is ok as the interpolant only contains variables if both corresponding atoms in  $\Gamma$  and  $\Delta$  do. Variables are the only terms in the interpolant that can “change their color”, so we don’t know a priori if there are constraints on the quantifier to overbind them with.

Convention w.r.t. a clause  $C$  which has been derived from  $C_1$  and  $C_2$ :  $\bar{Q}_n = Q_1 z_1 \dots Q_n z_n$ , such that the  $z_i$  correspond to the maximal terms  $t_i$  in  $\text{PI}(C)$ . Same terms must be overbound by same variable, see 101a for counterexample to per-occurrence-overbinding. The  $z_i$  are ordered such that

1. the orderings in the  $Q_{n_1}$  and  $Q_{n_2}$  are respected (no circular relations can occur in combination with merging as a term is only smaller than another term if it is smaller in length as well, which excludes cycles)

2. as well as ordering constraints of terms newly introduced in  $\text{PI}(C)$  (i.e. those that were not present in  $\text{PI}(C_1)$  and  $\text{PI}(C_2)$ ).

Basically, track dependencies and define actual order later.

Resolution.

$$\frac{C_1 : D \vee l \quad C_2 : E \vee \neg l'}{C : (D \vee E)\sigma} \quad \sigma = \text{mgu}(l, l')$$

$\bar{Q}_{n_1} \text{PI}(C_1)^*$

$\bar{Q}_{n_2} \text{PI}(C_2)^*$

1.  $l$  and  $l'$   $\Gamma$ -colored:

$$\text{PI}(C) \equiv (\text{PI}(C_1) \vee \text{PI}(C_2))\sigma$$

$$\text{PI}(C)^* \equiv (\text{PI}(C_1)^* \vee \text{PI}(C_2)^*)\sigma \text{ (just replace maximal terms)}$$

intended meaning of  $\sigma$ : to change the free variables still in the  $\text{PI}(C_i)$

TODO: basically do nothing here since no new atoms (revisit after mixed colored case has been dealt with)

Let  $t_1, \dots, t_{n_1}$  be terms overbound in  $\text{PI}(C_1)$  and  $s_1, \dots, s_{n_2}$  terms overbound in  $\text{PI}(C_2)$ .

$$\{z_1, \dots, z_n\} = \{t_1, \dots, t_{n_1}\}\sigma \cup \{s_1, \dots, s_{n_2}\}\sigma \quad // \text{ common terms are merged}$$

order relations as in  $C_1, C_2$

$$\bar{Q}_n \text{PI}(C)^* \equiv \bar{Q}_n(\text{PI}(C_1)^* \vee \text{PI}(C_2)^*)$$

2.  $l$  and  $l'$   $\Delta$ -colored:

similar to first case

3.  $l$  and  $l'$  grey: nothing here

$\triangle$

### 3 current proof attempts

**Lemma 7.** *If an atom  $A$  appears in the interpolant, it appeared in both original clause sets, once positively and once negatively.*

*$A$  is contained in some instance of the respective clauses in  $\Gamma$  and  $\Delta$ .*

**Lemma 8.** *Let  $C \in \Phi$  for some initial clause set  $\Phi$ .*

1. *Let  $x$  be an occurrence of a variable in  $C$  and  $x'$  another occurrence of the same variable in a different position but at the same term depth. Then  $\Phi \models QyC[x/y][x'/y]$  for  $Q \in \{\forall, \exists\}$ .*
2. *Let  $x$  be an occurrence of a variable in  $C$  with the lowest depth and  $x'$  another occurrence of the same variable with a higher depth. Let  $t$  be the maximal colored term which contains  $x'$ .  $t$  is  $\Phi$ -colored since it appears in  $\Phi$ . Then  $\Phi \models Qy\exists zC[x/y]\{t/z\}$  for  $Q \in \{\forall, \exists\}$ .*

**Lemma 9.** *Let  $t$  be a maximal colored term in  $C$  in  $\Phi$ . It is  $\Phi$ -colored. Let  $x_1, \dots, x_n$  be the variables which occur in  $t$ . Then  $\Phi \models Q\bar{x}\exists yC\{t/y\}$  for  $Q \in \{\forall, \exists\}$ .*

actually

We have that  $\Gamma \models \forall \bar{x} \ell_{\Delta, x} [\text{PI}(C) \vee C]$ .

Note that both are lifted.

We have that  $\Gamma \models \forall \bar{x} \ell_{\Delta, x} [C] \vee C$ .

Let  $t$  be a maximal  $\Gamma$ -term. It in general contains  $\Gamma$ -colored and grey terms, and also  $\Delta$ -terms. The latter have entered it by unification.

If  $t$  contains no  $\Delta$  terms, we can just overbind it existentially and give a witness.

Otherwise it contains  $\Delta$ -terms. Then there is a variable in  $t$  at position say  $p$  which also occurs elsewhere in  $C$ , say at position  $q$ .

If  $q$  is the outermost term or if it has only grey term ancestors, then quantifying over whatever is in  $q$  before quantifying over whatever is in  $p$  is fine. Hence there is an arrow.

$q$  can not be contained in a  $\Delta$  term since dependencies cannot be introduced and must be there from the beginning, where no color mix is possible.

So otherwise  $q$  is contained in a maximal  $\Gamma$ -term  $s$ . For finding witnesses, we will put the same one for the variable at both  $q$  and  $t$ . As  $q$  introduces a  $\Delta$ -term, at some point, there had had to be a unification with a formula from  $\Delta$  (this then could have been passed on through “mirroring”).

Conjecture: there are arrows along the path from the origin of the  $\Delta$ -path to  $q$ .

Hence whatever is placed in  $q$  and  $p$  is quantified over earlier than the variables which replace  $t$  and  $s$ .

TODO: Proof or refute...

conjecture: put all terms that share variables and appear in the same clause and are all overbound with the same quantifier in the same quantifier block.

probably does not work when facing other dependencies, check that !

Notation:

$p_1$  is the position of  $s$  in  $t$

$p_2$  is the elsewhere position of the shared var

A unification where a  $\Gamma$ -colored term  $s$  enters  $t$  happens when

- the other unified clause has a variable at the position of  $t$  ( $p_1$ )
- a variable is both in  $t$  ( $p_1$ ) and elsewhere in the unified clause ( $p_2$ )
- $p_2$  is either in a grey term or as outermost or a  $\Delta$ -colored term
  - if  $p_2$  directly in grey term or as outermost, the ancestor of  $p_2$  will not be overbound (only  $p_2$ ). we need  $p_2$  as witness for overbinding  $t$ , but not the other way.

Hence quantifying over  $p_2$  first is ok.

- if  $p_2$  is in a  $\Delta$ -colored term, say in maximal  $\Delta$ -term  $s'$ , Then  $s'$  and  $t$  are overbound with the same quantifier and order between them doesn't matter.

for witness, we both need whatever the var is, and that we get by the inherited relation.

? there must be an inherited relation as since both  $s'$  and  $t$  are  $\Gamma$  and contain a  $\Delta$ -term, the  $\Delta$  term must have gotten into a  $\Gamma$ -colored term using aufschaukeln ?

## 4 structured proof

**Lemma 10.**  $\Gamma \models PI(C) \vee C$  for  $C$  in a prop proof.

*Proof.* See Huang. □

**Lemma 11.**  $\Gamma \models \forall x_1 \dots \forall x_n \ell_{\Delta, x}[PI(C) \vee C]$  for  $C$  in a prop proof.

*Proof.* Still the same as in Huang. □

**Lemma 12.**  $AI_{\text{mat}}(C) = \ell[PI(C)]$  for  $C \in \pi$ .

SUPPOSE NO VARIABLE OCCURS TWICE IN A COLORED TERM IN AN INITIAL CLAUSE SET

**Lemma 13.** if there is a max  $\Delta$ -term in a max  $\Gamma$ -term, there is an arrow from occurrence of  $\Delta$ -term to the occurrence of  $\Gamma$ -term in AI.

*Proof.* induction:

base case: no foreign terms

$C_1 : D \vee l$  and  $C_2 : E \vee \neg l$

resolution, same color: induction hypothesis!!!

resolution, different color:  $l$  and  $l'$  unified. Disregard grey terms. Supp one of the unification locations, a term has a variable and all the prefix is grey/shared in  $l$  and  $l'$ . Then term from other literal enters, possibly foreign. Then by replacing all variables in  $C_1$  or  $C_2$ , a  $\Delta$ -term might enter a  $\Gamma$ -term. But in this case, we have an arrow.

$\Gamma$ -terms and  $\Delta$ -terms are different and hence not unifiable.

Suppose same prefix (i.e. same colored prefix), then different variables each. Supp variable at one end, foreign colored term at other. then arrows of literals are merged, and by induction hypothesis, the term with the foreign colored term has an arrow to the foreign colored term.

// for each unification where possibly foreign terms are introduced, there is an arrow.

Either resolution with same color: as long as just same color resolutions, no new literals in interpolant (but new terms by resolution?)

if after chain of same color resolution a grey literal is resolved, transitive edges kick in (chain of arrows)

resolution with grey literals: resolved literals share grey or aufgeschaukelter prefix before the variable.

RESTRICTION APPLIES HERE: no aufgeschaukelte prefixes

if the variable occurs elsewhere in one of the clauses, a foreign term might have been introduced in a colored term. in this case, there is an arrow □

**Lemma 14.**  $\Gamma \models \bar{Q}_n \ell_{\Gamma, y}[\ell_{\Delta, x}[PI(C) \vee C]]$  for  $C$  in a prop proof.

i.e.

$\Gamma \models AI(C)$  for  $C$  in a prop proof.

*Proof.* Show that the existential quantifiers in  $\bar{Q}_n$  have witnesses.

If the clause  $C$  is the result of a resolution step of  $C_1 : D \vee l$  and  $C_2 : E \vee \neg l$ , then by induction hypothesis, we get that

supp  $l$  grey:  $AI(C) = Q_1 u_i \dots Q_m u_m (\neg \chi \wedge AI_{\text{mat}}(C_2)) \vee (l \wedge AI_{\text{mat}}(C_1))$

Know:  $\Gamma \models \ell_{\Gamma, y}[PI(C) \vee C]$

By Lemma 12: Know:  $\ell[PI(C) \vee C] = AI_{\text{mat}}(C) \vee \ell[C]$

Need:  $\Gamma \models AI(C) \vee C$  //  $C$  in binding of AI

i.e.  $\Gamma \models \{Q_{n_1} \cup Q_{n_2} \cup \text{new ones}\}(\ell[(l \wedge \ell[\text{PI}(C_2)])] \vee (\neg l \wedge \ell[\text{PI}(C_1)]) \vee D \vee E)$

existentially overbound variables in  $\ell[\text{PI}(C_i)]$  will still work if relative order is maintained, which it is.  
for new terms in  $l$ , we have lemma 11, which provide witnesses.

huang-style: show how a foreign term got into a colored term, and this is how it must have an arrow.

TODO: show for restricted version: all variables occur once as maximal “colored” term.

□

**Lemma 15.** *Suppose no colored term occurs in  $\text{PI}(C) \vee C$  for  $C \in \pi$ . Then  $\Gamma \models \text{AI}_{\text{mat}}(C) \vee \text{AI}_{\text{cl}}(C)$  // implicit universal quantification*

*Proof.* Proof by induction.

Base case:

For  $C \in \Gamma$ ,  $\text{AI}_{\text{mat}}(C) = \perp$  and  $\text{AI}_{\text{cl}} = \ell[C] = \ell_{\Gamma,y}[C]$ . By the restriction,  $\ell_{\Gamma,y}[C] = C$  and  $\Gamma \models C$ .

For  $C \in \Delta$ ,  $\text{AI}_{\text{mat}}(C) = \top$ .

Induction step:

We know:  $\Gamma \models \text{AI}_{\text{mat}}(C_i) \vee \text{AI}_{\text{cl}}(C_i)$ ,  $i \in \{1, 2\}$ .

- Suppose  $l$  and  $l'$  of opposite color.

$$\text{AI}_{\text{mat}}(C)' = \left( (\neg l \wedge \text{AI}_{\text{mat}}(C_1)) \vee (l \wedge \text{AI}_{\text{mat}}(C_2)) \right) \sigma$$

$$\text{AI}_{\text{cl}}(C)' = \left( (\text{AI}_{\text{cl}}(C_1) \setminus \{l\}) \vee (\text{AI}_{\text{cl}}(C_2) \setminus \{\neg l'\}) \right) \sigma \quad // \text{ setminus: remove clause with that ancestor}$$

To show:  $\Gamma \models \ell[\text{AI}_{\text{mat}}(C)' \vee \text{AI}_{\text{cl}}(C)']$ , but as  $\sigma$  does not introduce a colored term, this is the same as  $\Gamma \models \text{AI}_{\text{mat}}(C)' \vee \text{AI}_{\text{cl}}(C)'$ .

Suppose  $\Gamma \not\models (\text{AI}_{\text{cl}}(C_1) \setminus \{l\})\sigma$  and  $\Gamma \not\models (\text{AI}_{\text{cl}}(C_2) \setminus \{\neg l'\})\sigma$  as otherwise we would be done.

Then  $\Gamma \not\models (\text{AI}_{\text{cl}}(C_1) \setminus \{l\})$  and  $\Gamma \not\models (\text{AI}_{\text{cl}}(C_2) \setminus \{\neg l'\})$ .

Hence the induction hypothesis reduces to  $\Gamma \models \text{AI}_{\text{mat}}(C_1) \vee l$  and  $\Gamma \models \text{AI}_{\text{mat}}(C_2) \vee \neg l'$ .

Therefore also  $\Gamma \models (\text{AI}_{\text{mat}}(C_1) \vee l)\sigma$  and  $\Gamma \models (\text{AI}_{\text{mat}}(C_2) \vee \neg l')\sigma$ .

Now as  $l\sigma = l'\sigma$ , their interpretation is linked, so we get the result in a similar way as in Huang's proof.

□

## 5 old stuff, not sure if valuable

**Lemma 16.** *Let  $\Phi$  be a set of formulas,  $t$  be a term and  $\sigma$  a substitution. Then  $\ell_{\Phi,x}[\ell_{\Phi,x}[t]\sigma] = \ell_{\Phi,x}[t\sigma]$ . //*  
**WRONG**

*Proof.* With  $\sigma'$  as in Lemma ??:

$$\ell_{\Phi,x}[t\sigma] = \ell_{\Phi,x}[t]\sigma'.$$

As  $\sigma$  just depends on the terms to replace and the variables to replace them with:

$$\ell_{\Phi,x}[\ell_{\Phi,x}[t]\sigma] = \ell_{\Phi,x}[\ell_{\Phi,x}[t]]\sigma' = \ell_{\Phi,x}[t]\sigma'. \quad \square$$

OLD from first principles reasoning (complete this in case Lemma ?? is flawed):

(this was based on  $\Delta$ -terms)

$t$  contains  $\Delta$ -terms  $t_i$ , grey terms  $g_i$  and free variables  $v_i$ .  $\ell_{\Delta,x}[t]$  is  $t$ , where every maximal  $\Delta$ -term  $t_i$  is replaced by free variable  $x_i$ .

A substitution  $\sigma$  occurring in a resolution refutation does not affect any of the  $x_i$ , as these symbols do not occur in the initial clause sets.

Hence  $\ell_{\Delta,x}[t]\sigma$  differs from  $\ell_{\Delta,x}[t]$  only in the  $v_i$ , which are potentially substituted. The  $v_i$  can be substituted to other free variables  $v'_i$  (not the same ones as factors of clauses are variable disjoint), grey terms  $g'_i$  or  $\Delta$ -terms  $t'_i$  (these terms may contain grey or  $\Delta$ -terms or free variables). Assumption:  $\ell_{\Delta,x}[\ ]$  always replaces a certain  $\Delta$ -term with the same variable, globally.

Goal:  $\ell_{\Delta,x}[\ell_{\Delta,x}[t]\sigma] = \ell_{\Delta,x}[t\sigma]$ .

Suppose  $\sigma$  does not affect  $t$  (and also not the  $x_i$  as it occurs in a resolution rule application). Then  $\ell_{\Delta,x}[t\sigma] = \ell_{\Delta,x}[t] = \ell_{\Delta,x}[\ell_{\Delta,x}[t]]$  (

Otherwise  $\sigma$  changes a variable  $v_i$  at position  $p$  in  $t$ .  $\ell_{\Delta,x}[t]$  has at position  $p$   $v_i$  as well if the path to  $p$  does not contain colored symbol. So  $\ell_{\Delta,x}[\ell_{\Delta,x}[t]\sigma]$  and  $\ell_{\Delta,x}[t\sigma]$  coincide at  $p$ , irrespective of what  $\sigma$  introduces (var,  $\Delta$ -/grey-term).

If  $p$  points into a maximal  $\Delta$ -colored term,  $p \bmod k$  for some  $k \geq 1$  is a  $\Delta$ -colored term  $t_j$ . Hence in  $\ell_{\Delta,x}[t]$ ,  $\rho(p \bmod k) = x_j$ .

However in  $\ell_{\Delta,x}[t\sigma]$ ,  $\rho(p \bmod k) = x_l$ .

**Lemma 17.** *If  $l \in C$ , then  $\ell[l] \in \text{AI}_{\text{cl}}(C)$ . //* **WRONG**

*Proof.* By induction:

Base case by definition.

Let a literal  $\lambda$  be such that it hasn't been resolved upon in the deduction leading up to  $C$ .

Resolution  $C_1 : D \vee l$  and  $C_2 : E \vee \neg l'$  with  $l\sigma = l'\sigma$  give  $C : (D \vee E)\sigma$ .

$\lambda \not\sim l$ , as otherwise it would not be contained in  $C$ . W.l.o.g.  $\lambda \in C_1$ . Then  $(\lambda\sigma) \in C$  by the resolution rule.

Induction hypothesis:  $\ell[\lambda] \in \text{AI}_{\text{cl}}(C_1)$ .

So have to show  $\ell[\lambda\sigma] \in \text{AI}_{\text{cl}}(C)$  // as  $(\lambda\sigma) \in C$

By Lemma 16,  $\ell[\lambda\sigma] = \ell[\ell[\lambda]\sigma]$ . We show that in all resolution cases,  $\ell[\ell[\lambda]\sigma] \in \text{AI}_{\text{cl}}(C)$ .

$$\text{AI}_{\text{cl}}(C) = \ell\left[\left((\text{AI}_{\text{cl}}(C_1) \setminus \{\ell[l]\}) \vee (\text{AI}_{\text{cl}}(C_2) \setminus \{\neg\ell[l']\})\right)\sigma\right]$$

As  $\lambda \not\sim l$  and  $\ell[\lambda] \in \text{AI}_{\text{cl}}(C_1)$ ,  $\ell[\ell[\lambda]\sigma] \in \text{AI}_{\text{cl}}(C)$ .  $\square$



**Proposition 18.**  $\Gamma \models Q_1 z_1 \dots Q_n z_n \overline{\text{PI}(C) \vee C}(z_1, \dots, z_n)$  , quantifiers ordered as in 5, is a craig interpolant.

*Proof.* Induction.

Suppose Resolution.

$$\frac{C_1 : D \vee l \quad C_2 : E \vee \neg l'}{C : (D \vee E)\sigma} \quad \sigma = \text{mgu}(l, l')$$

$$\Gamma \models \bar{Q}_{n_1} \overline{\text{PI}(C_1) \vee D \vee l}$$

$$\Gamma \models \bar{Q}_{n_2} \overline{\text{PI}(C_2) \vee E \vee \neg l'}$$

to show:

$$\Gamma \models \bar{Q}_n \overline{\text{PI}(C) \vee (D \vee E)\sigma} \quad // \text{ somewhat imprecise on } \bar{Q}_n, \text{ but that's just useless quantifiers}$$

$$\Gamma \models (\bar{Q}_{n_1} \overline{\text{PI}(C_1) \vee D \vee l})\sigma$$

$$\Gamma \models (\bar{Q}_{n_2} \overline{\text{PI}(C_2) \vee E \vee \neg l'})\sigma$$

By resolution:

$$\Gamma \models (\bar{Q}_{n_1} \overline{\text{PI}(C_1)} \vee \bar{Q}_{n_2} \overline{\text{PI}(C_2)})\sigma \vee (D \vee E)\sigma$$

1. Suppose  $l, l'$  are from  $\Gamma$  alone: TODO
2. Suppose  $l$  and  $l'$  are colored with different colors and w.l.o.g  $l$  is  $\Gamma$ -colored and  $l'$  is  $\Delta$ -colored.

$$\begin{aligned} \bar{Q}_n \overline{\text{PI}(C)} &\equiv \bar{Q}_n [(\neg l' \wedge \text{PI}(C_1)^*) \vee (l \wedge \text{PI}(C_2)^*)]\sigma \\ &\equiv \bar{Q}_n (\overline{\neg l' \sigma} \wedge \overline{\text{PI}(C_1)\sigma}) \vee (\overline{l \sigma} \wedge \overline{\text{PI}(C_2)\sigma}) \end{aligned}$$

Adapt Huang proof to this, need to consider quantifiers:

If  $\Gamma \not\models D\sigma$  and  $\Gamma \not\models E\sigma$  (else we are done), then

$$\Gamma \models [(\neg l' \wedge \bar{Q}_{n_1} \overline{\text{PI}(C_1)}) \vee (l \wedge \bar{Q}_{n_2} \overline{\text{PI}(C_2)})]\sigma$$

As  $\bar{Q}_{n_1}$  and  $\bar{Q}_{n_2}$  disjoint and their variables do not appear in  $l$  or  $l'$ ,

$$\Gamma \models (\bar{Q}_{n_1} \bar{Q}_{n_2} [(\neg l' \wedge \overline{\text{PI}(C_1)}) \vee (l \wedge \overline{\text{PI}(C_2)})])\sigma$$

$$\Gamma \models \bar{Q}_{n_1} \bar{Q}_{n_2} [(\neg l' \sigma \wedge \overline{\text{PI}(C_1)\sigma}) \vee (l \sigma \wedge \overline{\text{PI}(C_2)\sigma})]$$

Consider the maximal terms of this expression which are  $\Delta$ -colored.

The  $\text{PI}(C_i)$ ,  $i \in \{1, 2\}$  contain no colored terms.  $\sigma$  can introduce one by replacing a free variable  $x$  by a  $\Delta$ -term  $t$ . But then overline replaces it with an universally quantified variable again, hence the formula is still entailed by  $\Gamma$ .

$$\Gamma \models \bar{Q}_{n_1} \bar{Q}_{n_2} [(\neg l' \sigma \wedge \overline{\text{PI}(C_1)\sigma}) \vee (l \sigma \wedge \overline{\text{PI}(C_2)\sigma})]$$

TODO: should work out similarly as huang if using  $P_P$  or it's the same as what i'm trying above.

□

**Proposition 19.**  $\Gamma \models Q_1 z_1 \dots Q_n z_n \text{PI}(C)^*(z_1, \dots, z_n) \vee C$ , quantifiers ordered as in 5, is a craig interpolant.

*Proof.* Induction.

Suppose Resolution.

$$\frac{C_1 : D \vee l \quad C_2 : E \vee \neg l'}{C : (D \vee E)\sigma} \quad \sigma = \text{mgu}(l, l')$$

$$\Gamma \models \bar{Q}_{n_1} \text{PI}(C_1)^* \vee D \vee l$$

$$\Gamma \models \bar{Q}_{n_2} \text{PI}(C_2)^* \vee E \vee \neg l'$$

to show:  $\Gamma \models \bar{Q}_n \text{PI}(C)^* \vee (D \vee E)\sigma$

$$\Gamma \models (\bar{Q}_{n_1} \text{PI}(C_1)^* \vee D \vee l)\sigma$$

$$\Gamma \models (\bar{Q}_{n_2} \text{PI}(C_2)^* \vee E \vee \neg l')\sigma$$

By resolution:

$$\Gamma \models (\bar{Q}_{n_1} \text{PI}(C_1)^* \vee \bar{Q}_{n_2} \text{PI}(C_2)^*)\sigma \vee (D \vee E)\sigma$$

1. Suppose  $l, l'$  are from  $\Gamma$  alone: TODO
2. Suppose  $l$  and  $l'$  are colored with different colors and w.l.o.g  $l$  is  $\Gamma$ -colored and  $l'$  is  $\Delta$ -colored.

$$\bar{Q}_n \text{PI}(C)^* \equiv \bar{Q}_n([(\neg l' \wedge \text{PI}(C_1)^*) \vee (l \wedge \text{PI}(C_2)^*)])\sigma^*$$

Adapt Huang proof to this, need to consider quantifiers:

If  $\Gamma \not\models D\sigma$  and  $\Gamma \not\models E\sigma$  (else we are done), then

$$\Gamma \models [(\neg l' \wedge \bar{Q}_{n_1} \text{PI}(C_1)^*) \vee (l \wedge \bar{Q}_{n_2} \text{PI}(C_2)^*)]\sigma$$

As  $\bar{Q}_{n_1}$  and  $\bar{Q}_{n_2}$  disjoint and their variables do not appear in  $l$  or  $l'$ ,

$$\Gamma \models (\bar{Q}_{n_1} \bar{Q}_{n_2} [(\neg l' \wedge \text{PI}(C_1)^*) \vee (l \wedge \text{PI}(C_2)^*)])\sigma$$

The  $\text{PI}(C_i)$ ,  $i \in \{1, 2\}$  contain no colored terms.  $\sigma$  can introduce one by replacing a free variable  $x$ .

Consider the maximal terms of this expression which are  $\Gamma$ -colored.

Either they only have grey subterms, then if they are existentially quantified, we can just use it as witness as the terms aren't replaced.

Otherwise they contain at least a  $\Gamma$ - or a  $\Delta$ -colored subterm.

Base case: simple.

Suppose Resolution.

$$\frac{C_1 : D \vee l \quad C_2 : E \vee \neg l'}{C : (D \vee E)\sigma} \quad \sigma = \text{mgu}(l, l')$$

$$\Gamma \models \bar{Q}_{n_1} \text{PI}(C_1)^* \vee D \vee l$$

$$\Gamma \models \bar{Q}_{n_2} \text{PI}(C_2)^* \vee E \vee \neg l'$$

to show:  $\Gamma \models \bar{Q}_n \text{PI}(C)^*\sigma \vee (D \vee E)\sigma$

Note that a term newly introduced in  $\text{PI}(C)$  occurs in either  $l$  or  $l'$ , but not in both.

Let  $t$  be a colored term in  $\text{PI}(C)$ , which has just been added W.l.o.g. let it occur in  $l$ , i.e. in  $C_1$ .

Case distinction:

1. Suppose  $l, l'$  are from  $\Gamma$  alone:

By induction hypothesis:

$$\Gamma \models (\bar{Q}_{n_1} \text{PI}(C_1)^* \vee D \vee l)\sigma$$

$$\Gamma \models (\bar{Q}_{n_2} \text{PI}(C_2)^* \vee E \vee \neg l')\sigma$$

By resolution:

$$\Gamma \models (\bar{Q}_{n_1} \text{PI}(C_1)^* \vee \bar{Q}_{n_2} \text{PI}(C_2)^*)\sigma \vee (D \vee E)\sigma$$

Suppose  $t$  is  $\Gamma$ -colored.

Then it will be replaced by  $x_i$  and existentially quantified. It appears in either  $\text{PI}(C_1)$  or  $\text{PI}(C_2)$ .

$t$  is a witness for  $x_i$  because it contains subterms  $t_1, \dots, t_n$ . If they are overbound as well, they are so before  $t$  and are available here.

TODO: derive properties using examples 103 or so

OTHER TRY:

Then  $\sigma$  replaces variables  $y_1, \dots, y_k$  in  $E \vee \neg l'$  with terms that contain  $t$ .

By the induction hypothesis,  $\Gamma \models Q_1 z_1 \dots Q_{n_2} z_{n_2} \text{PI}(C_2)^*(z_1, \dots, z_{n_2}) \vee E \vee \neg l'$ .

Hence  $\Gamma \models (Q_1 z_1 \dots Q_{n_2} z_{n_2} \text{PI}(C_2)^*(z_1, \dots, z_{n_2}) \vee E \vee \neg l')\sigma$ .

Also  $\Gamma \models Q_1 z_1 \dots Q_{n_2} z_{n_2} (\text{PI}(C_2)^*(z_1, \dots, z_{n_2})\sigma) \vee E\sigma \vee \neg l'\sigma$ .

Similarly,  $\Gamma \models Q_1 z_1 \dots Q_{n_1} z_{n_1} (\text{PI}(C_1)^*(z_1, \dots, z_{n_1})\sigma) \vee D\sigma \vee l\sigma$

$\Gamma \models Q_1 z_1 \dots Q_n z_n ((\neg l \wedge \text{PI}(C_2)) \vee (l \wedge \text{PI}(C_1)))^*(z_1, \dots, z_n)\sigma \vee D\sigma \vee l\sigma$

$l$  basically is the only new thing ( $l\sigma = l'\sigma$ ).

Either  $l$  does not contain any subterms of other terms, then it does not depend on anything and  $l$  serves as witness for itself.

Otherwise it does depend on other terms and we have to make sure that that term is available. Depending on another term means that it uses information that is only available from another term, i.e. it contains a subterm of another term. but then that subterm is quantified over before the variable that replaces  $t$  is, so it works out.

$t$  is  $\Delta$ -colored. Then it is replaced by a universally quantified variable. But it “was already universally quantified” in the induction hypothesis. There, it was some free variable, because that’s the only thing that can be substituted, but even with this free var, it worked out.

□

**Conjecture 20.**  $\Gamma \cup \Delta$  unsat,  $\pi$  propositional resolution refutation. Then  $\Gamma \models \bar{Q}_n \text{PI}(C)^* \vee C$  and  $\Delta \models \neg \bar{Q}_n \text{PI}(C)^* \vee C$  for all  $C$  in  $\pi$ .

*Proof.* Base case as in Huang.

Induction.

Suppose Resolution.

$$\frac{C_1 : D \vee l \quad C_2 : E \vee \neg l}{C : D \vee E}$$

$\Gamma \models \bar{Q}_{n_1} \text{PI}(C_1)^* \vee D \vee l$

$\Gamma \models \bar{Q}_{n_2} \text{PI}(C_2)^* \vee E \vee \neg l$

to show:  $\Gamma \models \bar{Q}_n \text{PI}(C)^* \vee D \vee E$ , i.e.

$\Gamma \models \text{sort}(Q_{n_1} \cup Q_{n_2} \cup \text{colored-terms}(l))((\neg l^* \wedge \text{PI}(C_1)^*) \vee (l^* \wedge \text{PI}(C_2)^*)) \vee D \vee E$

If  $\Gamma \not\models D$  and  $\Gamma \not\models E$  (else we are done), then

$\Gamma \models (\neg l \wedge \bar{Q}_{n_1} \text{PI}(C_1)^*) \vee (l \wedge \bar{Q}_{n_2} \text{PI}(C_2)^*)$

As  $\bar{Q}_{n_1}$  and  $\bar{Q}_{n_2}$  disjoint and their variables do not appear in  $l$  or  $l$ ,

$\Gamma \models \bar{Q}_{n_1} \bar{Q}_{n_2} [(\neg l \wedge \text{PI}(C_1)^*) \vee (l \wedge \text{PI}(C_2)^*)]$

Since we’ve pushed the variables outside, no colored terms appear in  $\text{PI}(C_i)^*$ .

Suppose  $l$  does not contain colored terms. Then  $l = l^*$  and we are done.

Otherwise let  $t$  be a maximal colored term in  $l$ .

By lemma 7,  $l$  appears in  $\Gamma$  with a certain polarity, say in clause  $E$ .  $l$  is an instance of  $E$ .

In fact,  $l$  is contained in  $C\sigma$  where  $\sigma$  is the composition of unifiers applied in the derivation up to the current point.

Hence  $\Gamma \models C\sigma$ .

1. Suppose  $t$  is  $\Gamma$ -colored.  $\Gamma \models l$  implies that  $\Gamma \models \exists y l\{t/y\}$
2. Suppose  $t$  is  $\Delta$ -colored.  $\Gamma \models \forall y l\{t/y\}$  because:

□