Contents

0.1 referenced lemmas from previous sections

⟨lemma:lift_commute⟩? ?⟨lemma:logic_commute⟩?

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

t@@lemmaCommutLiftSubst@data⟩? ?⟨thmt@@lemmaCommut⟨LemmaSubst⟩?

Lemma 2 (Commutativity of lifting and substitution)**.** *Let C be a clause and σ a substitution such that no lifting variable occurs in C or σ. Define σ' 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 $\ell^z_\Phi[C\sigma] = \ell^z_\Phi[C]\sigma'$.

Chapter **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 σ a substitution such that no lifting variable occurs in C or σ. Define σ' 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 σ. 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 σ' is required to conform to the equation $z_{f(x)}\sigma' = z_{f(a)}$, whereas in others, it must hold that $z_{f(x)}\sigma' = z_{f(x)}$. $\triangle$

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 $\mathrm{AI}_{\mathrm{mat}}$ and calculation of $\mathrm{AI}_{\mathrm{cl}}$

$\mathrm{AI}_{\mathrm{mat}}$ 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 $\mathrm{AI}_{\mathrm{cl}}(C)$ (a formal definition will be given below), and for a term t at position p in C, we denote $\mathrm{AI}_{\mathrm{cl}}(C)|_p$ by t_{AIcl}.

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

Definition 4 ($\mathrm{AI}^{\bullet}_{\mathrm{mat}}$ and $\mathrm{AI}^{\bullet}_{\mathrm{cl}}$)**.** Let π be a resolution refutation of $\Gamma \cup \Delta$.

For a clause C in π, $\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C)$ and $\mathrm{AI}^{\bullet}_{\mathrm{cl}}(C)$ are defined as follows:

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

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

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 σ such that $l\sigma = l'\sigma$, then $\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C)$ and $\mathrm{AI}^{\bullet}_{\mathrm{cl}}$ are defined as follows:

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

1. If l is Γ-colored: $\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C) \stackrel{\text{def}}{=} \ell[\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C_1)\sigma] \vee \ell[\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C_2)\sigma]$
2. If l is Δ-colored: $\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C) \stackrel{\text{def}}{=} \ell[\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C_1)\sigma] \wedge \ell[\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C_2)\sigma]$
3. If l is grey: $\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C) \stackrel{\text{def}}{=} (\neg\ell[l'_{\mathrm{AIcl}}\sigma] \wedge \ell[\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C_1)\sigma]) \ \vee \ (\ell[l_{\mathrm{AIcl}}\sigma] \wedge \ell[\mathrm{AI}^{\bullet}_{\mathrm{mat}}(C_2)\sigma])$

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

Note that in $\mathrm{AI}^{\bullet}_{\mathrm{mat}}$ and $\mathrm{AI}^{\bullet}_{\mathrm{cl}}$, it is possible that there for a colored term t in C that $t_{\mathrm{AIcl}} \neq z_t$ 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 $\Delta = \{\neg R(u) \vee Q(g(u))\}$:

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

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

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

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. △

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

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

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

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

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 $\mathrm{AI}^{\bullet}_{\mathrm{mat}}(\Box)$ 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: $\exists y_c \forall x^*((Q(x^*) \wedge R(y_c)) \ \vee\ (\neg Q(x^*) \wedge P(y_c)))$. △

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 ι 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:

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

[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:

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

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

⟨prop:tau_dom_ran⟩ **Proposition 8.** *Let ι 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 ($\mathrm{AI}_{\mathrm{mat}}$ and $\mathrm{AI}_{\mathrm{cl}}$)**.** Let π be a resolution refutation of $\Gamma \cup \Delta$. $\mathrm{AI}_{\mathrm{mat}}(\pi)$ is defined to be $\mathrm{AI}_{\mathrm{mat}}(\Box)$, where $\Box$ is the empty clause derived in π.

For a clause C in π, $\mathrm{AI}_{\mathrm{mat}}(C)$ and $\mathrm{AI}_{\mathrm{cl}}(C)$ are defined inductively as follows:

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

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

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

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

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

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

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

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

1.2 Lifting the Δ-terms

Definition 10. $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C)$ ($\mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$) for a clause C is defined as $\mathrm{AI}_{\mathrm{mat}}(C)$ ($\mathrm{AI}_{\mathrm{cl}}(C)$) with the difference that in its inductive definition, every lifting $\ell[\varphi]$ for a formula or term φ is replaced by a lifting of only the Δ-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 π of $\Gamma \cup \Delta$. $\mathrm{AI}_{\mathrm{mat}}(C)$ and $\mathrm{AI}_{\mathrm{cl}}(C)$ do not contain colored symbols. $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C)$ and $\mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$ do not contain Δ-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 σ to intermediary formulas is lifted. By Proposition 8, $\mathrm{ran}(\mathrm{au}(\iota))$ for inferences ι in π only consists of lifting variables.

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

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

Proof. $\ell^{z}_{\Phi}[F\sigma]$ is an instance of F: σ substitutes variables either for terms not containing Φ-colored symbols or by terms containing Φ-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. □

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

TODO: check which statement we actually need (resolved literal, clause?)

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

Lemma 14. ***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***

⟨lemma:literals_clause_simgeq⟩ *Let λ 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 λ_{AIcl} such that $\lambda_{\mathrm{AIcl}} \gtrsim \ell[\lambda]$, where $\gtrsim$ is defined as follows:*

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

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

introduce definition for characterising the relation between C and $\mathrm{AI}_{\mathrm{cl}}(C)$

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_{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 ι of $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ using a unifier σ such that $l\sigma = l'\sigma$, then let $\tau = \mathrm{au}(\iota)$. Let λ 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 λ_{AIcl} and λ 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_{AIcl} is a non-lifting variable as well and $t = t_{\mathrm{AIcl}}$. But then $\ell[t_{\mathrm{AIcl}}\sigma] = \ell[t\sigma]$. If τ 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 τ 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 $z_s = t$.

  this argument only holds for terms in the resolved literals, see remark in lemma statement

  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_{AIcl} is one as well such that $\ell[t]$ is an instance of t_{AIcl}. As lifting variables are not affected by the unifications occurring in resolution derivations, we only need to consider modifications by means of τ. But as we have seen in the previous case, if τ 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. □

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

Proof. As $l\sigma = l'\sigma$, it also holds that $\ell[l\sigma] = \ell[l'\sigma]$. By Lemma 14, we obtain that $\ell[l_{\mathrm{AIcl}}\sigma]\tau \gtrsim \ell[l\sigma]$ and $\ell[l'_{\mathrm{AIcl}}\sigma]\tau \gtrsim \ell[l'\sigma]$. Furthermore 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 $\ell[l'_{\mathrm{AIcl}}\sigma]\tau$ match. But by the definition of au, τ substitutes any lifting variable at position p of $\ell[l_{\mathrm{AIcl}}\sigma]$ and $\ell[l'_{\mathrm{AIcl}}\sigma]$ by the lifting variable $\ell[l\sigma]|_p$, thus making them equal. □

⟨lemma:gamma_entails_aide⟩ **Lemma 16.** *Let π be a resolution refutation of $\Gamma \cup \Delta$. Then for clauses C in π, $\Gamma \models \mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$.*

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 ι of resolution. Then it is of the following form:

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

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 Δ-colored symbols. Hence by Lemma 12, pulling the lifting inwards using Lemma 1 and applying τ, we obtain:

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

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

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

1. Suppose that l is Γ-colored. Then $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) = \ell[\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_2)\sigma]\tau$. As l is Γ-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 15, $\ell[l_{\mathrm{AIcl}^\Delta}\sigma]\tau = \ell[l'_{\mathrm{AIcl}^\Delta}\sigma]\tau$. Hence we can perform a resolution step on $(\circ)$ and $(*)$ to arrive at $\Gamma \models \ell[\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_1)_\Gamma)^*\sigma]\tau \vee \ell[\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_2)\sigma]\tau \vee \ell[\mathrm{AI}^{\Delta}_{\mathrm{cl}}((C_2)_\Gamma)^*\sigma]\tau$. This is however by Lemma 1 nothing else than $\Gamma \models \mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$.

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

2. Suppose that l is Δ-colored. Then $\mathrm{AI}^\Delta_{\mathrm{mat}}(C) = \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_1)\sigma]\tau \wedge \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_2)\sigma]\tau$. As l and l' are Δ-colored, ($\circ$) and ($*$) reduce to $\Gamma \models \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}^\Delta_{\mathrm{cl}}((C_1)_\Gamma)^*\sigma]\tau$ and $\Gamma \models \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_2)\sigma]\tau \vee \ell[\mathrm{AI}^\Delta_{\mathrm{cl}}((C_2)_\Gamma)^*\sigma]\tau$ respectively. These however imply that $\Gamma \models (\ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_1)\sigma]\tau \wedge \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_2)\sigma]\tau) \vee \ell[\mathrm{AI}^\Delta_{\mathrm{cl}}((C_1)_\Gamma)^*\sigma]\tau \vee \ell[\mathrm{AI}^\Delta_{\mathrm{cl}}((C_2)_\Gamma)^*\sigma]\tau$, which in turn is nothing else than $\Gamma \models \mathrm{AI}^\Delta_{\mathrm{mat}}(C) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(C)$.
3. Suppose that l is grey. Then $\mathrm{AI}^\Delta_{\mathrm{mat}}(C) = (\neg\ell[l'_{\mathrm{AIcl}^\Delta}\sigma]\tau \wedge \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_1)\sigma]\tau) \vee (\ell[l_{\mathrm{AIcl}^\Delta}\sigma]\tau \wedge \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_2)\sigma]\tau)$.
   Let M be a model of Γ. Suppose that $M \models \mathrm{AI}^\Delta_{\mathrm{cl}}(C)$ as otherwise we are done. Hence $M \models \ell[\mathrm{AI}^\Delta_{\mathrm{cl}}((C_1)_\Gamma)^*\sigma]\tau$ and $M \models \ell[\mathrm{AI}^\Delta_{\mathrm{cl}}((C_2)_\Gamma)^*\sigma]\tau$ and ($\circ$) and ($*$) reduce to $\Gamma \models \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_1)\sigma]\tau \vee \ell[l_{\mathrm{AIcl}^\Delta}\sigma]\tau$ and $\Gamma \models \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_2)\sigma]\tau \vee \ell[l'_{\mathrm{AIcl}^\Delta}\sigma]\tau$ respectively. As by Lemma 15 $\ell[l_{\mathrm{AIcl}^\Delta}\sigma]\tau = \ell[l'_{\mathrm{AIcl}^\Delta}\sigma]\tau$, a case distinction on the truth value of $\ell[l_{\mathrm{AIcl}^\Delta}\sigma]\tau$ in M shows that $M \models \mathrm{AI}^\Delta_{\mathrm{mat}}(C)$.

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

However by Lemma 15, $\ell[(l_{\mathrm{AIcl}})_\Gamma\sigma]\tau = \ell[(l'_{\mathrm{AIcl}})_\Gamma\sigma]\tau$, hence we can perform a factorisation step to arrive at $\Gamma \models \ell[\mathrm{AI}^\Delta_{\mathrm{mat}}(C_1)\sigma]\tau \vee \ell[\mathrm{AI}^\Delta_{\mathrm{cl}}((C_1)_\Gamma)^*\sigma]\tau \vee \ell[(l_{\mathrm{AIcl}})_\Gamma\sigma]\tau$. This however is nothing else than $\Gamma \models \mathrm{AI}^\Delta_{\mathrm{mat}}(C) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(C)$. □

As we have just seen, the formula $\mathrm{AI}^\Delta_{\mathrm{mat}}(C) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(C)$ 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 17. Let Γ and Δ be sets of first-order formulas. A *one-sided interpolant* of Γ and Δ is a first-order formula I such that

1. $\Gamma \models I$
2. $\mathrm{L}(I) \subseteq \mathrm{L}(\Gamma) \cap \mathrm{L}(\Delta)$ △

Proposition 18. *Let Γ and Δ be sets of first-order forumulas such that $\Gamma \cup \Delta$ is unsatisfiable. Then there is a one-sided interpolant of Γ and Δ which is a Π_1 formula.*

Proof. Let π be a resolution refutation of $\Gamma \cup \Delta$. By Lemma 16, $\Gamma \models \mathrm{AI}^\Delta_{\mathrm{mat}}(\pi) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(\pi)$, or in other words $\Gamma \models \forall x_1 \ldots \forall x_n\, \mathrm{AI}^\Delta_{\mathrm{mat}}(\pi) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(\pi)$, where $x_1, \ldots, x_n$ are the Δ-lifting variables occurring in $\mathrm{AI}^\Delta_{\mathrm{mat}}(\pi) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(\pi)$. By Lemma 11, the formula $\mathrm{AI}^\Delta_{\mathrm{mat}}(\pi) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(\pi)$ does not contain Δ-colored symbols.

Let $y_1, \dots y_m$ be the Γ-lifting variables of $\ell^y_\Gamma[\mathrm{AI}^\Delta_{\mathrm{mat}}(\pi) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(\pi)]$ and

$$I = \forall x_1 \dots \forall x_n \exists y_1 \dots \exists y_m \ell^y_\Gamma[\mathrm{AI}^\Delta_{\mathrm{mat}}(\pi) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(\pi)].$$

Note that I does not contain any Γ-terms. As $\mathrm{AI}^\Delta_{\mathrm{mat}}(\pi) \vee \mathrm{AI}^\Delta_{\mathrm{cl}}(\pi)$ contains witness terms for every existential quantifier in I with respect to Γ, $\Gamma \models I$. Hence I is a Π_1 formula which is a one-sided interpolant for $\Gamma \cup \Delta$. □

1.3 Arrows

TODO: transition to ordering of quantified lifting vars

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

Definition 19 ($\mathrm{AI}_{\mathrm{col}}$)**.** 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 ι 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 $\mathrm{AI}_{\mathrm{col}}(C) \stackrel{\mathrm{def}}{=} \{\ell[\varphi\sigma]\tau \mid \varphi \in \mathrm{AI}_{\mathrm{col}}(C_1)\}$. △

Definition 20 (AI_*)**.** For a clause C, $\mathrm{AI}_*(C)$ denotes $\mathrm{AI}_{\mathrm{mat}}(C)$, $\mathrm{AI}_{\mathrm{cl}}(C)$, $\mathrm{AI}_{\mathrm{col}}(C)$. △

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

Proposition 21. *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.*

TODO: define: descendant (usual stuff, factorisation is merge, resolution is de-facto merge which happens implicitly so no actual merge required)

Definition 22. 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_{AIcl} and the 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 σ 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 AI^{Δ}, 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 Φ-term t occurs in maximal Ψ-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 σ such that $l\sigma = l'\sigma$, then

$G_C \stackrel{\text{def}}{=} G_{C_1} \cup G_{C_2}$[4] △

Definition 23 ($\rightsquigarrow$)**.** 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 AI^{Δ}, Γ-terms are not colored terms in this sense.) △

ed_term_starts_somewhere_grey⟩ **Lemma 24.** *Let l and l' be variable disjoint literals and $\sigma = \mathrm{mgu}(l,l')$ such that for a variable x, t occurs grey in $x\sigma$.*

Then there is a sequence of variables $x_1,\dots,x_n$ with $x_1 = x$ such that for $1 \le i \le n-1$, t occurs grey in $x_i\sigma$ and $x_i \mapsto_{\mathrm{mgu}} r[x_{i+1}]$, where x_{i+1} occurs grey in $r[x_{i+1}]$. Furthermore, $x_n \mapsto_{\mathrm{mgu}} r_t$, where r_t contains the outermost symbol of t at a grey position.

TODO: prove here as well: if x_i occurs grey/in s.c. Φ-term, then x_{i+1} occs grey/in s.c. Φ-term due to literals same and term grey in unifier image.

Proof. TODO: accidentally proved below:

POSSIBLE BETTER STATEMENT: There is a sequence of variable $y_1,\dots,y_n$ such that $y_i\sigma$ contains x and $y_i \mapsto_{\mathrm{mgu}} r[y_{i+1}]$ for $1 \le i \le n-1$ where $r[y_{i+1}]$ is a grey term and $y_n \mapsto_{\mathrm{mgu}} r[x]$, where $r[x]$ is a grey term as well or a variable.

Inductive definition: Let $y_1 = y$. For each y_i, $y_i \mapsto_{\mathrm{mgu}} t$ for some t such that t is an abstraction of $y_i\sigma$, which is a term containing a grey occurrence of x. Hence either x occurs in t, then $i = n$. Otherwise x does not occur in t

[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.

and there is a variable in t such that $v\sigma$ contains a grey occurrence of x. Let $y_{i+1} = v$. Note that as σ only changes a finite number of variables, a variable can only be added to the sequence finitely often and cycles are not possible by the nature of the unification algorithm. □

⟨lored_y_sigma_contains_grey_x⟩ **Lemma 25.** *Let a single-colored Φ-term $s[y]$ occur in l or l' such that x occurs grey in $y\sigma$. Then at least one of the following statements holds:*

1. *there is a variable z such that x occurs grey in $z\sigma$ and z occurs grey in l or l'*
⟨27_z_grey⟩
2. *x occurs in a s.c. Φ-term*
⟨27_x_in_sc_phi⟩
3. *there is a variable z such that $z\sigma$ contains a grey occurrence of x and z occurs in either l or l' two times: once in s.c. Φ-term and once in s.c. Ψ-term.*
⟨27_mixed⟩

Proof. By Lemma 24, there is a sequence .... We distinguish on the coloring of y_n.

- Suppose that y_n occurs grey. Then we have established item 1 where $z = y_n$.
- Suppose that y_n occurs in a single-colored Φ-term. Then as $y_n \mapsto_{\text{mgu}} r[x]$ where $r[x]$ contains a grey occurrence of x, x does so as well and we have established item 2.
- Suppose that y_n occurs in a single-colored Ψ-term for $\Psi \neq \Phi$. TODO: this is now proved in lemma 24, drop here As $y_1 = y$, y_1 occurs in a single-colored Φ-term. As for $1 \leq i \leq n-1$, $y_i \mapsto_{\text{mgu}} r[y_{i+1}]$ where y_{i+1} occurs grey in $r[y_{i+1}]$, each successive variable occurs in the same coloring as the last one. As y_1 and y_n are contained in single-colored terms of different colors, there must be some j, $1 \leq j \leq n$, such that y_j occurs in a clause once in a single-colored Φ-term as well as in a single-colored Ψ-term, establishing item 3. □

⟨ma:y_sigma_contains_colored_x⟩ **Lemma 26.** *Let a variable y occur in l or l' such that x occurs in a single-colored Φ-term in $y\sigma$. Then at least one of the following statements holds:*

1. *there is a variable z such that x occurs grey in $z\sigma$ and z occurs grey in l or l'*
2. *a single-colored Φ-term in l or l' contains x*
⟨29_grey_x⟩
3. *there is a variable z such that $z\sigma$ contains a grey occurrence of x and z occurs in either l or l' two times: once in s.c. Φ-term and once in s.c. Ψ-term.*

Proof. TODO: rewrite without the sequence; should be just like an algo and only an induction if i know how to do it properly

We attempt to build a sequence of variables $y_1, \ldots, y_n$ such that $y_i \mapsto_{\text{mgu}} r[y_{i+1}]$, where $r[y_{i+1}]$ contains y_{i+1} and does not contain Ψ-terms. Furthermore for $1 \leq i \leq n-1$, $y_i\sigma$ contains a single-colored Φ-term containing x (and no Ψ-symbols) and $y_n\sigma$ contains a grey occurrence of x (and no Ψ-symbols).

Let $y_1 = y$. $y_i \mapsto_{\text{mgu}} t$.

- Suppose that t contains a single-colored Φ-term containing x. Then we have established item 2 and relinquish the partial sequence.

- Suppose that t contains a variable v such that x occurs grey in $v\sigma$ and v occurs in a single-colored Φ-term in t. Then by Lemma 25 gives the result.
- Suppose that t contains a variable v such that $v\sigma$ contains a single-colored Φ-term containing x and no Ψ-symbols. Then let $y_{i+1} = v$.

Note that since y_i contains a single-colored Φ-term containing x, one of the last two cases must be the case in case the first isn't.

□

⟨ma:smallest_colored_container⟩ **Lemma 27.** *Let a variable x occur in C once in a single-colored Γ-term and once in a single-colored Δ-term.*[5] *Then x occurs grey in $\mathrm{AI}_*(C)$.*

TODO: add formal details above and below if result works out

Proof. We proceed by induction on the resolution refutation:

Base case. Clauses contained in Γ do not contain Δ-terms and clauses contained in Δ do not contain Γ-terms.

Resolution/Factorisation. Suppose the clause C is the result of a resolution step ι of $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ or of a factorisation step ι of $C_1 : l \vee l' \vee D$. Let $\sigma = \mathrm{mgu}(\iota)$. TODO: avoid assigning C_1 twice here in final formulation

We consider an occurrence of a single-colored Φ-term containing x in C. There are three circumstances leading to this situation:

1. A single-colored Φ-term containing x occurs in a preceding clause.
2. A single-colored Φ-term $t[y]$ in a preceding clause contains a variable y such that x occurs grey in $y\sigma$.
3. ⟨27_2⟩ A variable z occurs in a preceding clause such that $z\sigma$ contains a single-colored Φ-term containing x.

⟨27_3⟩ We apply Lemma 25 in the case of 2 and Lemma 26 in the case of 3 to obtain that in any of the cases, at least one of the following statements hold:

[copy formulation from lemma once it's finished there]

Now suppose that x occurs in a single-colored Γ-term and in a single-colored Δ-term in C. By applying the reasoning as just given, we know that one of the three statments holds for both occurrences.

If for any one z grey with $z\sigma$ contains grey x, then done

old way:

If IH-case, then: IH

otw both s.c. Γ and Δ-term respectively $\Rightarrow$ IH as well.

if for any one col change case, then col change var grey by IH, and this is unified to x.

otw both IH case, so one in s.c. Γ and one in s.c. Δ, but due variable disjointness in same clause, that's why IH works here. □

[5] Note that these terms may be subterms of other terms.

application of lemma below: Suppose such a term occurs in a clause. Then suppose that it occurs in same s.c. term in literal, otw grey (we are done) or other color (then IH). then lemma!

a:u_sigma_contains_delta_term⟩ **Lemma 28.** *Context: resolved literals. Suppose a single-colored Γ-term contains a variable u such that a Δ-term s occurs grey in $u\sigma$. Then one of the following statements holds:*

1. *there is a variable z such that s occurs grey in $z\sigma$ and z occurs grey in l or l' TODO: possibly change this everywhere to in $l\sigma$, s occurs grey*
2. *a single-colored Γ-term in l or l' contains s outermost symbol of s and variables such that in total, with the unifier we get s*

Proof. Suppose sequence with each unifying to next one, last one: $u_n \mapsto_{\text{mgu}} r[s]$, where s occurs grey in r. also in lemma, successive variables in same coloring

u from lemma statement occs in $\bar{u}$.

Suppose u_i grey, then done as all $u_i\sigma$ contain grey s, hence case 1

Suppose one u_i occs s.c. Γ and s.c. Δ. by lemma 27, u_i occurs grey and s occs grey as above, hence case 1

Otw, all colored, and as successive vars same coloring, all same s.c. term. Start with Γ, hence all Γ. Hence case 2 (term contains var v at grey pos which has s in grey pos at $v\sigma$), hence s occs grey in Γ-term.

□

Lemma 29. *If in $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$ a Γ-term $t[x_s]_p$ contains a Δ-lifting variable x_s, then $x_s \rightsquigarrow_{G_C} t[x_s]_p$.*

Proof. We proceed by induction.

Base case. For $C \in \Gamma \cup \Delta$, consider that no mixed-colored terms occur in C and hence no Γ-term in $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$ can contain a Δ-lifting variable.

Resolution. Suppose the clause C is the result of a resolution step ι of $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ with $\sigma = \mathrm{mgu}(\iota)$ and $\tau = \mathrm{au}(\iota)$. There are two possible cases in which a Δ-lifting variable x_s can be subterm of a Γ-colored term $t[x_s]_p$ in $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$ such that this has not been the case in C_1 or C_2:

1. Suppose a maximal colored Γ-term in C_1 or C_2 contains a variable u such that s occurs grey in $u\sigma$.

   ⟨25_1⟩ Note that it suffices to show that x_s occurs grey in $\mathrm{AI}^{\Delta}_{*}(C)$, since if we suppose that it does so at position r, then $\mathcal{A}_1$ as defined in Definition 22 contains (r, q) such that $\mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)|_q$ is $t[x_s]_p$. As $\mathcal{A}_1 \subseteq G_C$, this implies $x_s \rightsquigarrow_{G_C} t[x_s]_p$.

   We apply Lemma 28.

   in case 1, s occs grey.

   in case 2, IH for that term, say s': $s' \rightsquigarrow_{G_{C_j}} \gamma'[s']$ s' is maximal Δ-term (else would be contained in r and we would talk about x_r). as Γ-terms not lifted, s' occurs "grey". As s is in range of subst, s occurs in literal being unified, by the definition of au, $\{x_s \mapsto x_r\} \in \tau$

as r is the term at the position of x_s in $\lambda\sigma$ for λ the resolved literal where s' occurs.

Hence there is a grey occurrence of x_s in $\mathrm{AI}^{\Delta}_{*}(C)$. TODO: check this

By Lemma **??**, there is a sequence of variable $u_1, \dots, u_n$ such that $u_1 = u$ and s occurs grey in $u_i\sigma$ for $1 \leq i \leq n$. Note that if any variable u_i occurs grey in C_1 or C_2, then at the corresponding position in C, the term at this position is a grey occurrence of s and we are done. Therefore suppose that $u_1, \dots, u_n$ occur only colored in C_1 and C_2.

this is the ramp!

Note that in the prefix of x_s in $t[x_s]_p$, no Δ-colored symbol occurs as otherwise x_s would not occur in this term. Hence the smallest colored term containing the occurrence of u in the predecessor of $t[x_s]$ is a Γ-term.

Lemma **??** furthermore asserts that u_i occurs in a resolved literal l_i at $l_i|_{\hat{u}_i}$ such that in the respective opposite resolved literal l'_i, $l'_i|_{\hat{u}_i}$ contains u_{i+1} for $1 \leq i \leq n-1$ and $l'_n|_{\hat{u}_n}$ contains the outermost symbol of s. Note that for $1 \leq i \leq n$, u_i occurs at least twice in its respective clause. Note also that as $l_i\sigma = l'_i\sigma$, $l|_{\hat{u}_i}$ and $l'|_{\hat{u}_i}$ share the prefix of $\hat{u}_i$, so if $l|_{\hat{u}_i}$ is contained in a Φ-colored term, then so is the grey occurrence of u_{i+1} in $l'|_{\hat{u}_i}$.

If one of the u_i occurs in a clause twice such that for one occurrence, the smallest colored term containing it is Γ-colored and for the other one, the smallest colored term containing it is Δ-colored, then by Lemma 27, u_i occurs grey in $\mathrm{AI}_{*}(C)$ and we are done. Therefore assume that this situation does not arise for any u_i, $1 \leq i \leq n$.

Hence as the smallest colored term containing the occurrences of u_1 must be Γ-terms, the same holds for u_n. But as $l'_n|_{\hat{u}_n}$ contains the outermost symbol of s, which is a Δ-term, and $l_n\sigma = l'_n\sigma$ and the smallest colored term containing $l_n|_{\hat{u}_n}$ is a Γ-term, $l'_n|_{\hat{u}_n}$ is contained in a Γ-term. Let $r[x_\varphi]$ be hte maximal colored term containing $l'_n|_{\hat{u}_n}$ and x_φ be the lifting variable at the position of the outermost symbol of s in $l'_{n\,\mathrm{AIcl}}|_{\hat{u}_n}$. Let C_j be the clause containing l'_n.

2. Suppose a variable u occurs in C_1 or C_2 such that $u\sigma$ contains a multi-colored Γ-term t.

   Then by Lemma **??**, a variable u_n occurs in a resolved literal l at $l|_{\hat{u}_n}$ such that in the other resolved literal l', $l'|_{\hat{u}_n}$ contains the outermost symbol of t.

   If $l'|_{\hat{u}_n}$ is a multi-colored Γ-term, then by the induction hypothesis, dots

   Otherwise as the outermost symbol of t is Γ-colored, $l'|_{\hat{u}_n}$ contains a Γ-colored term which contains a variable v such that a Δ-term occurs grey in $v\sigma$, where case 1 gives the result, or a multi-colored Γ-term s occurs grey in v. But as s is strictly smaller than t, this case can only repeat finitely often before the other case is reached.

Factorisation. If the clause C is the result of a factorisation of C_1, then TODO: □

1.4 Combining the results

doesn't it work to add arrows based on C (actual clause), then prove correctness via AI^{Δ} and AI^{Γ}, then just use 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 30. *Let $\overline{x}$ be the Δ-lifting variables and $\overline{y}$ be the Γ-lifting variables of* $\mathrm{AI}(C)$. *Let $\overline{x'}$ be the Δ-lifting variables of* $\mathrm{AI}^{\Delta}(C)$.

$\Gamma \models \forall\overline{x}\, \mathrm{AI}^{\Delta}(C)$ *implies* $\Gamma \models \forall\overline{x}|\exists\overline{y}\, \mathrm{AI}(C)$.

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

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

Theorem 31. *Let π be a resolution refutation of $\Gamma \cup \Delta$. Then $\mathrm{AI}_{\mathrm{mat}}(\pi)$ is an interpolant.*

Proof. This needs too many things I don't yet know how to make precise, so let's start with $\Gamma \models \ldots$

$\square$

⟨sec:arrow_quantifier_block⟩

Chapter 2

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 32.

// (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 Φ-colored term t occurs grey in $x\sigma$, add arrow from every grey occurrence of t in C to every Ψ*-colored* occurrence of x in C_i.

arrows 3: if a maximal Φ-colored term t occurs inside a maximal Ψ-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. △

Lemma 33. *If in* $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C) \vee \mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$ *a* Γ*-colored term* $t[x_s]$ *contains a* Δ*-lifting variable* x_s*, then* $x_s \rightsquigarrow t[x_s]$*.*

Proof.

Suppose term containing max colored term which is Δ-term is introduced into Γ-colored term.

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

existence 1: If u occurs grey in C_i, then there, δ_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 δ_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. Δ-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 Δ-term enters a Γ-term, there is an arrow. Later, the terms always look the same as they are affected by the same unifications.

TODO: ICI; check example

NEW THING:

chain: either contain variables v s.t. $v\sigma$ contains Δ-term, or term contains Δ-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 Γ), then if we hit a Δ-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 Δ-terms (see 212c).

NEW THING MORE FORMAL:

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

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

Note: arguing along subst run.

If $l'|_{\hat{u}}$ contains outermost symbol of δ_i, then have Δ-term in Γ-term and ind hyp. Otw. $l'|_{\hat{u}}$ contains var v s.t. δ_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 Γ-term introduced.

Then u in C_i s.t. $\gamma[\delta_i]$ in $u\sigma$. Hence by arrow 3, arrow from every grey δ_i to every u. TODO: need make sure that grey δ_i exists (exactly δ_i? what if lifted)

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

$\square$

2.1 Garbage

⟨single_col_x_in_unif_range_old⟩? **Lemma 34.** *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 Δ-term in $y\sigma$.*

Then there is a sequence $y_1, \dots, y_n$ and some k such that $1 \leq k \leq n$, for $1 \leq i \leq k$, $y_i\sigma$ contains a single-colored Δ-term containing x and $y_i\sigma$ does not contain Γ-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 Δ-term containing x occurs in l or l'*
2. ⟨25_delta_x⟩ *some single-colored Γ-term containing x occurs in l or l' and there is a color change: some y_i is contained in a single-col Δ-term and some y_{i+1} is contained in a single-col Γ-term*
   ⟨25_gamma_x⟩ *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.*

?⟨25_grey_x⟩? *additional conjecture: for the first y_i, but not y_1, the terms are contained in single-col Δ-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 Δ-term r, $y \mapsto_{\mathrm{mgu}} r$. r furthermore contains x or a variable z such that $z\sigma$ does not contain a Γ-symbol and contains a grey occurrence of x or a single-colored Δ-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 Δ-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 Γ-terms and contains a single-colored Δ-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 Δ-term, when we are done, so suppose it occurs in a single-colored Γ-term. Note that y_{j_n} is contained in a single-colored Φ-term if and only if $s[x]$ is. Note that y_k is contained in a single-colored Δ-term. As single-colored Δ-terms and single-colored Γ-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. when we have finished peeling, there is at least one peeling step

TODO: check indices of i, k

□

?⟨lemma:proof_along_mgu_old⟩? **Lemma 35.** *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.* varlt?

old text: Then there is a sequence of variables $x_1, \dots, x_n$ 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 $x_{i+1} = v$. □

⟨unified_term_starts_somewhere⟩? **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 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) □

comment
alternate version (unfinished)
Lemma **??** furthermore asserts that u_n occurs in a resolved literal λ at $\lambda|_{\hat{u}_n}$ such that $\lambda'|_{\hat{u}_n}$ contains the outermost symbol of the Δ-term s, where λ' 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 Γ-term. Let $r[x_\varphi]$ be the maximal colored term containing $\lambda'|_{\hat{u}_n}$ and x_φ 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, $x_\varphi \rightsquigarrow_{G_{C_j}} r[x_\varphi]$, hence x_φ occurs grey in $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C_j)$, $\mathrm{AI}^{\Delta}_{\mathrm{cl}}(C_j)$ or $\mathrm{AI}^{\Delta}_{\mathrm{col}}(C_j)$. As however x_φ occurs grey in λ'_{AIcl}, by the definition of au, $\{x_\varphi \mapsto x_s\} \in \tau$ as s is the term at the position of x_φ in $\lambda'\sigma$. this is only guaranteed in AI^{Δ}, not in AI

  Hence there is a grey occurrence of x_s in $\mathrm{AI}^{\Delta}_{\mathrm{mat}}(C)$, $\mathrm{AI}^{\Delta}_{\mathrm{cl}}(C)$ or $\mathrm{AI}^{\Delta}_{\mathrm{col}}(C)$ and we are done.
- Suppose that u_i for $1 \leq i \leq n$ is contained in a Δ-term which is contained in a Γ-term.

  TODO:
- Suppose $\lambda'|_{\hat{u}_n}$ is contained in a Δ-term. Due to $\lambda\sigma = \lambda'\sigma$, $\lambda|_{\hat{u}_n}$ is also contained in a Δ-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 Γ-occurrence and one is a Δ-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 (⊛): If for two variables x and y it holds that x occurs grey in $y\sigma$, then by Lemma **??**, there exists a sequence $x_1, \dots, x_n$ such that for $1 \leq i \leq n-1$, u_i occurs in $\lambda|_{\hat{u}_i}$ for a resolved literal λ such that the other resolved literal λ' has a grey occurrence of u_{i+1} at $\lambda'|_{\hat{u}_i}$. Hence if u_i occurs in a single-colored Φ-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 Φ-colored, then by the induction hypothesis, u_{i+1} occurs grey in $\mathrm{AI}_{*}(C_i)$ for $i \in \{1, 2\}$ and as $u_{i+1}\sigma$ contains a grey occurrence of x, x occurs grey in $\mathrm{AI}_{*}(C)$. Therefore we can assume that all variable of the sequence $x_1, \dots, x_n$ occur only colored and each of the x_i, $1 \leq i \leq n$ is contained in some single-colored Φ-term, as otherwise we are done.

We make another observation ($*$): If for two variables x an y it holds that $y\sigma = s[x]$ a single-colored Δ-term, then we can assume that x occurs grey or in some single-colored Δ-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 λ 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 λ', 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 Δ-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 Γ-term. meh

TODO: ICI: ind hyp should work for when z/x occur in a single-colored Γ-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 Γ-term $s[x]$ which contains x and C_1 or C_2 contains a single-colored Δ-term containing a variable y such that x occurs grey or in a single-colored Δ-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 λ at $\lambda|_{\hat{y}_n}$ such that $\lambda'|_{\hat{y}_n}$ contains a grey occurrence of x. By observation (⊛), $\lambda|_{\hat{y}_n}$ is contained in a single-colored Δ-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 $\mathrm{AI}_{*}(C_j)$ where C_j is the clause containing $s[x]$, and as x is not affected by σ, x also occurs grey in $\mathrm{AI}_{*}(C)$. clauses var-disjoint
  - Suppose that x occurs in a single-colored Δ-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 Δ-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 $\mathrm{AI}_{*}(C_1)$ by the induction hypothesis.

    If a single-colored Δ-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 $\mathrm{AI}_{*}(C_j)$ by the induction hypothesis, so assume that no such $t[x]$ occurs in C_1 or C_2.

    But as a single-colored Δ-term containing x occurs in $y\sigma$, there must be a single-colored Δ-term in C_1 or C_2 which contains a variable z such that x occurs grey or in a single-colored Δ-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 Γ-term $s[y]$ occurs in C_i, $i \in \{1, 2\}$ such that x occurs grey or in a single-colored Γ-term in $y\sigma$ and a single-colored Δ-term $t[z]$ occurs in C_j, $j \in \{1, 2\}$ such that x occurs grey or in a single-colored Δ-term in $z\sigma$.

- 2 other items from arrow-final-conjectures.

old semi-main lemma reasoning:

- Suppose a single-colored Φ-term $s[x]$ in C_1 or C_2 contains a grey occurrence of x and a single-colored Ψ-term $t[x]$ is introduced in C. This is possible by two means:
  1. A single-colored Ψ-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 Ψ-term containing x

We apply Lemma 26 in the first case and Lemma **??** Then by Lemma 26, 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 Φ-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 Ψ-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 25 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 Φ-term containing x and a single-colored Ψ-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:

- Φ-col $s[x]$ in l/l', exists Ψ-col $t[z]$ with $z\sigma$ contains grey x
- exists Φ-col $s[y]$ with $y\sigma$ contains grey x and exists Ψ-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. Φ and once in s.c. Ψ-term
  - some Φ-term $r[x]$ in C_i
- Φ-col $s[x]$ in l/l', exists z in C_i s.t. $z\sigma$ contains s.c. Ψ-term containing x
- exists y in C_j s.t. $y\sigma$ contains s.c. Φ-term $s[x]$ and exists z in C_i s.t. $z\sigma$ contains s.c. Ψ-term $t[x]$

  by new 25, either:

  - some Φ-term $r[x]$ in C_i
  - y_i grey in C_i OR y_i in once in s.c. Φ and once in s.c. Ψ-term
  - x occs grey

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

  otw both case 1, but then ind hyp