# Contents

## 0.1   referenced lemmas from previous sections

⟨lem:lemma:logiccommute⟩?

**Lemma 1** (Commutativity of lifting and logical operators)**.** *Let $A$ and $B$ be first-order formulas and $s$ and $t$ be terms. Then it holds that:*

    *1. $\ell^z_\Phi[\neg A] \Leftrightarrow \neg\ell^z_\Phi[A]$*

    *2. $\ell^z_\Phi[A \circ B] \Leftrightarrow (\ell^z_\Phi[A] \circ \ell^z_\Phi[B])$ for $\circ \in \{\wedge, \vee\}$*

    *3. $\ell^z_\Phi[s = t] \Leftrightarrow (\ell^z_\Phi[s] = \ell^z_\Phi[t])$*

?⟨thm@@lemmaCommutLiftSubst@data⟩?

**Lemma 2** (Commutativity of lifting and substitution)**.** *Let $C$ be a clause and $\sigma$ a substitution such that no lifting variable occurs in $C$ or $\sigma$. Define $\sigma'$ with $\mathrm{dom}(\sigma') = \mathrm{dom}(\sigma) \cup \{z_t \mid t\sigma \neq t\}$ such that for a variable $z$,*

$$x\sigma' = \begin{cases} z_{t\sigma} & \text{if } x = z_t \text{ and } t\sigma \neq t \\ \ell^z_\Phi[x\sigma] & \text{otherwise} \end{cases}$$

*Then $\ell^z_\Phi[C\sigma] = \ell^z_\Phi[C]\sigma'$.*

1

# Interpolant extraction from resolution proofs in one phase

While the previous chapter demonstrates that it is possible to extract propositional interpolants and lift them from the colored symbols later in order to obtain a proper interpolant, we now present a novel approach, which only operates with grey intermediary interpolants. This is established by lifting any term which is added to the interpolant.

By its nature, this approach requires an alternate strategy than the proof of the extraction in two phases as a commutation of substitution and lifting is no longer possible if lifting variables are present. Let us recall the corresponding lemma from the previous chapter:

**Lemma 2** (Commutativity of lifting and substitution)**.** *Let $C$ be a clause and $\sigma$ a substitution such that no lifting variable occurs in $C$ or $\sigma$. Define $\sigma'$ with $\mathrm{dom}(\sigma') = \mathrm{dom}(\sigma) \cup \{z_t \mid t\sigma \neq t\}$ such that for a variable $z$,*

$$x\sigma' = \begin{cases} z_{t\sigma} & \text{if } x = z_t \text{ and } t\sigma \neq t \\ \ell^z_\Phi[x\sigma] & \text{otherwise} \end{cases}$$

*Then $\ell^z_\Phi[C\sigma] = \ell^z_\Phi[C]\sigma'$.*

Consider the following illustration of a problem of the notion of applying this lemma to terms containing lifting variables:

**Example 3.** Let $\sigma = \{x \mapsto a\}$ and consider the terms $f(x)$ and $f(a)$, where $f$ and $a$ are colored symbols. Clearly $f(x)\sigma = f(a)$ and therefore necessarily $z_{f(x)}\sigma' = z_{f(a)}$.

But now consider $x_{f(x)}\sigma$. As $z_{f(x)}$ is a lifting variable, it is not affected by unifiers from resolution derivations and also not by $\sigma$. Hence $z_{f(x)}\sigma = z_{f(x)}$ and therefore $\ell[z_{f(x)}\sigma] = \ell[z_{f(x)}] = z_{f(x)}$, but $\ell[z_{f(x)}]\sigma' = z_{f(x)}\sigma' = z_{f(a)}$. So $\ell[z_{f(x)}\sigma] \neq \ell[z_{f(x)}]\sigma'$.

We see here that there are circumstances under which in order to commute lifting and substitution, the substitution $\sigma'$ is required to conform to the equation $z_{f(x)}\sigma' = z_{f(a)}$, whereas in others, it must hold that $z_{f(x)}\sigma' = z_{f(x)}$. △

## 1.1 Definition of the extraction algorithm

The extracted interpolants are prenex formulas, where the quantifier block and the matrix of the formula are calculated separately in each step of the traversal of the resolution refutation.

### 1.1.1 Extraction of the interpolant formula matrix $\text{AI}_{\text{mat}}$ and calculation of $\text{AI}_{\text{cl}}$

$\text{AI}_{\text{mat}}$ is inspired by the propositional interpolants PI from Definition **??**. Its difference lies in the fact that the lifting occurs in every step of the extraction. This however necessitates applying these liftings to the clauses of the resolution refutation as well. For a clause $C$ of the resolution refutation, we will denote the clause with the respective liftings applied by $\text{AI}_{\text{cl}}(C)$ (a formal definition will be given below), and for a term $t$ at position $p$ in $C$, we denote $\text{AI}_{\text{cl}}(C)|_p$ by $t_{\text{AIcl}}$.

Now we can define preliminary versions of $\text{AI}_{\text{mat}}^\bullet$ and $\text{AI}_{\text{cl}}^\bullet$:

**Definition 4** ($\text{AI}_{\text{mat}}^\bullet$ and $\text{AI}_{\text{cl}}^\bullet$)**.** Let $\pi$ be a resolution refutation of $\Gamma \cup \Delta$.
For a clause $C$ in $\pi$, $\text{AI}_{\text{mat}}^\bullet(C)$ and $\text{AI}_{\text{cl}}^\bullet(C)$ are defined as follows:

Base case. If $C \in \Gamma$, $\text{AI}_{\text{mat}}^\bullet(C) \overset{\text{def}}{=} \bot$. If otherwise $C \in \Delta$, $\text{AI}_{\text{mat}}^\bullet(C) \overset{\text{def}}{=} \top$.

In any case, $\text{AI}_{\text{cl}}^\bullet(C) \overset{\text{def}}{=} \ell[C]$.

Resolution. If the clause $C$ is the result of a resolution step of $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ using a unifier $\sigma$ such that $l\sigma = l'\sigma$, then $\text{AI}_{\text{mat}}^\bullet(C)$ and $\text{AI}_{\text{cl}}^\bullet$ are defined as follows:

$$\text{AI}_{\text{cl}}^\bullet(C) \overset{\text{def}}{=} \ell[(\text{AI}_{\text{cl}}^\bullet(C_1)\backslash\{l_{\text{AIcl}}\})\sigma] \ \vee \ \ell[(\text{AI}_{\text{cl}}^\bullet(C_2)\backslash\{l'_{\text{AIcl}}\})\sigma]$$

1. If $l$ is $\Gamma$-colored: $\text{AI}_{\text{mat}}^\bullet(C) \overset{\text{def}}{=} \ell[\text{AI}_{\text{mat}}^\bullet(C_1)\sigma] \vee \ell[\text{AI}_{\text{mat}}^\bullet(C_2)\sigma]$

2. If $l$ is $\Delta$-colored: $\text{AI}_{\text{mat}}^\bullet(C) \overset{\text{def}}{=} \ell[\text{AI}_{\text{mat}}^\bullet(C_1)\sigma] \wedge \ell[\text{AI}_{\text{mat}}^\bullet(C_2)\sigma]$

3. If $l$ is grey: $\text{AI}_{\text{mat}}^\bullet(C) \overset{\text{def}}{=} (\neg\ell[l'_{\text{AIcl}}\sigma] \wedge \ell[\text{AI}_{\text{mat}}^\bullet(C_1)\sigma]) \ \vee \ (\ell[l_{\text{AIcl}}\sigma] \wedge \ell[\text{AI}_{\text{mat}}^\bullet(C_2)\sigma])$

Factorisation. If the clause $C$ is the result of a factorisation of $C_1 : l \vee l' \vee D$ using a unifier $\sigma$ such that $l\sigma = l'\sigma$, then $\text{AI}_{\text{mat}}^\bullet(C) \overset{\text{def}}{=} \ell[\text{AI}_{\text{mat}}^\bullet(C_1)\sigma]$ and $\text{AI}_{\text{cl}}^\bullet(C) \overset{\text{def}}{=} \ell[(\text{AI}_{\text{cl}}(C_1)\backslash\{l'_{\text{AIcl}}\})\sigma]$. $\triangle$

Note that in $\text{AI}_{\text{mat}}^\bullet$ and $\text{AI}_{\text{cl}}^\bullet$, it is possible that there for a colored term $t$ in $C$ that $t_{\text{AIcl}} \neq z_t$ as illustrated by the following examples:

**Example 5.** We consider a resolution refutation of the initial clause sets $\Gamma = \{R(c), \neg Q(v)\}$ and $\Delta = \{\neg R(u) \vee Q(g(u))\}$:

$$\frac{\dfrac{R(c) \qquad \neg R(u) \vee Q(g(u))}{Q(g(c))} \ \text{res}, y \mapsto c \qquad \neg Q(v)}{\Box} \ \text{res}, v \mapsto g(c)$$

We now replace every clause $C$ by $\text{AI}_{\text{mat}}^\bullet(C) \mid \text{AI}_{\text{cl}}^\bullet(C)$ in order to visualize the steps of the algorithm:

$$\frac{\dfrac{\bot \mid R(y_c) \qquad \top \mid \neg R(u) \vee \neg Q(x_{g(u)})}{R(y_c) \mid Q(x_{g(u)})} \text{ res}, y \mapsto c \qquad \bot \mid \neg Q(v)}{\neg Q(x_{g(c)}) \wedge R(y_c) \mid \square} \text{ res}, v \mapsto g(c)$$

By quantifying $y_c$ existentially and $x_{g(c)}$ universally[1], we obtain an interpolant for $\Gamma \cup \Delta$: $\exists y_c \forall x_{g(c)} (\neg Q(x_{g_c}) \wedge R(y_c))$. Note however that $\ell[Q(g(c))] = Q(x_{g(c)})$, but $\mathrm{AI}_{\mathrm{mat}}(Q(g(c))) = Q(x_{g(u)})$. This example shows that this circumstance is not necessarily an obstacle for the correctness of this algorithm. $\quad \triangle$

⟨exa:2b⟩ **Example 6.** We consider a resolution refutation of the initial clause sets $\Gamma = \{R(c), P(c)\}$ and $\Delta = \{\neg R(u) \vee \neg Q(g(u)), \neg P(v) \vee Q(g(v))\}$:

$$\frac{\dfrac{\neg R(u) \vee \neg Q(g(u)) \qquad R(c)}{\neg Q(g(c))} \text{ res}, u \mapsto c \qquad \dfrac{\neg P(v) \vee Q(g(v)) \qquad P(c)}{Q(g(c))} \text{ res}, v \mapsto c}{\square} \text{ res}$$

We now again display $\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C) \mid \mathrm{AI}^{\bullet}_{\mathrm{cl}}(C)$ for every clause $C$ of the refutation:

$$\frac{\dfrac{\top \mid \neg R(u) \vee \neg Q(x_{g(u)}) \qquad \bot \mid R(y_c)}{R(y_c) \mid \neg Q(x_{g(u)})} \text{ res}, u \mapsto c \qquad \dfrac{\top \mid \neg P(v) \vee Q(x_{g(v)}) \qquad \bot \mid P(y_c)}{P(y_c) \mid Q(x_{g(v)})} \text{ res}, v \mapsto c}{(Q(x_{g(v)}) \wedge R(y_c)) \vee (\neg Q(x_{g(u)}) \wedge P(y_c)) \mid \square} \text{ res}$$

Note again that here, we have that $\ell[\neg Q(g(c))] = \neg Q(x_{g(c)}) \neq \mathrm{AI}^{\bullet}_{\mathrm{cl}}(\neg Q(g(c))) = \neg Q(x_{g(u)})$ and $\ell[Q(g(c))] = Q(x_{g(c)}) \neq \mathrm{AI}^{\bullet}_{\mathrm{cl}}(Q(g(c))) = Q(x_{g(v)})$. However in this instance, it is not possible to find quantifiers for the free variables of $\mathrm{AI}^{\bullet}_{\mathrm{mat}}(\square)$ such that by binding them, an interpolant is produced. For the naive approach, namely to use $\exists y_c \forall x_{g(v)} \forall x_{g(u)}$ as prefix, it holds that $\Gamma \models \exists y_c \forall x_{g(v)} \forall x_{g(u)} ((Q(x_{g(v)}) \wedge R(y_c)) \vee (\neg Q(x_{g(u)}) \wedge P(y_c)))$. This failure is possible as intuitively, resolution deductions are valid by virtue of the resolved literals being equal. The interpolant extraction procedure exploits this property not directly on the clauses but on the lifted clause, i.e. on $\mathrm{AI}_{\mathrm{cl}}(C)$ for a clause $C$. Note that by ensuring that for resolved literals $l$ and $l'$, it holds that $l_{\mathrm{AIcl}} = l'_{\mathrm{AIcl}}$, we can obtain an interpolant, for instance: $\exists y_c \forall x^* ((Q(x^*) \wedge R(y_c)) \vee (\neg Q(x^*) \wedge P(y_c)))$. $\quad \triangle$

In order to avoid the pitfall shown in Example 6 and to generalize the indicated solution, we define a function on resolved literals calculating a substitution, which ensures that the literals in the lifted clause, which correspond to the resolved literals, are equal.

**Definition 7** (au). Let $\iota$ be a resolution or factorisation rule application with $l$ and $l'$ as resolved or factorised literals, $\sigma = \mathrm{mgu}(\iota)$

For terms $s$ and $t$ where $s = \ell[l_{\mathrm{AIcl}}\sigma]|_p$ and $t = \ell[l'_{\mathrm{AIcl}}\sigma]|_p$ for some position $p$, we define:

$$\mathrm{au}'(s,t) \overset{\text{def}}{=} \begin{cases} \bigcup_{i=1}^n \mathrm{au}'(s_i, t_i) & \text{if } s \text{ is grey}, s = f_s(s_1, \ldots, s_n) \text{ and} \\ & t = f_t(t_1, \ldots, t_n)^2 \\ \{z_{s'} \mapsto z_r, z_{t'} \mapsto z_r\} & \text{if } s \text{ is a lifting variable } z_{s'}, t = z_{t'}, \text{ and} \\ & z_r = \ell[l\sigma]|_p \end{cases}$$

---

[1] The procedure for calculating the quantifier block is defined in section 1.4

For $\ell[l_{\mathrm{AIcl}}\sigma] = P(s_1, \ldots, s_n)$ and $\ell[l'_{\mathrm{AIcl}}\sigma] = P(t_1, \ldots, t_n)$, we define:

$$\mathrm{au}'(\ell[l_{\mathrm{AIcl}}\sigma], \ell[l'_{\mathrm{AIcl}}\sigma]) = \mathrm{au}'(P(\bar{s}), P(\bar{t})) \stackrel{\mathrm{def}}{=} \bigcup_{i=1}^{n} \mathrm{au}'(s_i, t_i)$$

$$\mathrm{au}(\iota) \stackrel{\mathrm{def}}{=} \mathrm{au}'(\ell[l_{\mathrm{AIcl}}\sigma], \ell[l'_{\mathrm{AIcl}}\sigma]) \hspace{3cm} \triangle$$

⟨prop:tau_dom_ran⟩ **Proposition 8.** *Let $\iota$ be a resolution or factorisation rule application with $l$ and $l'$ as resolved or factorised literals, $\sigma = \mathrm{mgu}(\iota)$ Then $\mathrm{dom}(\mathrm{au}(\iota))$ consists exactly of the lifting variables of $\ell[l_{\mathrm{AIcl}}\sigma]$ and $\ell[l'_{\mathrm{AIcl}}\sigma]$ and $\mathrm{ran}(\mathrm{au}(\iota))$ consists exactly of the lifting variables of $\ell[l\sigma]$.*

> possibly argue here why au is well-defined (but it follows more or less directly from a later lemma)

**Definition 9** (AI$_{\mathrm{mat}}$ and AI$_{\mathrm{cl}}$)**.** Let $\pi$ be a resolution refutation of $\Gamma \cup \Delta$. AI$_{\mathrm{mat}}(\pi)$ is defined to be AI$_{\mathrm{mat}}(\square)$, where $\square$ is the empty clause derived in $\pi$.
For a clause $C$ in $\pi$, AI$_{\mathrm{mat}}(C)$ and AI$_{\mathrm{cl}}(C)$ are defined inductively as follows:

Base case. If $C \in \Gamma$, AI$_{\mathrm{mat}}(C) \stackrel{\mathrm{def}}{=} \bot$. If otherwise $C \in \Delta$, AI$_{\mathrm{mat}}(C) \stackrel{\mathrm{def}}{=} \top$.

In any case, AI$_{\mathrm{cl}}(C) \stackrel{\mathrm{def}}{=} \ell[C]$.

Resolution. If the clause $C$ is the result of a resolution step $\iota$ of $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ using a unifier $\sigma$ such that $l\sigma = l'\sigma$, then let $\tau = \mathrm{au}(\iota)$ and define AI$_{\mathrm{mat}}(C)$ and AI$_{\mathrm{cl}}(C)$ as follows:

AI$_{\mathrm{cl}}(C) \stackrel{\mathrm{def}}{=} \ell[(\mathrm{AI}_{\mathrm{cl}}(C_1) \setminus \{l_{\mathrm{AIcl}}\})\sigma]\tau \;\vee\; \ell[(\mathrm{AI}_{\mathrm{cl}}(C_2) \setminus \{l'_{\mathrm{AIcl}}\})\sigma]\tau$

1. If $l$ is $\Gamma$-colored: AI$_{\mathrm{mat}}(C) \stackrel{\mathrm{def}}{=} \ell[\mathrm{AI}_{\mathrm{mat}}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{mat}}(C_2)\sigma]\tau$

2. If $l$ is $\Delta$-colored: AI$_{\mathrm{mat}}(C) \stackrel{\mathrm{def}}{=} \ell[\mathrm{AI}_{\mathrm{mat}}(C_1)\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}(C_2)\sigma]\tau$

3. If $l$ is grey: AI$_{\mathrm{mat}}(C) \stackrel{\mathrm{def}}{=} (\neg\ell[l'_{\mathrm{AIcl}}\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}(C_1)\sigma]\tau) \;\vee\; (\ell[l_{\mathrm{AIcl}}\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}(C_2)\sigma]\tau)$

Factorisation. If the clause $C$ is the result of a factorisation $\iota$ of $C_1 : l \vee l' \vee D$ using a unifier $\sigma$ such that $l\sigma = l'\sigma$, then let $\tau = \mathrm{au}(\iota)$ and define AI$_{\mathrm{mat}}(C)$ and AI$_{\mathrm{cl}}(C)$ as follows:

AI$_{\mathrm{mat}}(C) \stackrel{\mathrm{def}}{=} \ell[\mathrm{AI}_{\mathrm{mat}}(C_1)\sigma]\tau$

AI$_{\mathrm{cl}}(C) \stackrel{\mathrm{def}}{=} \ell[(\mathrm{AI}_{\mathrm{cl}}(C_1) \setminus \{l'_{\mathrm{AIcl}}\})\sigma]\tau \hspace{3cm} \triangle$

## 1.2 Lifting the $\Delta$-terms

**Definition 10.** AI$_{\mathrm{mat}}^{\Delta}(C)$ (AI$_{\mathrm{cl}}^{\Delta}(C)$) for a clause $C$ is defined as AI$_{\mathrm{mat}}(C)$ (AI$_{\mathrm{cl}}(C)$) with the difference that in its inductive definition, every lifting $\ell[\varphi]$ for a formula or term $\varphi$ is replaced by a lifting of only the $\Delta$-terms $\ell_{\Delta}[\varphi]$.   $\triangle$

---

[2]Note that constants are treated as function symbols of arity zero.

⟨`lemma:no_colored_terms`⟩ **Lemma 11.** *Let $C$ be a clause of a resolution refutation $\pi$ of $\Gamma \cup \Delta$. $\mathrm{AI}_{\mathrm{mat}}(C)$ and $\mathrm{AI}_{\mathrm{cl}}(C)$ do not contain colored symbols. $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C)$ and $\mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$ do not contain $\Delta$-colored symbols.*

*Proof.* For $\mathrm{AI}_{\mathrm{mat}}(C)$ and $\mathrm{AI}_{\mathrm{cl}}(C)$, consider the following: In the base case of the inductive definitions of $\mathrm{AI}_{\mathrm{mat}}(C)$ and $\mathrm{AI}_{\mathrm{cl}}(C)$, no colored symbols occur. In the inductive steps, any colored symbol which is added by $\sigma$ to intermediary formulas is lifted. By Proposition 8, $\mathrm{ran}(\mathrm{au}(\iota))$ for inferences $\iota$ in $\pi$ only consists of lifting variables.

For $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C)$ and $\mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$, a similar argument goes through by reading colored as $\Delta$-colored. $\square$

⟨`lemma:substitute_and_lift`⟩ **Lemma 12.** *Let $\sigma$ be a substitution and $F$ a formula without $\Phi$-colored terms such that for a set of formulas $\Psi$, $\Psi \models F$. Then $\Psi \models \ell_{\Phi}^{z}[F\sigma]$.*

*Proof.* $\ell_{\Phi}^{z}[F\sigma]$ is an instance of $F$: $\sigma$ substitutes variables either for terms not containing $\Phi$-colored symbols or by terms containing $\Phi$-colored symbols. For the first kind, the lifting has no effect. For the latter, the lifting only replaces subterms of the terms introduced by the substitution by a lifting variable such that the original structure of $F$ remains invariant as it by assumption does not contain colored terms. $\square$

**Lemma 13.** *Let $l$ and $l'$ be resolved or factorised literals in a resolution derivation step $\iota$ creating a clause $C$ and $\tau = \mathrm{au}(\iota)$. For any substitution $(z_s \mapsto z_t) \in \tau$,*

<span style="color:red">*TODO: check which statement we actually need (resolved literal, clause?)*</span>

*make sure that it works for positions in the resolved literals as well as in the clause*

**Lemma 14.** <span style="color:red">***either reduce to "equal up to index of lifting variables" or use elaborate version as given below with additional lemma about how every $x_s$ refers to the same term PLUS variable renaming convention***</span>

⟨`lemma:literals_clause_simgeq`⟩ *Let $\lambda$ be a literal in a clause $C$ occurring in a resolution refutation of $\Gamma \cup \Delta$. Then $\mathrm{AI}_{\mathrm{cl}}(C)$ contains a literal $\lambda_{\mathrm{AIcl}}$ such that $\lambda_{\mathrm{AIcl}} \gtrsim \ell[\lambda]$, where $\gtrsim$ is defined as follows:*

$$\varphi \gtrsim \varphi' \Leftrightarrow \begin{cases} P = P' \wedge \bigwedge_{i=1}^{n} s_i \gtrsim s_i' & \text{if } \varphi = P(s_1, \ldots, s_n) \text{ and } \varphi' = P'(s_1', \ldots, s_n') \\ f = f' \wedge \bigwedge_{i=1}^{n} s_i \gtrsim s_i' & \text{if } \varphi = f(s_1, \ldots, s_n) \text{ and } \varphi' = f'(s_1', \ldots, s_n') \\ x = x' & \text{if } \varphi, \varphi' \text{ are non-lifting variables, } \varphi = x \text{ and } \varphi' = x' \\ s' \text{ is an instance of } s & \text{if } \varphi, \varphi' \text{ are lifting variables, } \varphi = z_s \text{ and } \varphi' = z_{s'} \end{cases}$$

*For resolved or factorised literals $\lambda$ of an inference $\iota$ with $\tau = \mathrm{au}(\iota)$ we furthermore have that $\ell[\lambda_{\mathrm{AIcl}}\sigma]\tau \gtrsim \ell[\lambda\sigma]$.*

<span style="color:red">*introduce definition for characterising the relation between $C$ and $\mathrm{AI}_{\mathrm{cl}}(C)$*</span>

*Proof.* We proceed by induction on the resolution refutation.

Base case. If for a clause $C$ either $C \in \Gamma$ or $C \in \Delta$ holds, then $\mathrm{AI}_{\mathrm{cl}}(C) = \ell[C]$. Therefore for every literal $l$ in $C$, there exists a literal $l_{\mathrm{AIcl}}$ in $\mathrm{AI}_{\mathrm{cl}}(C)$ such that $\ell[l] = l_{\mathrm{AIcl}}$, which implies $l_{\mathrm{AIcl}} \gtrsim \ell[l]$.

Resolution. If the clause $C$ is the result of a resolution step $\iota$ of $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ using a unifier $\sigma$ such that $l\sigma = l'\sigma$, then let $\tau = \mathrm{au}(\iota)$. Let $\lambda$ be a literal in $C_1$ or $C_2$. Note that every literal in $C$ is of the form $\lambda\sigma$. By the induction hypothesis, there is a literal in $\mathrm{AI}_{\mathrm{cl}}(C_1)$ or $\mathrm{AI}_{\mathrm{cl}}(C_2)$ respectively such that $\lambda_{\mathrm{AIcl}} \gtrsim \ell[\lambda_{\mathrm{AIcl}}]$. If $\lambda \notin \{l, l'\}$, then $\ell[\lambda_{\mathrm{AIcl}}\sigma]\tau$ is contained in $\mathrm{AI}_{\mathrm{cl}}(C)$. Hence in any case, it remains to show that $\ell[\lambda_{\mathrm{AIcl}}\sigma]\tau \gtrsim \ell[\lambda\sigma]$.

We perform an induction on the structure of $\lambda_{\mathrm{AIcl}}$ and $\lambda$ by letting $p$ be the position of the current term in the induction and $t_{\mathrm{AIcl}} = \lambda_{\mathrm{AIcl}}|_p$ as well as $t = \lambda|_p$.

- Suppose that $t$ is a non-lifting variable. As by the induction hypothesis $\ell[t_{\mathrm{AIcl}}] \gtrsim t$, $t_{\mathrm{AIcl}}$ is a non-lifting variable as well and $t = t_{\mathrm{AIcl}}$. But then $\ell[t_{\mathrm{AIcl}}\sigma] = \ell[t\sigma]$. If $\tau$ is trivial on $\ell[t_{\mathrm{AIcl}}\sigma]$, we are done as then $\ell[t_{\mathrm{AIcl}}\sigma]\tau = \ell[t\sigma]$, so assume that it is not.

  But by the definition of au, the substitutions in $\tau$ only update lifting variables to correspond to the terms in the clause of the actual resolution derivation. More formally, $\ell[t_{\mathrm{AIcl}}\sigma]\tau = z_s$ for some term $s$ implies that $\ell[\lambda\sigma]|_p = z_s$, but then $z_s = t$.

  **<span style="color:red">this argument only holds for terms in the resolved literals, see remark in lemma statement</span>**

  > outsource this thought to lemma after definition of au in case needed elsewhere

- Suppose that $t$ is colored term. Then $\ell[t]$ is a lifting variable and by the induction hypothesis, $t_{\mathrm{AIcl}}$ is one as well such that $\ell[t]$ is an instance of $t_{\mathrm{AIcl}}$. As lifting variables are not affected by the unifications occurring in resolution derivations, we only need to consider modifications by means of $\tau$. But as we have seen in the previous case, if $\tau$ substitutes $\ell[t_{\mathrm{AIcl}}\sigma]$, then it does so by $t$.

  > lemma

  Hence we obtain that $\ell[t_{\mathrm{AIcl}}\sigma]\tau \gtrsim \ell[t\sigma]$.

- Suppose that $t$ is a grey term of the form $f(s_1, \ldots, s_n)$. Then $\ell[t] = f(\ell[s_1], \ldots, \ell[s_n])$ and by the induction hypothesis, $t_{\mathrm{AIcl}} = f(r_1, \ldots, r_n)$ such that $\bigwedge_{i=1}^n r_i \gtrsim \ell[s_i]$. By the induction hypothesis applied to the parameters of $\ell[t]$ and $\ell[t_{\mathrm{AIcl}}]$, we obtain that $\ell[r_i\sigma]\tau \gtrsim \ell[s_i\sigma]$ for $1 \le i \le n$. Hence $f(\ell[r_1\sigma], \ldots, \ell[r_n\sigma]) \gtrsim f(\ell[s_1\sigma], \ldots, \ell[s_n\sigma])$, which however is nothing else than $\ell[t_{\mathrm{AIcl}}\sigma] \gtrsim \ell[t\sigma]$.

Factorisation. If the clause $C$ is the result of a factorisation, then we can argue analoguously as for resolution.    $\square$

⟨d_literal_like_lifted_literal⟩ **Lemma 15.** *Let $l$ be a resolved or factorised literal of a resolution derivation step $\iota$ employing the unifier $\sigma$ such that $l\sigma = l'\sigma$ and let $\tau = \mathrm{au}(\iota)$. Then $\ell[l_{\mathrm{AIcl}}\sigma]\tau = \ell[l\sigma]$.*

*Proof.* By Lemma 14, we obtain that $\ell[l_{\mathrm{AIcl}}\sigma]\tau \gtrsim \ell[l\sigma]$. Note that the $\gtrsim$-relation guarantees that pairs of predicates and terms in this relation are equal up to the index of their lifting variables. Hence it only remains to show that the lifting variables of $\ell[l_{\mathrm{AIcl}}\sigma]\tau$ and $\ell[l\sigma]$ match. But the definition of au, $\tau$ substitutes any lifting variable at position $p$ of $\ell[l_{\mathrm{AIcl}}\sigma]$ by the lifting variable $\ell[l\sigma]|_p$. $\qquad\square$

⟨lemma:resolved_literals_equal⟩ **Lemma 16.** *Let $l$ and $l'$ be the resolved or factorised literals of a resolution derivation step $\iota$ employing the unifier $\sigma$ such that $l\sigma = l'\sigma$ and let $\tau = \mathrm{au}(\iota)$. Then $\ell[l_{\mathrm{AIcl}}\sigma]\tau = \ell[l'_{\mathrm{AIcl}}\sigma]\tau$.*

*Proof.* By Lemma 15, we obtain that $\ell[l_{\mathrm{AIcl}}\sigma]\tau = \ell[l\sigma]$ and $\ell[l'_{\mathrm{AIcl}}\sigma]\tau = \ell[l'\sigma]$. But due to $l\sigma \equiv l'\sigma$, it holds that $\ell[l\sigma] = \ell[l'\sigma]$. $\qquad\square$

⟨lemma:gamma_entails_aide⟩ **Lemma 17.** *Let $\pi$ be a resolution refutation of $\Gamma \cup \Delta$. Then for clauses $C$ in $\pi$, $\Gamma \models \mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$.*

*Proof.* We proceed by induction of the strengthening $\Gamma \models \mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C_{\Gamma})^3$.

Base case. For $C \in \Gamma$, $\mathrm{AI}^{\Delta}_{\mathrm{cl}}(C_{\Gamma}) = \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C) = \ell_{\Delta}[C] = C$, so $\Gamma \models \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C_{\Gamma})$. Otherwise $C \in \Delta$ and hence $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) = \top$.

Resolution. Suppose the last rule application is an instance $\iota$ of resolution. Then it is of the following form:

$$\frac{C_1 : D \vee l \qquad C_2 : E \vee \neg l'}{C : (D \vee E)\sigma} \qquad l\sigma = l'\sigma$$

Let $\tau = \mathrm{au}(\iota)$. We introduce the following abbreviations:

$\mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_1)_{\Gamma})^* = \mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_1)_{\Gamma}) \backslash \{(l_{\mathrm{AIcl}^{\Delta}})_{\Gamma}\}$

$\mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_2)_{\Gamma})^* = \mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_2)_{\Gamma}) \backslash \{\neg(l'_{\mathrm{AIcl}^{\Delta}})_{\Gamma}\}$

Note that $\mathrm{AI}^{\Delta}_{\mathrm{cl}}(C) = \ell_{\Delta}[\mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_1)_{\Gamma})^*\sigma]\tau \vee \ell_{\Delta}[\mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_2)_{\Gamma})^*\sigma]\tau$.

Employing these, the induction hypothesis yields $\Gamma \models \mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_1) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_1)_{\Gamma})^* \vee (l_{\mathrm{AIcl}^{\Delta}})_{\Gamma}$ as well as $\Gamma \models \mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_2) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_2)_{\Gamma})^* \vee \neg(l'_{\mathrm{AIcl}^{\Delta}})_{\Gamma}$. By Lemma 11, $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_i)$ and $\mathrm{AI}^{\Delta}_{\mathrm{cl}}(C_i)$ for $i \in \{1, 2\}$ do not contain $\Delta$-colored symbols. Hence by Lemma 12, pulling the lifting inwards using Lemma 1 and applying $\tau$, we obtain:

$\Gamma \stackrel{(\circ)}{\models} \ell[\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_1)_{\Gamma})^*\sigma]\tau \vee \ell[(l_{\mathrm{AIcl}^{\Delta}})_{\Gamma}\sigma]\tau$

$\Gamma \stackrel{(*)}{\models} \ell[\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_2)\sigma]\tau \vee \ell[\mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_2)_{\Gamma})^*\sigma]\tau \vee \neg\ell[(l'_{\mathrm{AIcl}^{\Delta}})_{\Gamma}\sigma]\tau$

We continue by a case distinction on the color of $l$:

---

[3]Recall that as in Lemma **??**, $D_{\Phi}$ denotes the clause created from the clause $D$ by removing all literals which are not contained $\mathrm{L}(\Phi)$.

1. Suppose that $l$ is $\Gamma$-colored. Then $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) = \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_2)\sigma]\tau$. As $l$ is $\Gamma$-colored, $(l_{\mathrm{AIcl}^{\Delta}})_{\Gamma} = l_{\mathrm{AIcl}^{\Delta}}$ and as $l\sigma = l'\sigma$, also $(l'_{\mathrm{AIcl}^{\Delta}})_{\Gamma} = l'_{\mathrm{AIcl}^{\Delta}}$. By Lemma 16, $\ell[l_{\mathrm{AIcl}^{\Delta}}\sigma]\tau = \ell[l'_{\mathrm{AIcl}^{\Delta}}\sigma]\tau$. Hence we can perform a resolution step on $(\circ)$ and $(*)$ to arrive at $\Gamma \models \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_1)_{\Gamma})^*\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_2)\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_2)_{\Gamma})^*\sigma]\tau$. This is however by Lemma 1 nothing else than $\Gamma \models \mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$.

2. Suppose that $l$ is $\Delta$-colored. Then $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) = \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1)\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_2)\sigma]\tau$. As $l$ and $l'$ are $\Delta$-colored, $(\circ)$ and $(*)$ reduce to $\Gamma \models \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_1)_{\Gamma})^*\sigma]\tau$ and $\Gamma \models \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_2)\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_2)_{\Gamma})^*\sigma]\tau$ respectively. These however imply that $\Gamma \models (\ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1)\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_2)\sigma]\tau) \vee \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_1)_{\Gamma})^*\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_2)_{\Gamma})^*\sigma]\tau$, which in turn is nothing else than $\Gamma \models \mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$.

3. Suppose that $l$ is grey. Then $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) = (\neg \ell[l'_{\mathrm{AIcl}^{\Delta}}\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1)\sigma]\tau) \vee (\ell[l_{\mathrm{AIcl}^{\Delta}}\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_2)\sigma]\tau)$.
   Let $M$ be a model of $\Gamma$. Suppose that $M \models \mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$ as otherwise we are done. Hence $M \models \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_1)_{\Gamma})^*\sigma]\tau$ and $M \models \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_2)_{\Gamma})^*\sigma]\tau$ and $(\circ)$ and $(*)$ reduce to $\Gamma \models \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1)\sigma]\tau \vee \ell[l_{\mathrm{AIcl}^{\Delta}}\sigma]\tau$ and $\Gamma \models \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_2)\sigma]\tau \vee \ell[l'_{\mathrm{AIcl}^{\Delta}}\sigma]\tau$ respectively. As by Lemma 16 $\ell[l_{\mathrm{AIcl}^{\Delta}}\sigma]\tau = \ell[l'_{\mathrm{AIcl}^{\Delta}}\sigma]\tau$, a case distinction on the truth value of $\ell[l_{\mathrm{AIcl}^{\Delta}}\sigma]\tau$ in $M$ shows that $M \models \mathrm{AI}_{\mathrm{mat}}^{\Delta}(C)$.

**Factorisation.** Suppose the last rule application is an instance of factorisation. Then it is of the following form:

$$\frac{C_1 : l \vee l' \vee D}{C : (l \vee D)\sigma} \quad \sigma = \mathrm{mgu}(l, l')$$

Let $\tau = \mathrm{au}(\iota)$. We introduce the abbreviation $\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_1)_{\Gamma})^* \stackrel{\mathrm{def}}{=} \mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_1)_{\Gamma}) \setminus \{(l_{\mathrm{AIcl}})_{\Gamma}, (l'_{\mathrm{AIcl}})_{\Gamma}\}$ and express the induction hypothesis as follows: $\Gamma \models \mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_1)_{\Gamma})^* \vee (l_{\mathrm{AIcl}})_{\Gamma} \vee (l'_{\mathrm{AIcl}})_{\Gamma}$. By Lemma 11, Lemma 12 and Lemma 1 and after applying $\tau$ to the induction hypothesis, we obtain that $\Gamma \models \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_1)_{\Gamma})^*\sigma]\tau \vee \ell[(l_{\mathrm{AIcl}})_{\Gamma}\sigma]\tau \vee \ell[(l'_{\mathrm{AIcl}})_{\Gamma}\sigma]\tau$.

However by Lemma 16, $\ell[(l_{\mathrm{AIcl}})_{\Gamma}\sigma]\tau = \ell[(l'_{\mathrm{AIcl}})_{\Gamma}\sigma]\tau$, hence we can perform a factorisation step to arrive at $\Gamma \models \ell[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{cl}}^{\Delta}((C_1)_{\Gamma})^*\sigma]\tau \vee \ell[(l_{\mathrm{AIcl}})_{\Gamma}\sigma]\tau$. This however is nothing else than $\Gamma \models \mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$. $\qquad \square$

As we have just seen, the formula $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$ now satisfies one condition of interpolants. Using this, we are able to formulate a result on one-sided interpolants, which are defined as follows:

**Definition 18.** Let $\Gamma$ and $\Delta$ be sets of first-order formulas. A *one-sided interpolant* of $\Gamma$ and $\Delta$ is a first-order formula $I$ such that

1. $\Gamma \models I$

2. $\mathrm{L}(I) \subseteq \mathrm{L}(\Gamma) \cap \mathrm{L}(\Delta)$ $\qquad \triangle$

**Proposition 19.** *Let $\Gamma$ and $\Delta$ be sets of first-order forumulas such that $\Gamma \cup \Delta$ is unsatisfiable. Then there is a one-sided interpolant of $\Gamma$ and $\Delta$ which is a $\Pi_1$ formula.*

*Proof.* Let $\pi$ be a resolution refutation of $\Gamma \cup \Delta$. By Lemma 17, $\Gamma \models \mathrm{AI}_{\mathrm{mat}}^{\Delta}(\pi) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(\pi)$, or in other words $\Gamma \models \forall x_1 \ldots \forall x_n\, \mathrm{AI}_{\mathrm{mat}}^{\Delta}(\pi) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(\pi)$, where $x_1, \ldots, x_n$ are the $\Delta$-lifting variables occurring in $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(\pi) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(\pi)$. By Lemma 11, the formula $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(\pi) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(\pi)$ does not contain $\Delta$-colored symbols.

Let $y_1, \ldots y_m$ be the $\Gamma$-lifting variables of $\ell_{\Gamma}^{y}[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(\pi) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(\pi)]$ and

$$I = \forall x_1 \ldots \forall x_n \exists y_1 \ldots \exists y_m \ell_{\Gamma}^{y}[\mathrm{AI}_{\mathrm{mat}}^{\Delta}(\pi) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(\pi)].$$

Note that $I$ does not contain any $\Gamma$-terms. As $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(\pi) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(\pi)$ contains witness terms for every existential quantifier in $I$ with respect to $\Gamma$, $\Gamma \models I$. Hence $I$ is a $\Pi_1$ formula which is a one-sided interpolant for $\Gamma \cup \Delta$. $\qquad\square$

## 1.3  Arrows

<span style="color:red">TODO: transition to ordering of quantified lifting vars</span>

In order to establish the required ordering on the lifting variables, we annotate the literals with arrows. More formally:

**Definition 20** ($\mathrm{AI}_{\mathrm{col}}$)**.** The set of colored literals with respect to a clause $C$ in a resolution derivation is defined as follows:

Base case.  For $C \in \Gamma \cup \Delta$, $\mathrm{AI}_{\mathrm{col}}(C) \stackrel{\mathrm{def}}{=} \varnothing$.

Resolution.  Suppose the clause $C$ is the result of a resolution step $\iota$ of $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ with $\sigma = \mathrm{mgu}(\iota)$ and $\tau = \mathrm{au}(\iota)$. Then:

$\mathrm{AI}_{\mathrm{col}}(C) \stackrel{\mathrm{def}}{=} \{\ell[\varphi\sigma]\tau \mid \varphi \in \mathrm{AI}'_{\mathrm{col}}(C)\}$, where

$$\mathrm{AI}'_{\mathrm{col}}(C) \stackrel{\mathrm{def}}{=} \begin{cases} \mathrm{AI}_{\mathrm{col}}(C_1) \cup \mathrm{AI}_{\mathrm{col}}(C_2) \cup \{l_{\mathrm{AIcl}}, l'_{\mathrm{AIcl}}\} & \text{if } l \text{ is a colored literal} \\ \mathrm{AI}_{\mathrm{col}}(C_1) \cup \mathrm{AI}_{\mathrm{col}}(C_2) & \text{if } l \text{ is a grey literal} \end{cases}$$

Factorisation. If the clause $C$ is the result of a factorisation of $C_1$, then $\mathrm{AI}_{\mathrm{col}}(C) \stackrel{\mathrm{def}}{=} \{\ell[\varphi\sigma]\tau \mid \varphi \in \mathrm{AI}_{\mathrm{col}}(C_1)\}$. $\qquad\qquad\qquad\triangle$

**Definition 21** ($\mathrm{AI}_*$)**.** For a clause $C$, $\mathrm{AI}_*(C)$ denotes $\mathrm{AI}_{\mathrm{mat}}(C)$, $\mathrm{AI}_{\mathrm{cl}}(C)$, $\mathrm{AI}_{\mathrm{col}}(C)$. $\qquad\qquad\qquad\triangle$

This definition is convenient as it adheres to the following proposition:

**Proposition 22.** *Let $l$ be a literal in a clause in $\Gamma \cup \Delta$. Then for a clause $C$ in a resolution refutation of $\Gamma \cup \Delta$, $\mathrm{AI}_*(C)$ contains a literal derived from $l$.*

<span style="color:red">TODO: define: descendant (usual stuff, factorisation is merge, resolution is de-facto merge which happens implicitly so no actual merge required)</span>

**Definition 23.** We define a directed graph $G_C$ for every clause $C$ of the derivation. The nodes are of the form $l.tp$, where $l$ denotes a literal and $tp$ a position of a term in $l$, which is not contained in a colored term. The node $l.tp$ in a graph $G_C$ refers to the literal in $\mathrm{AI}_{\mathrm{mat}}(C)$, $\mathrm{AI}_{\mathrm{cl}}(C)$ or $\mathrm{AI}_{\mathrm{col}}(C)$ which is a descendant of $l$. Note that there exists exactly one for every literal of every clause which is an ancestor of $C$. Hence given $C$, $l.tp$ is a well-defined position and the position will usually just be denoted by $p$ or $q$ as abbreviation of $l.tp$. For literals in $\mathrm{AI}_{\mathrm{cl}}(C)$, we usually denote the literal by $l_{\mathrm{AIcl}}$ and the

> write this more formally, there is a relation like $\gtrsim$ here. possibly write that lemma like this

corresponding literal in $C$ by $l$. Note that set of literals in $\mathrm{AI}_{\mathrm{cl}}(C)$ is exactly the set of literals of $C$.

Note that term positions are well defined since arcs do not point into colored terms and are hence not removed by liftings and in the course of the derivation, terms in literals are only modified by substitutions, which does not remove any term which might invalidate a term position.

?⟨def:arrows⟩?

Base case. For $C \in \Gamma \cup \Delta$, we define $G_C$ to be the empty graph.

Resolution. If the clause $C$ is the result of a resolution step of $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ using a unifier $\sigma$ such that $l\sigma = l'\sigma$, we define:

TODO: find meaningful name for index when usage of $\mathcal{A}_1$ is clear

// old idea, basically requires to know term behind lifting var $\mathcal{A}_1 \stackrel{\text{def}}{=} \{(p, q) \mid$ maximal colored term $t$ occurs in $x\sigma$ for some variable $x$, $p$ grey occurrence of $t$ in $C$ (NOTE: does not only mean $C$ actually), $q$ maximal colored term containing colored occurrence of $x$ (where the color of $x$ is different from the color of $t$) in $C_1$ or $C_2\}$

NB: this will only work for $\mathrm{AI}^\triangle$, c.f. 212c:

$\mathcal{A}_1 \stackrel{\text{def}}{=} \{(p, q) \mid$ maximal colored term $t$ occurs in $x\sigma$ for some variable $x$, $p$ grey occurrence of $z_t$ in $\mathrm{AI}_*(C)$, $q$ maximal colored term containing colored occurrence of $x$ (where the color of $x$ is different from the color of $t$) in $C_1$ or $C_2\}$

$\mathcal{A}_2 \stackrel{\text{def}}{=} \{(p, q) \mid$ maximal $\Phi$-term $t$ occurs in maximal $\Psi$-term $s$ in $x\sigma$ for some variable $x$, $p$ grey occurrence of $t$ in $C$, $q$ grey occurrence of $x$ or maximal colored term containing colored occurrence of $x$ in $C_1$ or $C_2$, $(\Phi, \Psi) \in \{(\Gamma, \Delta), (\Delta, \Gamma)\}\}$

$G_C \stackrel{\text{def}}{=} G_{C_1} \cup G_{C_2} \cup \mathcal{A}_1 \cup \mathcal{A}_2$

Factorisation. If the clause $C$ is the result of a factorisation of $C_1 : l \vee l' \vee D$ using a unifier $\sigma$ such that $l\sigma = l'\sigma$, then

$$G_C \stackrel{\text{def}}{=} G_{C_1} \cup G_{C_2}{}^4 \qquad\qquad\qquad \triangle$$

**Definition 24** ($\rightsquigarrow$). For terms $s$ and $t$, $s \rightsquigarrow_{G_C} t$ holds if there is some $p, q$ in the edge set of $G_C$ such that $s$ is a subterm of the term at $p$ and $t$ is a subterm of the term at $t$ such that $s$ and $t$ are not contained in colored terms. (NOTE: in $\mathrm{AI}^\triangle$, $\Gamma$-terms are not colored terms in this sense.) $\triangle$

**Lemma 25.** *Let $l$ and $l'$ be literals such that $\sigma = \mathrm{mgu}(l, l')$ and let $\Lambda = l \vee l'$.*

*Suppose a single-colored $\Phi$-term $s[y]$ containing a variable $y$ occurs in $\Lambda\sigma_{(0,i-1)}$ where $1 \le i \le n$ and $\sigma_0 = \mathrm{id}$ such that a variable $x$ occurs grey in $y\sigma_i$.*

*Then if $x$ only occurs only in single-colored $\Psi$-terms in $\Lambda\sigma_{(0,i-1)}$, $y$ also occurs in a single-colored $\Psi$-term in $\Lambda\sigma_{(0,i-1)}$.*

*Proof.* There must be an occurrence $\hat{y}$ of $y$, say w.l.o.g. in $l\sigma_{(0,i-1)}$, such that $l'\sigma_{(0,i-1)}|_{\hat{y}} = y\sigma_i$. Note that $l\sigma_{(0,i-1)}|_{\hat{y}}$ and $l'\sigma_{(0,i-1)}|_{\hat{y}}$ agree on the prefix and that $x$ occurs grey in $l'\sigma_{(0,i-1)}|_{\hat{y}}$.

---

[4]Note however that the literal $l$ in $C$ has $l$ as well as $l'$ in $C_1$ as predecessors, i.e. the arrows from both of these literals apply implicitly.

Now suppose that $x$ only occurs in single-colored $\Psi$-terms in $\Lambda\sigma_{(0,i-1)}$. Then $l\sigma_{(0,i-1)}|_{\hat{y}}$ is a single-colored $\Psi$-term containing $y$. $\qquad\square$

⟨lemma:color_change⟩ **Lemma 26.** *Let $\pi$ be a resolution refutation of $\Gamma \cup \Delta$ and $C_1$ and $C_2$ be the clauses used in a resolution or factorisation step $\iota$ with $\sigma = \mathrm{mgu}(\iota)$. Let $\Lambda$ be the set of literals contained in $\Gamma \cup \Delta$ with all unifications of $C_1$ and $C_2$ applied. Then if a variable $x$ is a color-changing variable[5] in $\Lambda\sigma_{(0,i)}$, $x$ also occurs grey in $\Lambda\sigma_{(0,i)}$.*

*Proof.* We proceed by induction. Note that in the initial clause sets, no foreign colored terms occur.

We consider a resolution or factorisation step. We perform a nested induction over the construction steps of $\sigma = \sigma_1 \cdots \sigma_n$ with $C_1$ and $C_2$ as induction start.

Suppose that $x$ does not occur grey in $\Lambda\sigma_{(0,i-1)}$ as otherwise we are done. We show that if a variable $x$ occurs in a single-colored $\Phi$-term in $\Lambda\sigma_{(0,i)}$, then (1) it does so in $\Lambda\sigma_{(0,i-1)}$ or (2) there is a color-changing variable $y$ in $\Lambda\sigma_{(0,i-1)}$ such that $x$ occurs grey in $y\sigma_i$. Consider the situations which produce a single-colored $\Phi$-term containing a variable:

- Suppose single-colored $\Phi$-colored term containing $x$ is present in $\Lambda\sigma_{(0,i-1)}$. Then it is as well in $\Lambda\sigma_{(0,i)}$.

- Suppose that a variable $y$ occurs a single-colored $\Phi$-term in $\Lambda\sigma_{(0,i-1)}$ such that $x$ occurs grey in $y\sigma_i$. Suppose furthermore that $x$ does not occur in a single-colored $\Phi$-term in $\Lambda\sigma_{(0,i-1)}$ as otherwise we are done. As by assumption it does not occur grey in $\Lambda\sigma_{(0,i-1)}$, $x$ only occurs in single-colored $\Psi$-terms in $\sigma_{(0,i-1)}$. But as $x$ occurs grey in $y\sigma_i$, there must be an occurrence $\hat{y}$ of $y$ in a resolved or factorised literal, say $l\sigma_{(0,i-1)}$, such that for the other resolved or factorised literal $l'$, $l'\sigma_{(0,i-1)}|_{\hat{y}}$ is a subterm where $x$ occurs grey. But as $l'\sigma_{(0,i-1)}|_{\hat{y}}$ is contained in a single-colored $\Psi$-term, so is $l\sigma_{(0,i-1)}$, hence $y$ is a color-changing variable in $(C_1 \cup C_2)\sigma_{(0,i-1)}$.

- Suppose that a variable $y$ occurs in $(C_1 \cup C_2)\sigma_{(0,i-1)}$ such that $x$ occurs in a single-colored $\Phi$-term in $y\sigma_i$. There must be an occurrence of $y\sigma_i$ in $(C_1 \cup C_2)\sigma_{(0,i-1)}$, but this is nothing else than single-colored $\Phi$-term containing $x$.

Suppose now that $x$ occurs in $\Lambda\sigma_{(0,i)}$ in a single-colored $\Phi$-term as well as in a single-colored $\Psi$-term. If this is the case in $\Lambda\sigma_{(0,i-1)}$, then by the induction hypothesis, $x$ occurs grey in $\Lambda\sigma_{(0,i-1)}$ and consequently also in $\Lambda\sigma_{(0,i)}$.

If otherwise $x$ does not occur in a single-colored $\Phi$- or $\Psi$-term in $\Lambda\sigma_{(0,i-1)}$, then by the reasoning given above, there is a color-changing variable $y$ in $\Lambda\sigma_{(0,i-1)}$ such that $x$ occurs grey in $y\sigma_i$. By the induction hypothesis, then $y$ occurs grey in $\Lambda\sigma_{(0,i-1)}$, which directly implies that $x$ occurs grey in $\Lambda\sigma_{(0,i)}$. $\quad\square$

**Conjectured Lemma 27.** *Let $C$ be a clause in a resolution refutation of $\Gamma \cup \Delta$. If in $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$ a $\Gamma$-term $t[x_s]$ contains a $\Delta$-lifting variable $x_s$, then $x_s$ occurs grey in $\mathrm{AI}_{*}^{\Delta}(C)$.*

---

[5]Recall that a variable is a color-changing if it occurs both in a single-colored $\Gamma$-term and a single-colored $\Delta$-term

*Proof.* Note that if a respective term occurs in $\text{AI}_{\text{mat}}^\Delta(C) \vee \text{AI}_{\text{cl}}^\Delta(C)$, the corresponding non-lifted term is a $\Gamma$-term containing a $\Delta$-term.

Note that it suffices to show that at the derivation step which introduces $s$ as subterm of $t[s]$ $x_s$ occurs grey in $\text{AI}_*^\Delta(C)$ as any potential later modification of $x_s$ is only performed by the substitution $\tau$. However $\tau$ is applied globally in $\text{AI}_*^\Delta$, so it affects each occurrence of $x_s$ in the same manner.

We proceed by induction. Note that for $C \in \Gamma \cup \Delta$, no $\Delta$-lifting variable occurs in a $\Gamma$-term in $\text{AI}_{\text{mat}}^\Delta(C) \vee \text{AI}_{\text{cl}}^\Delta(C)$.

For the induction step, suppose that the condition holds for the clauses $C_1$ and $C_2$ used in a resolution or factorisation step $\iota$. Let $\sigma = \text{mgu}(\iota)$. We continue by induction over the construction steps of $\sigma = \sigma_1 \cdots \sigma_n$ and consider the situations which produce $\Delta$-terms in $\Gamma$-terms:

- Suppose a maximal colored single-colored $\Gamma$-term $t[u]$ in $\Lambda\sigma_{(0,i-1)}$ contains a variable $u$ such that a $\Delta$-term $s'$ occurs grey in $u\sigma_i$ such that $s'\sigma_{(i+1,n)} = s$.

  We assume that $u$ does not occur grey in $\Lambda\sigma_{(0,i-1)}$ as otherwise we are done. If $u$ occurs in $\Lambda\sigma_{(0,i-1)}$ in a single-colored $\Delta$-term, then by Lemma 26, $x$ occurs grey in $\Lambda\sigma_{(0,i)}$ and we are done as well.

  Therefore suppose that $u$ only occurs in single-colored $\Gamma$-terms in $\Lambda\sigma_{(0,i-1)}$. As $u \in \text{dom}(\text{mgu})$, $u$ occurs in a resolved or factorised literal, say at $\hat{u}$ in $l\sigma_{(0,i-1)}$. The other resolved or factorised literal $l'\sigma_{(0,i-1)}$ contains a grey occurrence of $s'$ at the subterm $l'\sigma_{(0,i-1)}$. But as $l\sigma_{(0,i-1)}|_{\hat{u}}$ and $l'\sigma_{(0,i-1)}|_{\hat{u}}$ agree on the prefix, $s'$ occurs in a single-colored $\Gamma$-term in $l'\sigma_{(0,i-1)}$. So by the induction hypothesis, $s'$ occurs grey in $\Lambda\sigma_{(0,i-1)}$. Note that if $s'$ is introduced by $\sigma_{(0,i-1)}$, then due to $l\sigma|_{\hat{u}} = s$, $\sigma$ introduces a grey occurrence of $s$, which in the corresponding literal in $\text{AI}_*^\Delta$ is lifted to yield $x_s$, in which case we are done.

  Otherwise $s'$ has a predecessor $s''$ in $C_1$ or $C_2$ such that $s''$ is a $\Delta$-term which is contained in a $\Gamma$-term and $s''\sigma_{(0,i-1)} = s'$. The lifting variable in $\text{AI}_*^\Delta(C_1)$ or $\text{AI}_*^\Delta(C_2)$ corresponding to $s''$ in general is of the form $x_r$ with $r \neq s$. But Lemma 15, we have that $\ell_\Delta[l_{\text{AIcl}}\sigma]\tau = \ell_\Delta[l\sigma]$ for the resolved or factorised literal $l$ with $\tau = \text{au}(\iota)$. Since $x_r$ occurs in $l_{\text{AIcl}}$ and lifting variables are only modifed by $\tau$, it must be the case that $\{x_r \mapsto x_s\} \in \tau$. But then $x_s$ occurs in $\ell_\Delta[l_{\text{AIcl}}\sigma]\tau$, which is contained in $\text{AI}_{\text{col}}^\Delta(C)$ and hence in $\text{AI}_*^\Delta(C)$.

- Suppose that a variable $u$ occurs in $C_1$ or $C_2$ either grey or in a maximal colored single-colored $\Gamma$-term such that $u\sigma$ contains a multi-colored $\Gamma$-term $t$.

  Then $u$ occurs in a resolved or factorised literal $\lambda\sigma_{(0,i-1)}$ at $\hat{u}$ such that at the other resolved or factorised literal $\lambda'\sigma_{(0,i-1)}$, $\lambda'\sigma_{(0,i-1)}|_{\hat{u}} = t$. But then by the induction hypothesis, $\text{AI}_*^\Delta(C)$ contains grey occurrences for every lifting variable in $t$ and as $t$ occurs in the resolved or factorised literal, but a similar reasoning as given in the other case, $\tau$ substitutes these lifting variables to exactly the ones occurring in $t\sigma$. $\square$

**Example 28.** $R(h(y)) \vee P(f(y))$

$\neg P(f(x_{g(x)})) \lor Q(x_{g(x)})$ such that in the actual clause, it is $g(a)$ and not $g(x)$ any more. Then $\{x_{g(x)} \mapsto x_{g(a)}\} \in \tau$ as desired. $\triangle$

**Conjectured Lemma 29.** *Let $C$ be a clause in a resolution refutation of $\Gamma \cup \Delta$. If in $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$ a $\Gamma$-term $t[x_s]_p$ contains a $\Delta$-lifting variable $x_s$, then $x_s \rightsquigarrow_{G_C} t[x_s]_p$.*

TODO: big fat todo

## 1.4  Combining the results

**Definition 30.** Let $C$ be a clause in a resolution refutation of $\Gamma \cup \Delta$ and $\bar{x}$ be the $\Delta$-lifting variables and $\bar{y}$ the $\Gamma$-lifting variables occurring in $\mathrm{AI}_{\mathrm{mat}}(C)$. $Q(C)$ denotes an arrangement of the elements of $\{\forall x_t \mid x_t \in \bar{x}\} \cup \{\exists y_t \mid y_t \in \bar{y}\}$ such that for two lifting variable $z_s$ and $z_r$ $z_s \rightsquigarrow_{G_C} z_r$ implies that $z_s$ is listed before $z_r$. $\triangle$

### doesn't it work to add arrows based on $C$ (actual clause), then prove correctness via $\mathrm{AI}^{\Delta}$ and $\mathrm{AI}^{\Gamma}$, then just use $\mathrm{AI}$ wihout actually needing the one-sided ones?

there's also a similar result in `-presentable`: $\ell[C] \sim \ell_{\Gamma}[\mathrm{AI}^{\Delta}(C)]$

**Lemma 31.** *Let $\overline{x}$ be the $\Delta$-lifting variables and $\overline{y}$ be the $\Gamma$-lifting variables of $\mathrm{AI}(C)$. Let $\overline{x'}$ be the $\Delta$-lifting variables of $\mathrm{AI}^{\Delta}(C)$.*
*$\Gamma \models \overline{\forall x}\, \mathrm{AI}^{\Delta}(C)$ implies $\Gamma \models \overline{\forall x|\exists y}\, \mathrm{AI}(C)$.*

*Proof.* (sketch) (TODO: don't use $\mathrm{AI}^{\Delta}$) We need to show that every $y$ in $\mathrm{AI}$ corresponds to the same term in $\mathrm{AI}^{\Delta}$ and that every $x$ in $\mathrm{AI}^{\Delta}$ corresponds to the same $x'$ in $\mathrm{AI}$

Then we can insert the terms for $y$ in $\mathrm{AI}$ and they will be equal to $\mathrm{AI}^{\Delta}$. Then as there are less restrictions on the $\mathrm{AI}^{\Delta}$ than there are on the $\mathrm{AI}$, we are done. $\square$

**Theorem 32.** *Let $\pi$ be a resolution refutation of $\Gamma \cup \Delta$. Then $\mathrm{AI}_{\mathrm{mat}}(\pi)$ is an interpolant.*

*Proof.* $\square$

⟨`sec:arrow_quantifier_block`⟩

# outline of arrow part

### 2.0.1 Variable occurrences

Need for var $x$ the set of colored occs and grey occs in initial clauses. lift clauses as usual s.t. to not see any of the colored structure, hence remember only in which max colored term the var is.

for resolution/factorisation, check unifier:

- if $x$ occurs grey in $y\sigma$, then the set of occurrences of $y$ is added to the ones of $x$, col to col and grey to grey
- if $x$ occurs colored in $y\sigma$, then the set of occurrences of $y$ is added to the ones of $x$, col and grey to col

**Definition 33.**

// (apparently not needed) arrows 1: if $x$ occurs in $y\sigma$, add arrow from every *grey* occurrence of $x$ in $C$ to every colored occurrence of $y$ in $C_i$.

arrows 2: if a maximal $\Phi$-colored term $t$ occurs grey in $x\sigma$, add arrow from every grey occurrence of $t$ in $C$ to every $\Psi$-*colored* occurrence of $x$ in $C_i$.

arrows 3: if a maximal $\Phi$-colored term $t$ occurs inside a maximal $\Psi$-colored term $s$ in $x\sigma$, add an arrow from every grey occurrence of $t$ in $C$ to every occurrence of $x$ in $C_i$. $\triangle$

**Lemma 34.** *If in* $\mathrm{AI}_{\mathrm{mat}}^{\Delta}(C) \vee \mathrm{AI}_{\mathrm{cl}}^{\Delta}(C)$ *a* $\Gamma$-*colored term* $t[x_s]$ *contains a* $\Delta$-*lifting variable* $x_s$, *then* $x_s \rightsquigarrow t[x_s]$.

*Proof.*

Suppose term containing max colored term which is $\Delta$-term is introduced into $\Gamma$-colored term.

Then $\Gamma$-colored occ of $u$ in $C_i$ s.t. $\delta_i$ grey in $u\sigma$ ($\delta_i$ is max col term). Hence by arrow 2, arrow from every grey $\delta_i$ to every colored $u$. TODO: as below, need existence

existence 1: If $u$ occurs grey in $C_i$, then there, $\delta_i$ occurs grey in $C$ (this is the necessary color change case $x, f(x)$) and hence the arrow actually exists.

existence 2 proper:

need to show that $\delta_i$ occurs grey given the assumptions.

unification algo produces a chain: $u \mapsto t$, $v \mapsto s$, ...

$u$ only occurs colored in $C_i$. Hence also at $l|_{\hat{u}}$. Therefore $l'|_{\hat{u}}$ is a colored occurrence as well.

chain of colored variables:

if var occurs at some point grey s.t. $\Delta$-term is still complete, then we are done.

if var occurs at some point at position we are unifying with, then we are done by the induction hypothesis.

AUX LEMMA: if a $\Delta$-term enters a $\Gamma$-term, there is an arrow. Later, the terms always look the same as they are affected by the same unifications.

<span style="color:red">TODO: ICI; check example</span>

NEW THING:

chain: either contain variables $v$ s.t. $v\sigma$ contains $\Delta$-term, or term contains $\Delta$-term already (such that outermost symbol matches with the one we get in the end)

in both cases: if term occurs grey, we are done. in this case, we get exactly the lifting var we want.

if term occurs colored (can only be in $\Gamma$), then if we hit a $\Delta$-symbol, we can use the ind hyp. Here, we get the lifting var which just is there. NOTE: different from whether both colors are lifted or just $\Delta$-terms (see 212c).

NEW THING MORE FORMAL:

If for some $u$, $\delta_i$ grey in $u\sigma$ and $u$ occurs in $\Gamma$-term, then $\delta_i$ occurs grey somewhere.

Prf. either $u$ occurs grey, then we are done. Otw. $u$ only occurs colored in $\Gamma$-terms. so $l'|_{\hat{u}}$ also colored.

Note: arguing along subst run.

If $l'|_{\hat{u}}$ contains outermost symbol of $\delta_i$, then have $\Delta$-term in $\Gamma$-term and ind hyp. Otw. $l'|_{\hat{u}}$ contains var $v$ s.t. $\delta_i$ grey in $v\sigma$. Note that now, we can apply the same argument to $v$ and this recursion terminates as mgu algo has terminated.

Suppose multi-colored $\Gamma$-term introduced.

Then $u$ in $C_i$ s.t. $\gamma[\delta_i]$ in $u\sigma$. Hence by arrow 3, arrow from every grey $\delta_i$ to every $u$. <span style="color:red">TODO: need make sure that grey $\delta_i$ exists (exactly $\delta_i$? what if lifted)</span>

existence: $l'|_{\hat{u}}$ is an abstraction of $u\sigma$ different from $u$. if contains multi-colored term $\Rightarrow$ ind hyp. Otw induction, $\Delta$-term must come at some point. we either have other case, or some multi-colored term appears.

$\square$

## 2.1 Garbage

**Lemma 35.** *Let $l$ and $l'$ be variable disjoint literals and $\sigma = \mathrm{mgu}(l, l')$ and $x$ and $y$ be variables such that $x$ occurs in a single-colored $\Delta$-term in $y\sigma$.*

*Then there is a sequence $y_1, \ldots, y_n$ and some $k$ such that $1 \leq k \leq n$, for $1 \leq i \leq k$, $y_i\sigma$ contains a single-colored $\Delta$-term containing $x$ and $y_i\sigma$ does not contain $\Gamma$-symbols, and for $k + 1 \leq i \leq n$, $y_i\sigma$ contains a grey occurrence of $x$.*

*Furthermore, at least one of the following statements holds:*

*1. some single-colored $\Delta$-term containing $x$ occurs in $l$ or $l'$*

*2. some single-colored $\Gamma$-term containing $x$ occurs in $l$ or $l'$ and there is a color change: some $y_i$ is contained in a single-col $\Delta$-term and some $y_{i+1}$ is contained in a single-col $\Gamma$-term*

*possible new text: $y_i$ (and also $y_{i+1}$ occurs grey, and they are unified to $x$ as $i > k$*

*3. $x$ occs grey.*

*additional conjecture: for the first $y_i$, but not $y_1$, the terms are contained in single-col $\Delta$-terms. when the colored tiers are peeled off, the remaining $y_i$ are grey occs of $x$. this is where color changes are possible.*

*Proof.* Let $y_1 = y$.

that for some single-colored $\Delta$-term $r$, $y \mapsto_{\mathrm{mgu}} r$. $r$ furthermore contains $x$ or a variable $z$ such that $z\sigma$ does not contain a $\Gamma$-symbol and contains a grey occurrence of $x$ or a single-colored $\Delta$-term containing $x$.

We build the sequence inductively: By Lemma **??**, there is an occurrence of $y_{i_n}$ of $y_i$ such that $y_{i_n} \mapsto_{\mathrm{mgu}} r$, where $r$ shares the outermost symbol with $y_i\sigma$. As $y_i\sigma$ is a single-colored $\Delta$-term containing $x$, $r$ either contains $x$ in which case $i = k = n$ and item 1 holds and we are done. Otherwise $r$ contains a variable $z$ such that $z\sigma$ contains a grey occurrence of $x$ or $z\sigma$ does not contain $\Gamma$-terms and contains a single-colored $\Delta$-term which contains $x$. Hence $y_{i+1} = z$ and in the first case, $k = i + 1$. Note that the length of $z\sigma$ is a strictly smaller than the length of $y\sigma$, hence the second case can not occur infinitely often.

If we hit the first case and $k = i + 1$, then we continue defining the sequence inductively. Let $y_j$ be such that $y_j\sigma$ contains a grey occurrence of $x$. By Lemma **??**, there is an occurrence $y_{j_n}$ of $y_j$ such that $y_{j_n} \mapsto_{\mathrm{mgu}} s[x]$, where $s[x]$ contains a grey occurrence of $x$. If $s[x]$ occurs grey or in a single-colored $\Delta$-term, when we are done, so suppose it occurs in a single-colored $\Gamma$-term. Note that $y_{j_n}$ is contained in a single-colored $\Phi$-term if and only if $s[x]$ is. Note that $y_k$ is contained in a single-colored $\Delta$-term. As single-colored $\Delta$-terms and single-colored $\Gamma$-terms are not unifiable, there is some $i$, $i < k \leq n$ such that $y_i$ and $y_{i+1}$ occur grey in either $l$ or $l'$, so 2 is the case.

TODO: check indices of $i, k$

□

when we have finished peeling, there is at least one peeling step

varlt?

**Lemma 36.** *Let $l$ and $l'$ be variable disjoint literals and $\sigma = \mathrm{mgu}(l, l')$ such that for a variable $x$, $x\sigma$ contains a grey occurrence of a term $t$.*

*old text: Then there is a sequence of variables $x_1, \ldots, x_n$ with $x_1 = x$ such that for $1 \leq i \leq n$, $t$ occurs grey in $x_i\sigma$ and $x_i$ occurs in one of the literals, say $l_i$, at $l_i|_{\hat{x}_i}$ such that with $l'_i$ being the respective other literal, $l'_i|_{\hat{x}_i}$ contains $x_{i+1}$ for $1 \leq i \leq n - 1$ and $l'_n|_{\hat{x}_n}$ contains the outermost symbol of $t$.*

*new text: Then there is a sequence of variables $x_1, \ldots, x_n$ with $x_1 = x$ such that for $1 \leq i \leq n$, $t$ occurs grey in $x_i\sigma$ and $x_i \mapsto_{\mathrm{mgu}} r[x_{i+1}]$ or $i = n \wedge x_n \mapsto_{\mathrm{mgu}} r_t$, where $r_t$ contains the outermost symbol of $t$*

*Proof.* Let $x_1 = x$ and note that $t$ occurs in $x\sigma$ by assumption. We now consider the execution of the mgu algorithm as defined in **??** and show that for an $x_i$ in the sequence, either we can find an element $x_{i+1}$ which matches the requirement for the sequence or there is an occurrence of $x_i$ which is unified with a term containing the outermost symbol of $t$.

As the mgu algorithm produces a unifier which modifies $x_i$, $x_i$ must occur in a literal, say in $l_i$ at $l_i|_{\hat{x}_i}$, such that at the other literal $l_i'$, $l_i'|_{\hat{x}_i}$ is an abstraction of a term containing $t$ which is different from $x_i$. We distinguish two cases:

- Suppose that $l_i'|_{\hat{x}_i}$ contains the outermost symbol of $t$. Then let $x_n = x_i$.

- Otherwise $l_i'|_{\hat{x}_i}$ contains a variable $v$ such that $t$ occurs grey in $v\sigma$. Let $x_{i+1} = v$. $\qquad\square$

⟨nified_term_starts_somewhere⟩? **Lemma 37.** *Let $l$ and $l'$ be variable disjoint literals and $\sigma = \mathrm{mgu}(l, l')$ such that for a variable $x$, $x\sigma$ contains a term $t$.*

*new text: Then there is a sequence of variables $x_1, \ldots, x_n$ with $x_1 = x$ such that for $1 \leq i \leq n$, $t$ occurs in $x_i\sigma$ and $x_i \mapsto_{\mathrm{mgu}} r[x_{i+1}]$ or $i = n \wedge x_n \mapsto_{\mathrm{mgu}} r_t$, where $r_t$ contains the outermost symbol of $t$*

*Proof.* TODO: (but is virtually a subset of some lemma below) $\qquad\square$

comment

alternate version (unfinished)

Lemma **??** furthermore asserts that $u_n$ occurs in a resolved literal $\lambda$ at $\lambda|_{\hat{u}_n}$ such that $\lambda'|_{\hat{u}_n}$ contains the outermost symbol of the $\Delta$-term $s$, where $\lambda'$ is the respective other resolved literal. As $u_n$ is a colored occurrence and $\lambda\sigma = \lambda'\sigma$, $\lambda'|_{\hat{u}_n}$ is a colored occurrence as well.

- Suppose $\lambda'|_{\hat{u}_n}$ is contained in a $\Gamma$-term. Let $r[x_\varphi]$ be the maximal colored term containing $\lambda'|_{\hat{u}_n}$ and $x_\varphi$ be the lifting variable at the position of the outermost symbol of $s$ in $\lambda'|_{\hat{u}_n}$ in $\mathrm{AI}_{\mathrm{cl}}(C_j)$ for $j = 1$ or $j = 2$. So by the induction hypothesis, $x_\varphi \leadsto_{G_{C_j}} r[x_\varphi]$, hence $x_\varphi$ occurs grey in $\mathrm{AI}_{\mathrm{mat}}^\Delta(C_j)$, $\mathrm{AI}_{\mathrm{cl}}^\Delta(C_j)$ or $\mathrm{AI}_{\mathrm{col}}^\Delta(C_j)$. As however $x_\varphi$ occurs grey in $\lambda'_{\mathrm{AIcl}}$, by the definition of au, $\{x_\varphi \mapsto x_s\} \in \tau$ as $s$ is the term at the position of $x_\varphi$ in $\lambda'\sigma$.

  this is only guaranteed in $\mathrm{AI}^\Delta$, not in $\mathrm{AI}$

  Hence there is a grey occurrence of $x_s$ in $\mathrm{AI}_{\mathrm{mat}}^\Delta(C)$, $\mathrm{AI}_{\mathrm{cl}}^\Delta(C)$ or $\mathrm{AI}_{\mathrm{col}}^\Delta(C)$ and we are done.

- Suppose that $u_i$ for $1 \leq i \leq n$ is contained in a $\Delta$-term which is contained in a $\Gamma$-term.

  TODO:

- Suppose $\lambda'|_{\hat{u}_n}$ is contained in a $\Delta$-term. Due to $\lambda\sigma = \lambda'\sigma$, $\lambda|_{\hat{u}_n}$ is also contained in a $\Delta$-term. As by assumption none of the $u_i$, $1 \leq i \leq n$ is a grey occurrence, there must be a clause which contains two occurrences of $u_i$ such that one of them is a $\Gamma$-occurrence and one is a $\Delta$-occurrence.

- Suppose that one is only gamma and the other only delta

- Suppose that mixed

comment

old proof of smallest colored container

We start by making an observation ($\circledast$): If for two variables $x$ and $y$ it holds that $x$ occurs grey in $y\sigma$, then by Lemma **??**, there exists a sequence $x_1, \ldots, x_n$ such that for $1 \leq i \leq n-1$, $u_i$ occurs in $\lambda|_{\hat{u}_i}$ for a resolved literal $\lambda$ such that the other resolved literal $\lambda'$ has a grey occurrence of $u_{i+1}$ at $\lambda'|_{\hat{u}_i}$. Hence if $u_i$ occurs in a single-colored $\Phi$-colored term in $\lambda|_{\hat{u}_i}$, then $u_{i+1}$ does so too in $\lambda'|_{\hat{u}_i}$ as $\lambda\sigma = \lambda'\sigma$. As $u_{i+1}$ also occurs in $\lambda'|_{\hat{u}_{i+1}}$ for $1 \leq 1 \leq n-1$, i.e. in the same clause as $\lambda'|_{\hat{u}_i}$, then if $\lambda'|_{\hat{u}_{i+1}}$ occurs in a single-colored term which is not $\Phi$-colored, then by the induction hypothesis, $u_{i+1}$ occurs grey in $\mathrm{AI}_{\circledast}(C_i)$ for $i \in \{1, 2\}$ and as $u_{i+1}\sigma$ contains a grey occurrence of $x$, $x$ occurs grey in $\mathrm{AI}_{\circledast}(C)$. Therefore we can assume that all variable of the sequence $x_1, \ldots, x_n$ occur only colored and each of the $x_i$, $1 \leq i \leq n$ is contained in some single-colored $\Phi$-term, as otherwise we are done.

We make another observation ($\circledast$): If for two variables $x$ an $y$ it holds that $y\sigma = s[x]$ a single-colored $\Delta$-term, then we can assume that $x$ occurs grey or in some single-colored $\Delta$-term in $C_1$ or $C_2$. Proof: We proceed by induction on the size of $s[x]$. By Lemma **??**, there is an occurrence of $y_n$ of $y$ in a resolved literal $\lambda$ in say $\lambda[\hat{y}_n]$ such that $\lambda'[\hat{y}_n]$ contains the outermost symbol of $s[x]$.

Suppose for the induction start that $s[x]$ is of size 2. Note that this is the smallest size for a single-colored term containing a variable. Then $\lambda'|_{\hat{y}_n}$ either is $s[x]$, in which case we are done, or $\lambda'|_{\hat{y}_n}$ is $s[z]$ for a variable $z$ such that $z\sigma = x$. Hence $z$ occurs elsewhere in $\lambda'$, say in $\lambda'|_{\hat{z}}$, such that $\lambda|_{\hat{z}}$ is $x$. So if $\lambda'|_{\hat{z}}$ is a grey occurrence or $\lambda'|_{\hat{z}}$ is contained in a single-colored $\Delta$-term, then due to $\lambda\sigma = \lambda'\sigma$, $\lambda|_{\hat{z}}$ is a corresponding occurrence of $x$. Otherwise $\lambda'|_{\hat{z}}$ is contained in a single-colored $\Gamma$-term. meh

TODO: ICI: ind hyp should work for when $z/x$ occur in a single-colored $\Gamma$-term, otw check what we need to have as lemma statement. all is in the resolved literal, so it's gone from the clause in the next step.

We distinguish between all four cases which produce a clause on which the lemma applies:

- Suppose that w.l.o.g. $C_1$ contains a single-colored $\Gamma$-term $s[x]$ which contains $x$ and $C_1$ or $C_2$ contains a single-colored $\Delta$-term containing a variable $y$ such that $x$ occurs grey or in a single-colored $\Delta$-colored in $y\sigma$. Note that the case of an opposite assignment of colors can be argued in a symmetric manner.

    - Suppose that $x$ occurs grey in $y\sigma$: Then by Lemma **??**, there is a variable $y_n$ which occurs in a resolved literal $\lambda$ at $\lambda|_{\hat{y}_n}$ such that $\lambda'|_{\hat{y}_n}$ contains a grey occurrence of $x$. By observation ($\circledast$), $\lambda|_{\hat{y}_n}$ is contained in a single-colored $\Delta$-term. But then so is $\lambda'|_{\hat{y}_n}$, and as clauses are variable disjoint, $s[x]$ also occurs in this clause. So by the induction hypothesis, there is a grey occurrence of $x$ in $\mathrm{AI}_{\circledast}(C_j)$ where $C_j$ is the clause containing $s[x]$, and as $x$ is not affected by $\sigma$, $x$ also occurs grey in $\mathrm{AI}_{\circledast}(C)$.

    > clauses var-disjoint

    - Suppose that $x$ occurs in a single-colored $\Delta$-term $y\sigma$:

      Then by Lemma **??**, either $x$ occurs grey, in which case we are done, or some $y_i$ occurs grey in $l$ or $l'$ such that $y_i\sigma$ contains a grey occurrence of $x$, in which case we are done, or $x$ occurs in a single-colored $\Delta$-term $t[x]$. Then however as $s[x]$ occurs in $C_1$ and clauses are variable disjoint, $t[x]$ occurs in $C_1$ as well and $x$ occurs grey in $\mathrm{AI}_{\circledast}(C_1)$ by the induction hypothesis.

      If a single-colored $\Delta$-term $t[x]$ containing $x$ occurs in $C_1$ or $C_2$, say in $C_j$, then as clauses are variable disjoint, it must be the same clause as $s[x]$. But then $x$ occurs grey in $\mathrm{AI}_{\circledast}(C_j)$ by the induction hypothesis, so assume that no such $t[x]$ occurs in $C_1$ or $C_2$.

      But as a single-colored $\Delta$-term containing $x$ occurs in $y\sigma$, there must be a single-colored $\Delta$-term in $C_1$ or $C_2$ which contains a variable $z$ such that $x$ occurs grey or in a single-colored $\Delta$-term in $z\sigma$. Hence this case is repeated, but as $z\sigma$ is strictly smaller than $y\sigma$, this case can only repeat finitely often.

- Suppose that a single-colored $\Gamma$-term $s[y]$ occurs in $C_i$, $i \in \{1, 2\}$ such that $x$ occurs grey or in a single-colored $\Gamma$-term in $y\sigma$ and a single-colored $\Delta$-term $t[z]$ occurs in $C_j$, $j \in \{1, 2\}$ such that $x$ occurs grey or in a single-colored $\Delta$-term in $z\sigma$.

- 2 other items from arrow-final-conjectures.

# old semi-main lemma reasoning:

- Suppose a single-colored $\Phi$-term $s[x]$ in $C_1$ or $C_2$ contains a grey occurrence of $x$ and a single-colored $\Psi$-term $t[x]$ is introduced in $C$. This is possible by two means:

    1. A single-colored $\Psi$-term $t[z]$ in $C_1$ or $C_2$ contains a variable $z$ such that $x$ occurs grey in $z\sigma$
    2. A variable $u$ occurs in $C_1$ and $C_2$ such that $u\sigma$ contains a single-colored $\Psi$-term containing $x$

We apply Lemma **??** in the first case and Lemma **??** Then by Lemma **??**, at least one of the given three statments holds.

(1) As there is a grey occurrence of $z$ in $C_1$ or $C_2$, there is a grey occurrence of $x$ in $\mathrm{AI}_*(C)$.

(2) then this term occurs in the same clause as $s[x]$ as clauses are variable disjoint and $x$ occurs grey by the induction hypothesis

(3) then by IH, there is a grey occurrence of $z$ in $C_1$ or $C_2$ and hence a grey occurrence of $x$ in $\mathrm{AI}_*(C)$.

- Suppose a single-colored $\Phi$-term $s[y]$ in $C_1$ or $C_2$ contains a variable $y$ such that $x$ occurs grey in $y\sigma$ and a single-colored $\Psi$-term $t[z]$ in $C_1$ or $C_2$ contains a variable $z$ such that $x$ occurs grey in $z\sigma$.

  Then we can apply Lemma **??** to both of $s[y]$ and $t[z]$.

  If any one yields case (1), we are done (as above).

  If any one yields case (3), we are done (IH, as above).

  Hence suppose that both yield case 2. Thus there is a single-colored $\Phi$-term containing $x$ and a single-colored $\Psi$-term containing $x$ in $C_1$ or $C_2$. Note that as clauses are variable disjoint, both these terms must occur in the same clause, say in $C_j$. But then by the induction hypothesis, $x$ occurs grey in $\mathrm{AI}_*(C_j)$ and so also in $\mathrm{AI}_*(C)$.

TODO: ICI; finish this proof
 new distinction:

- $\Phi$-col $s[x]$ in $l/l'$, exists $\Psi$-col $t[z]$ with $z\sigma$ contains grey $x$

- exists $\Phi$-col $s[y]$ with $y\sigma$ contains grey $x$ and exists $\Psi$-col $t[z]$ with $z\sigma$ contains grey $x$

  by new 24 (for col occs of $y$), either

    - $x$ occs grey
    - $y_i$ grey in $C_i$ OR $y_i$ in once in s.c. $\Phi$ and once in s.c. $\Psi$-term
    - some $\Phi$-term $r[x]$ in $C_i$

- $\Phi$-col $s[x]$ in $l/l'$, exists $z$ in $C_i$ s.t. $z\sigma$ contains s.c. $\Psi$-term containing $x$

- exists $y$ in $C_j$ s.t. $y\sigma$ contains s.c. $\Phi$-term $s[x]$ and exists $z$ in $C_i$ s.t. $z\sigma$ contains s.c. $\Psi$-term $t[x]$

  by new 25, either:

    - some $\Phi$-term $r[x]$ in $C_i$
    - $y_i$ grey in $C_i$ OR $y_i$ in once in s.c. $\Phi$ and once in s.c. $\Psi$-term
    - $x$ occs grey

  any of both case 2 or 3 $\Rightarrow$ done.

  otw both case 1, but then ind hyp