1 Arrow-Algo

1. In the original clauses, find all occurrences of variables.

   Common case: If a variable appears as outermost symbol or only has grey ancestor-terms, add an arrow from it to all other occurrences.

   Uncommon case: if there is more than one occurrence of a variable under a Φ-colored term, add a *weak* dependency between them all (symmetric relation).

   NOTE: this creates double arrows for occurrences at same depth. This appears to be necessary for terms which are only variables, and doesn't hurt if the variable is contained in a term.

2. For each step in the derivation:

   a) Build propositional interpolant using $\mathrm{PI}(C_i)^*$, $i \in \{1, 2\}$, i.e. use ancestor PI without colored terms.

   b) If ancestors of atom added to $\mathrm{PI}(C)$ had arrows, merge them to atom in $\mathrm{PI}(C)$ (i.e. arrows starting in and leading to this atom).

   c) Replace colored terms in $\mathrm{PI}(C)$ (from new atom and unifier applied to $\mathrm{PI}(C_i)^*$) with fresh variables, except if a term has a double ended arrow to another overbinding variable, then use that variable.

      An arrow starts (ends) in one of the new variables if it starts (ends) somewhere in the term it replaced.

   d) Collect quantifiers: from $\mathrm{PI}(C_i)^*$, $i \in \{1, 2\}$ and ones from atom added to $\mathrm{PI}(C)$. Order such that arrows only point to variables to the right AND weakly connected variables appear in the same quantifier block.

      $\bar{Q}_n = \mathrm{sort}(Q_{n_1} \cup Q_{n_2} \cup \text{colored-terms}(l))$

1.1 algo more formally

Every literal in any initial clause set has a globally unique id/number
Ex: $P(y, a, f(z, g(y, b))) \vee Q(x)$
Term position:
0.2.1.0 means first literal, 3rd arg, 2nd arg, fst arg: y
0.1 is a
0.2.1 is $g(z, b)$
ρ calculates the position of a term or the term of a position, depending on the argument type.
for a position p_i, $\rho(p_i)$ denotes whatever p_i refers to in its respective clause.
for a term t, $\rho(t)$ denotes the position in t in its respective clause.
for a position p, $\rho_{\text{lit}}(p)$ denotes the position of the literal
for a position p, $\rho_{\text{term}}(p)$ denotes the position of the term in p_i
$\Rightarrow p = \rho_{\text{lit}}(p). \rho_{\text{term}}(p)$
for a position p, $p \mod i$ denotes p with i least significant places cut off, $0.2.1.0 \mod 2 = 0.2$

1.2 Arrows:

$\mathcal{A}$ is a set of ordered pairs of term positions which point to positions in terms in literals
$\mathcal{W}$ is a set of unordered pairs of term positions which point to positions in terms in literals
w.r.t a refutation π of $\Gamma \cup \Delta$:

1. For each initial clause C in $\Gamma \cup \Delta$:

   Add to $\mathcal{A}$ all (p_1, p_2) in C such that p_1 contains only grey symbol and $\rho(p_1)$ is a variable and $\rho(p_1) = \rho(p_2)$ but $p_1 \neq p_2$.

   Add to $\mathcal{W}$ all $\{p_1, p_2\}$ such that there is a colored symbol in p_1 and a possible different one in p_2 and $\rho(p_1)$ is a var and $\rho(p_1) = \rho(p_2)$.

2. For each C resulting from a resolution step from $C_1 : D \vee l$ and $C_2 : E \vee \neg l$ to $C = D \vee E$ with prop interpolant $\text{PI}(\cdot)$:

   Note: literals are added to the interpolant if they occur in both ancestors.

   Merge the respective ids of l and $\neg l$, i.e. their arrows. Their term structure will be the same, so all arrows point to valid positions.

1.3 algo

NOTE: for now, we assume that every colored-term has a globally unique id i and will be replaced by a variable with this index. This restriction is useful now and could potentially be lifted later, but it's not severe anyway.

Note: when a literal is added to the interpolant, the colored terms in one literal might have already been replaced with a certain variable before. we definitely have to use the same variable for both literals, and if one literal has other dependencies, we should stick with the variable we have.
PROBLEM: terms already replaced by variables still change! need to use same variable anyway, so note above not accurate!

1.3.1 $\mathrm{AI}_{\mathrm{mat}}$ and $\mathrm{AI}_{\mathrm{cl}}$

Here, we define $\mathrm{AI}_{\mathrm{mat}}$, which represents the *matrix* of what will be the interpolant, and $\mathrm{AI}_{\mathrm{cl}}$, which represents the *clauses* in the refutation applied with the same unifications as $\mathrm{AI}_{\mathrm{mat}}$.

1. For each initial clause C, $\mathrm{AI}_{\mathrm{mat}}(C) = \mathrm{PI}(C)$ and $\mathrm{AI}_{\mathrm{cl}}(C) = \ell[C]$.

2. For each C resulting from a resolution step from $C_1 : D \vee l$ and $C_2 : E \vee \neg l$ to $C = (D \vee E)\sigma$ with $l\sigma = l'\sigma$ with prop interpolant $\mathrm{PI}(\cdot)$:

   $l \ldots$ literal in original clause

   $l_{\mathrm{cl}} \ldots$ literal in $\mathrm{AI}_{\mathrm{cl}}$ (with unifications and liftings carried over, such that ind hyp goes through)

   $l\sigma = l'\sigma$, but l_{cl} and l'_{cl} might have been overbound with different variables. Still, they in a sense refer to the same ground literal, so we "can just" "unify" them.

   Prose explanation of formal definition below: Shape must be the same in the sense that grey terms are the same, otherwise there is Φ-replacing-var vs Φ-replacing-var (let arbitrary one win) or Φ-replacing-var vs Φ-term (let var win) or Φ-term and Φ-term (replace both with same var). Also apply this to liftings in $\mathrm{AI}_{\mathrm{cl}}$ and $\mathrm{AI}_{\mathrm{mat}}$ here.

   au is defined on terms which are used as parameters for literals l_{cl} and l'_{cl} which occur in $\mathrm{AI}_{\mathrm{cl}}(C_1)$, $\mathrm{AI}_{\mathrm{cl}}(C_2)$ such that for their corresponding l and l', $l\sigma = l'\sigma$. Note that if one of the arguments of au has assigned a color, the other one either has none or the same color. There cannot be a conflict as otherwise their original form would not be unifiable. Note that $\mathrm{au}(a, b)$ is well-defined, i.e. never maps a variable to two different values as each occurrence of some x_j refers to a term with possible free variables, and since across the definition of au, always the same substitution σ is used as reference, every occurrence of x_j will be mapped to the same variable. (NOTE: this is what yet unproven conjectures in the other pdf are trying to formalize.)

   Let $l_{\mathrm{cl}} = A(a_1, \ldots, a_n)$, $l'_{\mathrm{cl}} = A(b_1, \ldots, b_n)$

   $\mathrm{au}(A(a_1, \ldots, a_n), A(b_1, \ldots, b_n)) = \bigcup_{i=1}^{n} \mathrm{au}(a_i, b_i)$

$$\mathrm{au}(a_{\mathrm{cl}}, b_{\mathrm{cl}}) = \begin{cases} \bigcup_{j=1}^{n} \mathrm{au}(s_j, t_j) & \text{if } a_{\mathrm{cl}} = f_s(\bar{s}) \text{ grey and } b_{\mathrm{cl}} = f_t(\bar{t}) \text{ (includes } f_s \text{ being a constant)} \\ \{x_j \mapsto x_k\} & \text{if } a_{\mathrm{cl}} = x_j \text{ from lifting and } b_{\mathrm{cl}} \text{ a term with } \ell[b_{\mathrm{cl}}\sigma] = x_k \\ & \text{(note that } b_{\mathrm{cl}} \text{ is a variable as otherwise it would be colored} \\ & \text{and lifted) } 1 \\ \{x_k \mapsto x_j\} & \text{if } b_{\mathrm{cl}} = x_k \text{ from lifting and } a_{\mathrm{cl}} \text{ a term with } \ell[a_{\mathrm{cl}}\sigma] = x_j \\ & \text{(note that } a_{\mathrm{cl}} \text{ is a variable as otherwise it would be colored} \\ & \text{and lifted) } 1 \\ \{x_j \mapsto x_m, x_k \mapsto x_m\} & \text{if } a_{\mathrm{cl}} = x_j \text{ and } b_{\mathrm{cl}} = x_k\text{, both from lifting, and } x_m \text{ is the} \\ & \text{lifted term in the unified literal, i.e. } x_m = \ell[a] = \ell[b].\ 2 \\ & \text{More formally, } p = \rho(a) = \rho(b) \text{ and } \rho(\rho_{\mathrm{lit}}(\ell[l\sigma]).\,\rho_{\mathrm{term}}(p)) = \\ & \rho(\rho_{\mathrm{lit}}(\ell[l'\sigma]).\,\rho_{\mathrm{term}}(p)) = t_m \text{ and } \ell[t_m] = x_m. \\ \emptyset & \text{if } \ell[a\sigma] = \ell[b\sigma] = x_j \end{cases}$$

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

- if l and l' don't have the same color:

  $\mathrm{AI}_{\mathrm{mat}}(C) = (\neg\ell[l_{\mathrm{cl}}\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}(C_1)\sigma]\tau) \vee (\ell[l_{\mathrm{cl}}\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}(C_2)\sigma)]\tau$

- if l and l' are Γ-colored :

  $\mathrm{AI}_{\mathrm{mat}}(C) = \ell[\Big(\mathrm{AI}_{\mathrm{mat}}(C_1) \vee \mathrm{AI}_{\mathrm{mat}}(C_2)\Big)\sigma]\tau$

- if l and l' are Δ-colored:

  $\mathrm{AI}_{\mathrm{mat}}(C) = \ell[\Big(\mathrm{AI}_{\mathrm{mat}}(C_1) \wedge \mathrm{AI}_{\mathrm{mat}}(C_2)\Big)\sigma]\tau$

1.3.2 AI

$\mathrm{AI}(C) = Q_1 u_1 \ldots Q_m u_m\, \mathrm{AI}_{\mathrm{mat}}(C)$
$u_1, \ldots, u_m$ are comprised of all x_i and y_i PLUS all free variables in $\mathrm{AI}_{\mathrm{mat}}(C)$.
Q_i is $\exists$ if $u_i = y_i$ for some i, $\forall$ if $u_i = x_i$ for some i. For free variables, Q_i as in C_1/C_2. if $u_i\sigma \neq u_i$, u_i becomes one of x_j or y_j or grey term.
$(p_1, p_2) \in \mathrm{TransitiveClosure}(\mathcal{A})$ implies that $u_i < u_j$ if u_i replaces t_i and u_j replaces t_j and p_1 points into t_i and p_2 points into t_j.

2 proof of propositional aspect of AI

t_subst_commute⟩ **Lemma 1** (Restated from proof without propositional refutation, lemma 1). *Let C be a clause and σ a substitution. Let $t_1, \dots, t_n$ be all maximal Δ-terms in this context, i.e. those that occur in C or $C\sigma$, and $x_1, \dots, x_n$ the corresponding fresh variables to replace the t_i (i.e. none of the x_i occur in C). Define σ' such that for a variable z,*

$$z\sigma' = \begin{cases} x_l & \text{if } z = x_k \text{ and } t_k\sigma = t_l \\ \ell_{\Delta,x}[z\sigma] & \text{otherwise} \end{cases}$$

Then $\ell_{\Delta,x}[C\sigma] = \ell_{\Delta,x}[C]\sigma'$.

Remark (Restriction of Lemma 1). Lemma 1 does not hold in case x_i occurs in C. This can easily be seen using the following counterexample:
Let $\sigma = \{x \mapsto a\}$ and $t_1 = f(x)$ and $t_2 = f(a)$. Then clearly $t_1\sigma = t_2$ and therefore $x_1\sigma' = x_2$. But now consider $x_1\sigma$. As x_1 has its place in the domain of variables to replace colored terms, and σ is taken from a resolution refutation, they do not affect each other. Hence $x_1\sigma = x_1$ and therefore $\ell_{\Delta,x}[x_1\sigma] = x_1$, but $\ell_{\Delta,x}[x_1]\sigma' = \ell_{\Delta,x}[x_1]\sigma' = x_2$.
However such a situation arises naturally if we lift colored terms after every step of the interpolant extraction procedure, as there, the intermediate relative interpolants clearly contains variables to overbind terms, but we also need to treat terms that enter the interpolant by means of unification. △

t_logic_commute⟩ **Lemma 2** (corresponds to Lemma 4.8 in thesis and Lemma 11 in Huang). *Let A and B be first-order formulas and s and t be terms. Then it holds that:*

1. $\ell_{\Phi,x}[\neg A] \Leftrightarrow \neg\ell_{\Phi,x}[A]$
2. $\ell_{\Phi,x}[A \circ B] \Leftrightarrow (\ell_{\Phi,x}[A] \circ \ell_{\Phi,x}[B])$ *for* $\circ \in \{\wedge, \vee\}$
3. $\ell_{\Phi,x}[s = t] \Leftrightarrow (\ell_{\Phi,x}[s] = \ell_{\Phi,x}[t])$

o_colored_terms⟩ **Lemma 3.** $\mathrm{AI}_{\mathrm{mat}}(C)$ *and* $\mathrm{AI}_{\mathrm{cl}}(C)$ *contain only grey terms and variables replacing colored terms. They do not contain colored terms. // true and used*

?⟨corr:lift_ai⟩? **Corollary 4.** *For a clause C in a resolution refutation π of $\Gamma \cup \Delta$:*

1. $\mathrm{AI}_{\mathrm{mat}}(C) = \ell[\mathrm{AI}_{\mathrm{mat}}(C)]$.
2. $\mathrm{AI}_{\mathrm{cl}}(C) = \ell[\mathrm{AI}_{\mathrm{cl}}(C)]$.

// true and unused

g_vars_in_subst⟩ **Lemma 5.** *Lifting variables do not occur in any substitution of a resolution refutation. // true and used, also generally relevant* titute_and_lift⟩ **Lemma 6.** *Let F be a formula without colored terms such that for a set of formulas Φ, $\Phi \models F$. Then $\Phi \models \ell[F\sigma]$ for a substitution σ.*

Proof. Note that substitutions only replace variables. Term positions, which are replaced by grey terms by σ are not affected by the lifting and hold due to being special cases of F.
Term positions, which are replaced by colored term by σ are again reduced to variables. All occurrences of a certain variable in F are substituted by the same term, so as the lifting replaces a certain term always be the same variables, all these occurrences of the variable are replaced by the same variable. □

_clause_similar⟩ **Lemma 7.** *If a literal l occurs in a clause C from a resolution refutation, then $\mathrm{AI}_{\mathrm{cl}}(C)$ contains corresponding a literal l_{cl} such that $l_{\mathrm{cl}} \sim \ell[l]$, where $\sim$ means equal up to the index of the variables which replace colored terms.* // *true and somewhat used*

Proof. Base case: By Definition of $\mathrm{AI}_{\mathrm{cl}}$.
Let C be the result of a resolution step from $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ to $C = (D \vee E)\sigma$. Every literal of C is derived from a literal in C_1 or C_2. Let λ be a literal in C_1. The case for a literal in C_2 is analogous. Note that $\lambda \neq l$ as otherwise λ would not be contained in C.
By assumption $\lambda \in C_1$. Then by the resolution rule application, $\lambda\sigma \in C$.
By the induction hypothesis, there is a $\lambda_{\mathrm{cl}} \in \mathrm{AI}_{\mathrm{cl}}(C_1)$ such that $\lambda_{\mathrm{cl}} \sim \ell[\lambda]$. By the definition of $\mathrm{AI}_{\mathrm{cl}}$, $\ell[\lambda_{\mathrm{cl}}\sigma]\tau \in \mathrm{AI}_{\mathrm{cl}}(C)$ with $\tau = \mathrm{au}(l, l')$.
So we have to show that $\ell[\lambda\sigma] \sim \ell[\lambda_{\mathrm{cl}}\sigma]\tau$.
Remark on τ: τ only replaces lifting terms by other lifting by other lifting terms NB: this is where variable indices may not match.
We perform an induction on the depth of terms in λ (except the non-maximal colored terms).

- Suppose t is a term of size 1 in λ and it is a lifting variable, say z_i. Then by Lemma 5, $t\sigma = t$ and also $\ell[t\sigma] = t$.

  As by the induction hypothesis $t_{\mathrm{cl}} \sim \ell[t]$, $t_{\mathrm{cl}} = z_j$. Hence, $\ell[t_{\mathrm{cl}}\sigma] = t_{\mathrm{cl}}$. By the remark on τ, $\ell[t\sigma] \sim \ell[t_{\mathrm{cl}}\sigma]\tau$.

- Suppose t is a term of size 1 in λ and it is a non-lifting variable, say u.

  As $\ell[u] \sim u_{\mathrm{cl}}$ and u is a variable, $u = u_{\mathrm{cl}}$. But then $u\sigma = u_{\mathrm{cl}}\sigma$ and also $\ell[u\sigma] = \ell[u_{\mathrm{cl}}\sigma]$ and $\ell[u\sigma] \sim \ell[u_{\mathrm{cl}}\sigma]\tau$.

- Suppose t is a term of size 1 in λ and it is a constant. Suppose t is grey. Then it is unaffected by both liftings and substitutions. Otherwise suppose t is colored. Then $\ell[t\sigma]$ is a lifting variable, but as $t_{\mathrm{cl}} = \ell[t]$, so is $\ell[t_{\mathrm{cl}}\sigma]\tau$.

  NB: From the point on where t was lifted, t_{cl} even always refers to exactly the lifting var $\ell[t] = x_k$ for some k. Cf. Lemma 12. Hence this case is no obstacle to showing the statement with $\ell[t] = t_{\mathrm{cl}}$ (and not just $\ell[t] \sim t_{\mathrm{cl}}$).

- Suppose t is of the form $f(t_1, \dots, t_n)$ in λ. Then by the induction hypothesis, $\ell[t_i\sigma] \sim \ell[(t_i)_{\mathrm{cl}}\sigma]\tau$ for $1 \leq i \leq n$.
  - Suppose f is grey. Then f is neither affected by substitutions nor by liftings.
  - Suppose f is colored. We only consider the case of occurrences of maximal colored terms as the other ones are discarded by the lifting. As $t_{\mathrm{cl}} \sim \ell[t]$, t_{cl} is a lifting variable. Hence also $\ell[t_{\mathrm{cl}}\sigma]\tau$ is a lifting variable. But so is $\ell[t\sigma]$. □

    NB: Note that even if it was the case that $\ell[t] = t_{\mathrm{cl}}$ (and not just $\ell[t] \sim t_{\mathrm{cl}}$), $\ell[t\sigma]$ might not be equal to $\ell[t_{\mathrm{cl}}\sigma]$, but only $\ell[t\sigma] \sim \ell[t_{\mathrm{cl}}\sigma]$.
    E.g. $t = f(x)$, $\ell[t] = x_1$, $t_{\mathrm{cl}} = x_1$, $\sigma = \{x \mapsto a\}$. Then $\ell[t\sigma] = \ell[f(a)] = x_2$, but $\ell[t_{\mathrm{cl}}\sigma] = x_1$. τ does not fix this, but could potentially if it is more careful than σ'.
    See also Example 16.

Example 8. TODO: example with terms in π vs AI, similar to 206a and last part of 208a:
$f(x)$ vs x_j
$f(g(y))$ vs x_j (actual term is changed but lifting variable stays the same
$f(g(h(z)))$ vs x_k (now x_j appears in resolution, either this occurrence or another occurrence of this var)
$f(g(h(a)))$ vs x_k (again actual term is changed without changing the lifting variable) △

ma:tau_variable⟩ **Lemma 9.** *For 1 kind of τ entries, b is a non-lifting variable.*

raction_of_term⟩ **Lemma 10.** *Let $a_{\mathrm{cl}} = z_j$ a lifting variable. Then $a = t_j\rho$ for some substitution ρ. Even more, if a substitution $z_j \mapsto z_k$ for lifting variables z_j and z_k occurs, z_k refers to exactly t_k and there exists a substitution ρ' such that $t_k = t_j\rho'$.* // used

specialisation⟩?

NB: this probably also hold in $\mathrm{AI}_{\mathrm{mat}}$ and for terms not occurring $\mathrm{AI}_{\mathrm{cl}}$ as well.

Proof. Base case: z_j is introduced to lift t_j, ρ is the identity function.
Induction step: Suppose z_j refers to $t_j\rho$ for some ρ.
Suppose $t_j\rho$ changes in the course of the resolution derivation. Then it changes by means of unification, say by the unifier σ. Then it changes to $t_j\rho\sigma$, so now z_j refers to $t_j\rho\sigma$.
Suppose z_j changes. By the construction of $\mathrm{AI}_{\mathrm{cl}}/\mathrm{AI}_{\mathrm{mat}}$, lifting variables are not affected by the resolution unifications or the liftings, but only by τ.

- Suppose 1 kind of τ entry, $(z_j \mapsto z_k)$. If z_j is changed to z_k in a resolution step, then the term z_j refers to, say a, is changed to $a\sigma$. We show that $a\sigma = t_k$.

  $a_{\mathrm{cl}} = z_j$. By the induction hypothesis, $a = t_j\rho$ for some ρ. By the resolution step au, $b\sigma = a\sigma$.

  By the definition of au, $\ell[b_{\mathrm{cl}}\sigma] = z_k$ and hence $b_{\mathrm{cl}}\sigma = t_k$. By Lemma 9, b is a non-lifting variable and as by Lemma 7 $\ell[b] \sim b_{\mathrm{cl}}$, $b = b_{\mathrm{cl}}$.

  But then $t_k = b_{\mathrm{cl}}\sigma = b\sigma = a\sigma = t_j\rho\sigma$.

- Suppose (2) kind of τ entry, $(z_j \mapsto z_k)$.

  By the induction hypothesis, $a = t_j\rho$ for some ρ.

  $a\sigma = t_k$.

  Hence $t_j\rho\sigma = a\sigma = t_k$. $\square$

By Lemma 7, we have that $l_{\text{cl}} \sim \ell[l]$. But we can also show that the terms in l only become more specialised, i.e. if a lifting variable z_j occurs in l_{cl}, the corresponding term in $\ell[l]$ is a specialisation of t_j,

_variable_terms⟩ **Lemma 11.** *$(x_j \mapsto x_k) \in \mathrm{au}(l_{\text{cl}}, l'_{\text{cl}})$ with $j \neq k$ implies that t_j (corresponding to x_j) contains a free variable.*
In other words: If t_j does not contain a free variable, then if it is lifted to x_j, τ will never change it to some x_k with $k \neq j$. *// true but unused in actual proof of the procedure*

Proof. Let $\mathrm{au}(s_{\text{cl}}, r_{\text{cl}})$ introduce $\{x_j \mapsto x_k\}$. We perform a case distinction:

- Suppose s_{cl} is a lifting variable and but r_{cl} is not. Then by the definition of au, $s_{\text{cl}} = x_j$ and $\ell[r_{\text{cl}}\sigma] = x_k$. Suppose that t_j does not have a free variable. Then in the resolution derivation, from the point on as it is lifted by x_j the original term does not change, hence $s = t_j$. As in consequence s does not contain free variables, $s\sigma = s$. As $s\sigma = r\sigma$ by the resolution rule application, we have that $s = r\sigma$. But as $\ell[s] = \ell[s_j] = x_j$, we must also have that $\ell[r\sigma] = x_j$. Hence $r\sigma = t_j$.

  As r_{cl} occurs in $\mathrm{AI}_{\text{cl}}(C)$ for some C, it is not a colored term, but as it is not a lifting variable and $r_{\text{cl}}\sigma$ is a colored term, r_{cl} must be a variable. By Lemma 7, as $r \in C$, r_{cl} is equal to s up to the index of lifting variables, hence $r = r_{\text{cl}}$.

  But then $r_{\text{cl}}\sigma = r\sigma = t_j$, so $\ell[r_{\text{cl}}\sigma] = x_j$. But then $j = k$, a contradiction.

- Suppose r_{cl} is a lifting variable but s_{cl} is not. This case can be argued analogously.

- Suppose that both s_{cl} and r_{cl} are a lifting variables. Then by the definition of au, $s = r = t_k$ such that $\ell[t_k] = x_k$. Suppose that t_j does not contain a free variable. Then from the point on where s_{cl} has been lifted, t_j does not change. Therefore $s = t_j$. But then $k = j$, a contradiction. $\square$

mma:jka5a5halat⟩ **Lemma 12.** *If a grey or maximal colored term t in a clause C does not contain a free variable, then for t_{cl} in AI_{cl}, we have that $\ell[t] = t_{\text{cl}}$.* *// this is more of a comment, prove properly before actually using it*

Proof. Either t is there from beginning, then $t_{\text{cl}} = \ell[t]$. Otherwise it was introduced by a substitution, but then it was also introduced in $\mathrm{AI}_{\text{cl}}(C)$ and lifted there.
Substitutions do not affect t_{cl} due to Lemma 5. Hence also liftings do not affect t_{cl}. By Lemma 11, τ does not change t_{cl}.
$\Rightarrow$ t as well as t_{cl} remain invariant $\square$

affected_by_tau⟩ **Lemma 13.** *Let u be a variable in a literal l being unified in a resolution step with l' using σ and let $\ell[u\sigma] = z_k$. Then $z_k\tau = z_k$.* // true and used

Proof. NB: much weird stuff here, but last two paragraphs of first item seem to make sense
NB: possibly try to show that $u\sigma = t_j$ in first item
$u\sigma = t_k$.
as $l\sigma = l'\sigma$, in l' there is t with $u\sigma = t\sigma$, so $t\sigma = t_k$.

- Suppose that $(z_k \mapsto z_j) \in \tau$ (first kind). Suppose that $k \neq j$ as otherwise we are done. Then we have a pair of corresponding terms $(a_{\text{cl}}, b_{\text{cl}})$ in l_{cl} and l'_{cl} respectively such that $a_{\text{cl}} = z_k$ and b_{cl} is such that $\ell[b_{\text{cl}}\sigma] = z_j$ but b_{cl} is not a lifting variable. As by Lemma 3 b_{cl} is not colored, it must be a variable. and as by Lemma 7 $b_{\text{cl}} \sim \ell[b]$, $b = b_{\text{cl}}$. Also $b = t_j$

  By Lemma 7 $a_{\text{cl}} \sim \ell[a]$. By Lemma 10 , there is a substitution ρ such that $t_k\rho = a$. So a is a specialisation of t_k.

  Furthermore due to $l\sigma = l'\sigma$, $a\sigma = b\sigma$. Hence $\ell[b_{\text{cl}}\sigma] = \ell[b\sigma] = \ell[a\sigma]$.

  $a\sigma = b\sigma$, so $t_k\rho\sigma = b\sigma = t_j$.

  By Lemma 11, t_k contains a free variable. As each incomparable clause is variable disjoint, C_1 and C_2 are variable disjoint. Hence t_k can only occur in one of them, w.l.o.g. let it occur in C_1 say at p. really?

  $u\sigma = t\sigma = t_k$

  we only have $(z_k \mapsto z_j)$ if the underlying term is unified, i.e. the variable is replaced. This variable only occurs in this clause (or related ones). But $t\sigma = t_k$, i.e. either t still contains this variable or σ introduces it.

  t cannot still contain it as $t_k\sigma$ removes it, and σ cannot introduce it as C_1 and C_2 are variable disjoint, and it could only add it if a variable from the other clause is unified with it, but that variabel then cannot occur in t as it's from the other clause.

- Suppose that $(z_k \mapsto z_j) \in \tau$ (second kind). Suppose that $k \neq j$ as otherwise we are done.

  Similar reasoning: there is a variable in t_k occurring at least twice (as z_k). but then it must be unified to the same variable. so there, the same terms are present and the lifting variables are set accordingly by this crude method.

  So u is substituted for whatever happens on this other side, i.e. $u\sigma = t_j$

□

ting_variables⟩? **Lemma 14.** *The set of lifting variables, which refer to terms which have free variables, is disjoint for every incomparable clause.* // true but ok to have unused

Proof. The free variables for every inital clause is disjoint. ☐

Apparently, τ establishes equality for the terms in the literals being resolved on and quasi-equality for other literals in the remaining clause:

s_clauses_equal⟩ **Lemma 15.** *Let l_{cl}, l'_{cl} be the literal in $\mathrm{AI}_{\mathrm{cl}}(C_1)$ and $\mathrm{AI}_{\mathrm{cl}}(C_2)$ corresponding to l and l' where C is the result of a resolution step from $C_1 : D \vee l$ and $C_2 : E \vee \neg l'$ to $C = (D \vee E)\sigma$ (i.e. $l\sigma = l'\sigma$). Let $\tau = \mathrm{au}(l_{\mathrm{cl}}\sigma, l'_{\mathrm{cl}}\sigma)$. Then $\ell[l_{\mathrm{cl}}\sigma]\tau = \ell[l'_{\mathrm{cl}}\sigma]\tau$. // true and used*

Proof. Let s_{cl} be a (sub-)term of a parameter of $l_{\mathrm{cl}}\sigma$ and t_{cl} the term at the same term position in $l'_{\mathrm{cl}}\sigma$. Let s and t be their corresponding (sub)-term at the same term position in $l\sigma$ and $l'\sigma$. We show that $\ell[s_{\mathrm{cl}}\sigma]\tau = \ell[t_{\mathrm{cl}}\sigma]\tau$ by induction on the structure of s_{cl} and t_{cl} respectively.
Note that by Lemma 3, s_{cl} and t_{cl} do not contain colored terms. This also implies that only grey terms can contain subterms.
By Lemma 7, $l_{\mathrm{cl}} \sim \ell[l]$ and $l'_{\mathrm{cl}} \sim \ell[l']$.

Lifting variables. Suppose that $s_{\mathrm{cl}} = z_i$ and/or $t_{\mathrm{cl}} = z_j$ for some i and j. Suppose that $s_{\mathrm{cl}} \neq t_{\mathrm{cl}}$ as otherwise we are done. By the resolution rule application $s\sigma = t\sigma$. Cases:

- $s_{\mathrm{cl}} = z_i$ and $t_{\mathrm{cl}} = z_j$ with $i \neq j$. As σ affects neither s_{cl} nor t_{cl}, $\ell[s_{\mathrm{cl}}\sigma] = s_{\mathrm{cl}}$ and $\ell[t_{\mathrm{cl}}\sigma] = t_{\mathrm{cl}}$. We show that $s_{\mathrm{cl}}\tau = t_{\mathrm{cl}}\tau$.
  Note that the function au visits all subterms and combines all mappings it encounters. Hence $\mathrm{au}(s_{\mathrm{cl}}, t_{\mathrm{cl}})$ is part of the final substitution τ. However due to the just established circumstances, $\mathrm{au}(s_{\mathrm{cl}}, t_{\mathrm{cl}}) = \{z_i \mapsto z_m, z_j \mapsto z_m\}$ with m as in the definition of au, so $s_{\mathrm{cl}}\tau = t_{\mathrm{cl}}\tau$. NB: this is the somewhat crude step where all lifting variables in the resolved literal are just reset.
- W.l.o.g. $s_{\mathrm{cl}} = z_i$ and t_{cl} is not a lifting variable. As $t_{\mathrm{cl}} \sim \ell[t]$, t is not a colored term. But due to $s_{\mathrm{cl}} \sim \ell[s]$, s is a colored term. As $s\sigma = t\sigma$, t must be a variable and $t\sigma$ a colored term. So $\ell[t\sigma] = z_k$ for some k. Note that the function au visits all subterms and combines all mappings it encounters. By the construction of au, at $\mathrm{au}(s_{\mathrm{cl}}, t_{\mathrm{cl}})$, $\{z_i \mapsto z_k\}$ is added. Therefore $\ell[s_{\mathrm{cl}}\sigma]\tau = \ell[z_i\sigma]\tau = \ell[z_i]\tau = z_i\tau = z_k$.
  Due to $t_{\mathrm{cl}} \sim \ell[t]$ and as t is a variable, $t = t_{\mathrm{cl}}$. Then $\ell[t_{\mathrm{cl}}\sigma]\tau = \ell[t\sigma]\tau = z_k\tau$. By Lemma 13, $z_k\tau = z_k$.

Grey terms. Suppose that at least one of s_{cl} and t_{cl} is a grey term.

- Suppose that both s_{cl} and t_{cl} are grey terms: By $s_{\mathrm{cl}} = \ell[s]$ and $t_{\mathrm{cl}} = \ell[t]$, and as $s\sigma = t\sigma$, their outermost symbol is the same in all these terms. The equality of the parameters is established by the induction hypothesis. Note that grey constants can be treated as grey functions without parameters.
- Suppose that exactly one of s_{cl} and t_{cl} is a grey terms. W.l.o.g. let s_{cl} be a grey term. Then as $s\sigma = t\sigma$, $s_{\mathrm{cl}} = \ell[s]$ and $t_{\mathrm{cl}} = \ell[t]$, t_{cl} is a variable and $t = t_{\mathrm{cl}}$. Furthermore,

$t_{\text{cl}}\sigma$ is a grey term. Even more, their outermost symbol is the same due to $s\sigma = t\sigma$. Equality of potential parameters in s_{cl} is established by the induction hypothesis.

Variables. Suppose that both s_{cl} and t_{cl} are variables. Suppose that σ is non-trivial on at least s_{cl} or t_{cl}, as otherwise we would be done. Due to $s_{\text{cl}} = \ell[s]$ and $t_{\text{cl}} = \ell[t]$, $s = s_{\text{cl}}$ and $t = t_{\text{cl}}$. As $s\sigma = t\sigma$, the outermost symbol of both $s_{\text{cl}}\sigma$ is the same as the one of $t_{\text{cl}}\sigma$. As the equality of potential parameters of $s_{\text{cl}}\sigma$ and $t_{\text{cl}}\sigma$ is established by the induction hypothesis, we are done. $\square$

_different_term⟩ **Example 16.** We illustrate that the given procedure, if a lifting variable x_k occurs in $\text{AI}_{\text{cl}}(C)$, it does not necessaily mean that t_k occurs in C:
$\Gamma = \{P(f(x)) \vee Q(x)\}$
$\Delta = \{\neg P(y), \neg Q(a)\}$

$$\dfrac{\dfrac{\bot \mid P(f(x)) \vee Q(x) \qquad \top \mid \neg P(y)}{P(x_1) \mid Q(x)} \qquad \top \mid \neg Q(a)}{Q(y_2) \vee P(x_1) \mid \square}$$

Here, x_1 first refers to $f(x)$ and later to $f(a)$. This however is not essential for the correctness of the procedure, and it would be tedious to fix all such x_1 see also corresponding remark in case distinction in Lemma 7. $\triangle$

Lemma 17. *Suppose no* Γ*-term occurs in* Γ*. Then* $\Gamma \models \text{AI}_{\text{mat}}(C) \vee \text{AI}_{\text{cl}}(C)$*.*

Proof. Proof by induction of the strenghtening: $\Gamma \models \text{AI}_{\text{mat}}(C) \vee \text{AI}_{\text{cl}}(C_\Gamma)$.
Base case:
For $C \in \Gamma$, $\text{AI}_{\text{mat}}(C) = \bot$ and $\text{AI}_{\text{cl}}(C) = \ell[C_\Gamma] = \ell_{\Gamma,y}[C]$. By the restriction, $\ell_{\Gamma,y}[C] = C$ and $\Gamma \models C$.
For $C \in \Delta$, $\text{AI}_{\text{mat}}(C) = \top$.
Induction step:

restriction used here

Resolution.

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

We introduce the following abbreviations, where l^* is the literal in AI_{cl} which corresponds to l:

$\text{AI}_{\text{cl}}(C_1)^* = \text{AI}_{\text{cl}}((C_1)_\Gamma) \setminus \{\ell[(l_{\text{cl}})_\Gamma]\}$

$\text{AI}_{\text{cl}}(C_2)^* = \text{AI}_{\text{cl}}((C_2)_\Gamma) \setminus \{\ell[\neg(l'_{\text{cl}})_\Gamma]\}$

$\tau = \text{au}((l_{\text{cl}})_\Gamma, (l'_{\text{cl}})_\Gamma)$

$\text{AI}_{\text{cl}}(C_\Gamma) = \ell[\Big(\text{AI}_{\text{cl}}(C_1)^* \vee \text{AI}_{\text{cl}}(C_2)^*\Big)\sigma]\tau.$

By Lemma 2, $\mathrm{AI}_{\mathrm{cl}}(C_\Gamma) = \ell[\mathrm{AI}_{\mathrm{cl}}(C_1)^*\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{cl}}(C_2)^*\sigma]\tau$.

By the induction hypothesis, $\Gamma \models \mathrm{AI}_{\mathrm{mat}}(C_i) \vee \mathrm{AI}_{\mathrm{cl}}(C_{i\Gamma})$, $i \in \{1,2\}$, or expressed differently:

$\Gamma \models \mathrm{AI}_{\mathrm{mat}}(C_1) \vee \mathrm{AI}_{\mathrm{cl}}(C_1)^* \vee (l_{\mathrm{cl}})_\Gamma$

$\Gamma \models \mathrm{AI}_{\mathrm{mat}}(C_2) \vee \mathrm{AI}_{\mathrm{cl}}(C_2)^* \vee \neg(l'_{\mathrm{cl}})_\Gamma$

By Lemma 3, $\mathrm{AI}_{\mathrm{mat}}(C_1)$ and $\mathrm{AI}_{\mathrm{cl}}(C_1)$ as well as $\mathrm{AI}_{\mathrm{mat}}(C_2)$ and $\mathrm{AI}_{\mathrm{cl}}(C_2)$ do not contain colored terms. Hence by Lemma 6, Lemma 2 and applying τ, we get that

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

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

By Lemma 15, $\ell[(l_{\mathrm{cl}})_\Gamma\sigma]\tau = \ell[(l'_{\mathrm{cl}})_\Gamma\sigma]\tau$.

- If l and l' grey:

  $\mathrm{AI}_{\mathrm{mat}}(C) = (\neg\ell[l_{\mathrm{cl}}\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}(C_1)\sigma]\tau) \vee (\ell[l_{\mathrm{cl}}\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{mat}}(C_2)\sigma)]\tau$

  Suppose for a model M of Γ that $M \not\models \mathrm{AI}_{\mathrm{cl}}(C)$, i.e. $M \not\models \ell[\mathrm{AI}_{\mathrm{cl}}(C_1)\sigma]\tau$ and $M \not\models \ell[\mathrm{AI}_{\mathrm{cl}}(C_2)\sigma]\tau$ as otherwise we would be done. Then by $(\circ)$ and $(*)$:

  $M \models \ell[\mathrm{AI}_{\mathrm{mat}}(C_1)\sigma]\tau \vee \ell[l_{\mathrm{cl}}\sigma]\tau$

  $M \models \ell[\mathrm{AI}_{\mathrm{mat}}(C_2)\sigma]\tau \vee \neg\ell[l'_{\mathrm{cl}}\sigma]\tau$

  By Lemma 15, $\ell[l_{\mathrm{cl}}\sigma]\tau = \ell[l'_{\mathrm{cl}}\sigma]\tau$. By a case distinction on the truth value of $\ell[l_{\mathrm{cl}}\sigma]\tau$ in M, we obtain that $M \models \mathrm{AI}_{\mathrm{mat}}(C)$.

- If l and l' are Γ-colored: $\mathrm{AI}_{\mathrm{mat}}(C) = \ell\Big[\Big(\mathrm{AI}_{\mathrm{mat}}(C_1) \vee \mathrm{AI}_{\mathrm{mat}}(C_2)\Big)\sigma\Big]\tau$

  By Lemma 15, we can do a resolution step on $\ell[l_{\mathrm{cl}}\sigma]\tau$ of $(\circ)$ and $(*)$ to arrive at

  $\Gamma \models \ell[\mathrm{AI}_{\mathrm{mat}}(C_1)\sigma]\tau \ \vee\ \ell[\mathrm{AI}_{\mathrm{cl}}(C_1)^*\sigma]\tau \ \vee\ \ell[\mathrm{AI}_{\mathrm{mat}}(C_2)\sigma]\tau \ \vee\ \ell[\mathrm{AI}_{\mathrm{cl}}(C_2)^*\sigma]\tau$

  This however is by Lemma 2 nothing else than $\Gamma \models \mathrm{AI}_{\mathrm{mat}}(C) \vee \mathrm{AI}_{\mathrm{cl}}(C)$

- If l and l' are Δ-colored: $\mathrm{AI}_{\mathrm{mat}}(C) = \ell\Big[\Big(\mathrm{AI}_{\mathrm{mat}}(C_1) \wedge \mathrm{AI}_{\mathrm{mat}}(C_2)\Big)\sigma\Big]\tau$

  As l is Δ-colored, $(\circ)$ and $(*)$ reduce to:

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

  $\Gamma \models \ell[\mathrm{AI}_{\mathrm{mat}}(C_2)\sigma]\tau \vee \ell[\mathrm{AI}_{\mathrm{cl}}(C_2)^*\sigma]\tau$

  But this implies that

  $\Gamma \models \Big(\ell[\mathrm{AI}_{\mathrm{mat}}(C_1)\sigma]\tau \wedge \ell[\mathrm{AI}_{\mathrm{cl}}(C_1)^*\sigma]\tau\Big) \ \vee\ \ell[\mathrm{AI}_{\mathrm{cl}}(C_1)^*\sigma]\tau \ \vee\ \ell[\mathrm{AI}_{\mathrm{cl}}(C_2)^*\sigma]\tau$

  This however is by Lemma 2 nothing else than $\Gamma \models \mathrm{AI}_{\mathrm{mat}}(C) \vee \mathrm{AI}_{\mathrm{cl}}(C)$. □

3 thoughts

Conjecture 18. *Double ended arrows are not important as terms are overbound with same variable anyway as always same unifier applies.*

4 arrow proof

_same_variables⟩ **Lemma 19.** *Whenever the same variable appears multiple times in* $\mathrm{PI}(C) \vee C$ *for* $C \in \pi$*, there are arrows.*

- *If both variables are contained only in grey terms, there is a double arrow // they unify to exactly the same*
- *If only one variable is only contained in grey terms, there is an arrow from it to the other one // either unify the one in grey term, then other one must be overbound later. if otherwise var in the colored term is unified, we can still overbind the grey one first.*
- *otherwise there are weak arrows between them // have same quantifier, so order does not matter, but want to keep dependencies on both the same*

Proof. By induction. Note: As required by resolution, all initial clauses are variable disjoint.
Base case: In the initial clause sets, consider for a clause C two different positions p_1 and p_2 pointing to the same variable. Then either:

- p_1 and p_2 contain only grey symbols. Then $(p_1, p_2) \in \mathcal{A}$.
- Only p_i, $i \in \{1, 2\}$ contains only grey symbols. Then $(p_i, p_{(i \mod 2)+1}) \in \mathcal{A}$.
- There are not only grey symbols in both p_1 and p_2, i.e. both contain at least a colored symbol. Then $\{p_1, p_2\} \in \mathcal{W}$.

Induction step: Suppose a clause C is the result of a resolution of $C_1 : D \vee l$ and $C_2 : E \vee \neg l$ with $l\sigma = l'\sigma$. $\mathrm{PI}(C)$ is $[\mathrm{PI}(C_1) \circ \mathrm{PI}(C)]\sigma$ or $[(l \wedge \mathrm{PI}(C_2)) \vee (\neg l \wedge \mathrm{PI}(C_1))]\sigma$.
Assumption: C_1 and C_2 are variable disjoint, i.e. variables are renamed in case C_1 and C_2 are derived from some common original clause and share variables.
By the induction hypothesis, there are appropriate arrows in both $\mathrm{PI}(C_i) \vee C_i$, $i \in \{1, 2\}$.
If the variables were present in C_1 or C_2, the arrow is still there, either in $\mathrm{PI}(C)$ (in the case of l or l'), C (in case of D and E) or in currently not shown literal (in case l and l' have the same color).
Otherwise, it was introduced by unification in $l\sigma$ or $\mathrm{PI}(C_i)\sigma$. In this case, there is some term position q in with $\rho(l).q$ a variable and $\rho(l').q$ a variable or a term containing variables (or other way around). Hence unification maps a variable to a variable or a term containing variables. The variable being unified is in $\mathrm{PI}(C_i) \vee C_i$ for some $i \in \{1, 2\}$. But by the induction hypothesis, all occurrences of each variable does already have appropriate arrows, which are still present. □

Lemma 20. *In* $\mathrm{PI}(C) \vee C$ *for* $C \in \pi$*, if there is a* Δ*-colored term* s *in a* Γ*-term* t*, then there is an arrow from* p_1 *to* p_2 *such that* $\rho(p_1) = s$ *and* $\rho(p_2) = s$ *and for some* i*,* $\rho(p_2 \mod i) = t$*.*

Note: p_1 might be in some clause, the prop interpolant or none of both.

Proof. By induction.
Base case: There are no foreign terms in the initial clause sets, so no arrows necessary.
Induction step:

Resolution. Suppose a clause C is the result of a resolution of $C_1 : D \vee l$ and $C_2 : E \vee \neg l$ with $l\sigma = l'\sigma$.

1. Suppose l is colored. This case is similar to the grey one, with the exception that the cases applying to l in PI do not apply.
2. Suppose l is grey. Then $\mathrm{PI}(C) = [(l \wedge \mathrm{PI}(C_2)) \vee (\neg l \wedge \mathrm{PI}(C_1))]\sigma$

   By the induction hypothesis, there are appropriate arrows in $\mathrm{PI}(C_1) \vee C_1$ and $\mathrm{PI}(C_2) \vee C_2$.

   We show that for all maximal Γ-terms in $\mathrm{PI}(C) \vee C$ with Δ-terms in them which were not present in $\mathrm{PI}(C_i) \vee C_i$, $i \in \{1, 2\}$, there is an arrow.

   Γ-terms and Δ-terms are not unifiable. Hence all pairs of terms (t_1, t_2) in the same positions in l and l' (if both positions exist) either point to the same symbol or (w.l.o.g.) t_1 is a variable and t_2 is a term. If there are Δ-terms in Γ-terms in the prefix, they are present in both ancestors and handled by the induction hypothesis.

   The only way a Δ-colored term may enter a Γ-colored term is in the situation where t_1 is a variable and t_2 a colored term. But then $\mathrm{mgu}(t_1, t_2)$ applied to t_1 yields t_2, i.e. "the parts of σ concerned with unifying t_1 and t_2" do not introduce new Δ-terms in Γ-terms. In other words, all such situation have been present in $\mathrm{PI}(C_i) \vee C_i$ for $i \in \{1, 2\}$ and since the arrows for l and l' are merged, they are present for $l\sigma$ in $\mathrm{PI}(C)$.

   This handles the case where terms t_1 and t_2 are unified. But unification also affects all other occurrences of variables, this means "the parts of σ not concerned with unifying t_1 and t_2". The relevant case for this lemma is when a Γ-term contains a variable, that is substituted by a term containing Δ-terms. But in this case, by Lemma 19, there is an arrow from the other occurrence of the variable to the one in the Γ-term: either double arrow in $\mathcal{A}$ if both prefixes are grey, one in $\mathcal{A}$ if one of the prefixes is grey or one in $\mathcal{W}$ if both prefixes contain a colored symbol. $\square$