

Rule-based Graph Repair using Minimally Restricted Consistency-Improving Transformations

Alexander Lauer

November 15, 2022

Abstract

Contents

1 Preliminaries	3
2 consistency increasing	4
2.1 extended alternating quantifier normal form	4
2.2 conditions up to layer	4
2.3 consistency increasing transformations and rules	8
2.4 Comparison with other consistency concepts	12
3 application condition	14
3.1 potentially minimal improving rules	16

1 Preliminaries

Definition 1.1 (subgraph). Let G_1 and G_2 be graphs. The graph G_2 is called a subgraph of G_1 if an injective morphism $f : G_2 \rightarrow G_1$ exists. We use the notation $G_2 \subseteq G_1$ if G_2 is a subgraph of G_1 and $G_2 \subset G_1$ if f is not bijective.

Definition 1.2 (uppergraph). Let G_1 and G_2 be graphs with $G_1 \subseteq G_2$. A graph C is called an upper-graph of G_1 w.r.t G_2 , if $G_1 \subset C \subseteq G_2$. The set of uppergraphs of G_1 w.r.t. G_2 is denoted by $\mathcal{U}(G_1, G_2)$. If $G_1 = G_2$, we set $\mathcal{U}(G_1, G_2) = \{G_1\}$.

Definition 1.3 (overlap). Let G and G' be graphs. A graph H is called an overlap of G and G' if morphisms $p : G \hookrightarrow H$ and $p' : G' \hookrightarrow H$ such that p and p' are jointly surjective. The set of all overlaps of G and G' is denoted by $\text{ol}(G, G')$.

Definition 1.4 (overlap at morphism). Let C, G and C' with $C \subset C'$ be graphs and $p : C \hookrightarrow G$ a morphism. A graph H is called an overlap of G and C' at p if a morphism $p' : C' \hookrightarrow H$ with $p'|_C = p$ exists. The set of all overlaps of G and C' at p is denoted by $\text{ol}_p(G, C')$.

Definition 1.5 (partial morphism). Let $f : G_1 \rightarrow G_2$ and $g : G_3 \rightarrow G_4$ be morphisms. The morphism g is called a partial morphism of f if $G_3 \subseteq G_1$, $G_4 \subseteq G_2$ and $f|_{G_3} = g$.

Definition 1.6 (nested graph condition). A graph condition over a graph C_0 is inductively defined as follows:

- *true* is a graph condition over every graph.
- $\exists(a : C_0 \hookrightarrow C_1, d)$ is a graph condition over C_0 if a is a injective graph morphism and d is a graph condition over C_1 .
- $\neg d$ is a graph condition over C_0 if d is a graph condition over C_0 .
- $d_1 \wedge d_2$ and $d_1 \vee d_2$ are graph conditions over C_0 if d_1 and d_2 are graph conditions over C_0 .

Conditions over the empty graph \emptyset are called constraints. Every injective morphism $p : C_0 \hookrightarrow G$ satisfies *true*. An injective morphism p satisfies $\exists(a : C_0 \hookrightarrow C_1, d)$ if there exists an injective morphism $q : C_1 \hookrightarrow G$ such that $q \circ a = p$ and q satisfies d . An injective morphism satisfies $\neg d$ if it does not satisfy d , it satisfies $d_1 \wedge d_2$ if it satisfies d_1 and d_2 and it satisfies $d_1 \wedge d_2$ if it satisfies d_1 or d_2 . A graph G satisfies a constraint c , $G \models c$, if $p : \emptyset \hookrightarrow G$ satisfies c . We use the abbreviation $\forall(a : C_0 \hookrightarrow C_1, d) := \neg\exists(a : C_0 \hookrightarrow C_1, \neg d)$.

The nesting level nl of a condition is defined as $\text{nl}(\text{true} = 0$ and $\text{nl}(\exists(a : P \rightarrow Q, d)) := \text{nl}(d) + 1$.

Definition 1.7 (alternating quantifier normal form (ANF))[1]. A graph condition c is in alternating normal form (ANF) if it is of the form

$$c = Q(a_1 : C_0 \hookrightarrow C_1, \overline{Q}(a_2 : C_1 \hookrightarrow C_2, Q(a_3 : C_2 \hookrightarrow C_3, \overline{Q}(a_4 : C_3 \hookrightarrow C_4, \dots))))$$

with $Q \in \{\exists, \forall\}$ and $\overline{Q} = \exists$ if $Q = \forall$, $\overline{Q} = \forall$ if $Q = \exists$.

2 consistency increasing

In this section, we introduce the notion of *consistency increasing* transformations and rules, which allows to increase the consistency of a constraint layer by layer.

2.1 extended alternating quantifier normal form

To prevent the need of case discrimination, a new normal form for conditions, called *extended alternating quantifier normal form* (EANF), will be introduced. The sets of conditions in ANF and EANF do intersect and we show that both sets are expressively equivalent.

Definition 2.1 (extended alternating quantifier normal form). A conditions c is in extended alternating quantifier normal form (EANF) if it is of the form $\forall(a_0 : C_0 \hookrightarrow C_1, d)$ and d is a condition in ANF.

Note that, given a condition c in EANF, every subcondition of c at layer $1 \leq k \leq \text{nl}(c)$ is universally bound, if k is an odd number and existentially bound, if k is an even number.

Lemma 2.2. Any condition in ANF can be transformed into an equivalent condition in EANF and vice versa.

Proof. “ \Rightarrow ”: Let a graph G and a constraint c in ANF be given. If c is universally, c is already in EANF.

If $c = \exists(a_0 : C_0 \hookrightarrow C_1, d)$, we show that c is equivalent to $c' := \forall(\text{id}_{C_0} : C_0 \hookrightarrow C_0, c)$.

1. Let $p : C_0 \hookrightarrow G$ be a morphism, such that $q \models c$. Therefore a morphism $q : C_0 \rightarrow G$ with $q \models e$ and $p = q \circ a_1$ exists. Then, $p \models d$, since p is the only morphism from C_0 to G with $p = p \circ \text{id}_{C_0}$ and $p \models c$.
2. Let $p : C_0 \hookrightarrow G$ be a morphism with $p \models c'$, therefore all morphisms $q : C_0 \hookrightarrow G$ with $p = q \circ \text{id}_{C_0}$ satisfy c . Since $p = p \circ \text{id}_{C_0}$, $p \models c$ follows immediately.

“ \Leftarrow ”: Let a graph G and a constraint c in EANF be given. If $c = \forall(a_0 : C_0 \hookrightarrow C_1, d)$ with $C_0 \neq C_1$, c is already in ANF.

Otherwise, the equivalence of c with d can be shown analogously to the first case. □

2.2 conditions up to layer

Definition 2.3 (Layer of a subcondition). Let c be a condition and d a subcondition of c . The layer of d is defined as $\text{lay}(d) := \text{nl}(c) - \text{nl}(d) - 1$.

Definition 2.4 (Subcondition at layer). Let c be a condition. The subcondition at layer k , denoted by $\text{sco}_c(k)$, is the subcondition d of c with $\text{lay}(d) = k$.

We define a notion of partial consistency, called *satisfaction at layer*, which will be used for the definition of consistency increasing. First, two operators are introduced to modify given constraints on a certain layer.

Definition 2.5 (substitution at layer). Let $c = Q(a : C_0 \rightarrow C_1, d)$ be a condition in ANF, such that the subcondition of c with layer $0 \leq k \leq \text{nl}(c)$ is a condition over C_k . Let e be a condition over C_k . The substitution in c at layer k with e , $\text{sub}(k, c, e)$, is recursively defined as:

1. If $k \leq 1$:

$$\text{sub}(0, c, e) := e$$

2. If $k > 1$:

$$\text{sub}(k, c, e) := Q(a : C_0 \rightarrow C_1, \text{sub}(k - 1, d, e))$$

Example 2.1. Let the conditions $c := \forall(a_0 : C_0 \rightarrow C_1, \exists(a_1 : C_1 \rightarrow C_2, \text{true}))$ and $d = \exists(a'_1 : C_1 \rightarrow C_3, e)$ be given. Then,

$$\text{sub}(2, c, d) = \forall(a_0 : C_0 \rightarrow C_1, \exists(a'_1 : C_1 \rightarrow C_3, e)).$$

Through this, we define *condition up to layer*. Intuitively, a condition is cut off at a certain layer, by replacing the subcondition at this layer by `true` or `false`, depending on the quantifier, the replaced subcondition is bound by.

Definition 2.6 (Condition up to layer). Let c be a condition in EANF and d be the subcondition of c at layer $0 \leq k \leq \text{nl}(c)$. The condition up to layer k of c , $\text{cond}(k, c)$, is defined as

$$\text{cond}(k, c) := \begin{cases} \text{sub}(k + 1, c, \text{true}) & \text{if } d \text{ is existentially bound} \\ \text{sub}(k + 1, c, \text{false}) & \text{if } d \text{ is universally bound.} \end{cases}$$

Example 2.2. Let the condition $c = \forall(a_0 : C_0 \rightarrow C_1, \exists(a_1 : C_1 \rightarrow C_2, d))$ be given. Then,

$$\text{cond}(2, c) = \forall(a_0 : C_0 \rightarrow C_1, \text{true}).$$

Now, we are ready to define satisfaction at layer, which is a key ingredient to define consistency increasing.

Definition 2.7 (Satisfaction at layer). Let G be a graph and c be a condition over C_0 . A morphism $p : C_0 \rightarrow G$ satisfies c at layer k , $p \models_k c$, if

$$p \models \text{cond}(k, c).$$

A graph G satisfies a constraint c at layer k , $G \models_k c$, if $q : \emptyset \rightarrow G$ satisfies $\text{cond}(k, c)$. The biggest k with $G \models_k c$ such that no $j > k$ with $G \models_j c$ exists is denoted by c_{\max}^G .

Example 2.3.

The following lemmas arise as a direct consequence of the definition of satisfaction at layer. If a graph satisfies a constraint up to a certain layer, let c be this condition up to this layer, that ends with $\forall(a : C \hookrightarrow C', \text{false})$, the graph satisfies all constraints starting with c .

Lemma 2.8. *Let G be a graph $p : C_0 \hookrightarrow G$ a morphism and c a condition over C_0 in EANF with $p \models_k c$. If the subcondition $d = Q(a_k : C_{k-1} \hookrightarrow C_k, e)$ of c at layer k is universally bound, then for any condition f over C_k it holds that*

$$p \models \text{sub}(k+1, c, f).$$

Proof. Let k be the smallest number with $\text{sco}_c(k) = \forall(a_k : C_{k-1} \hookrightarrow C_k, \text{false})$ being universally bound and $p \models_k c$. Let $q : G_{k-1} \hookrightarrow G$ be a morphism such that $q \models \forall(a_k : C_{k-1} \hookrightarrow C_k, \text{false})$. This must exist, since $p \models_k c$ and k is the smallest even number such that $p \models_k c$.

Therefore, there does not exist a morphism $q' : C_k \hookrightarrow G$ with $q = q' \circ a_k$. Hence, for every condition f over C_k a morphism $q' : C_k \hookrightarrow G$ with $q \not\models f$ and $q = q' \circ a_k$ cannot exist. It follows immediately that $q \models \forall(a_k : C_{k-1} \hookrightarrow C_k, f)$ and with that $p \models \text{sub}(k+1, c, f)$. For every $j > k$ even, such that $p \models_j c$, and every condition d over C_j , with the first part, it holds that $p \models \text{sub}(k+1, c, f)$ with $f = \text{sub}(j-k+1, \text{sco}_c(k+1), d)$ and since $\text{sub}(k+1, c, f) = \text{sub}(j+1, c, d)$ it follows that $p \models \text{sub}(j+1, c, d)$. \square

As a direct consequence of the previous lemma, a graph satisfying a condition up to layer ending with $\forall(a : C \hookrightarrow C', \text{false})$ also satisfies the whole constraint.

Lemma 2.9. *Let G be a graph, $p : C_0 \hookrightarrow G$ a morphism and c a condition over C_0 in EANF with $p \models_k c$. If $\text{sco}_c(k)$ is universally bound,*

$$p \models_k c \implies p \models c.$$

Proof. Follows immediately by using lemma 2.8 and setting f to $\text{sco}_c(k+1)$. \square

Lemma 2.10. *Let a graph G , a morphism $p : C_0 \hookrightarrow G$ and a constraint c in EANF be given. Then, $p \models c$ for all $j < c_{\max}^G$ such that $\text{sco}_c(j)$ is existentially bound.*

Proof. 1. The subcondition of c at layer c_{\max}^G is existentially bound: If a $j < c_{\max}^G$ with $p \models_j c$ exists such that $\text{sco}_c(j)$ is universally bound, let j_1 be the smallest of these. With lemma 2.8 follows that $p \models_{j_2} c$ for all $j_1 < j_2$. Let $\ell < j_1$, such that $\text{sco}_c(\ell)$ is existentially bound and let $d = \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, e)$ be the condition up to layer $j_1 - \ell$ of $\text{sco}_c(\ell)$. Since $\ell < j_1$, a morphism $q : C_\ell \hookrightarrow G$ with $q \models d$ must exists and therefore a morphism $q' : C_{\ell+1} \hookrightarrow G$ with $q = q' \circ a_k$ must exists. It follows that $q \models \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, \text{true})$ and with that $p \models_\ell c$.

2. The subcondition of c at layer c_{\max}^G is universally bound: With lemma 2.8 follows that $p \models_{k+1} c$. Since c is in EANF, 1. can be applied to $k+1$. \square

Through satisfaction at layer an increase of consistency can be detected in the following way. Let $G \implies H$ be a transformation. If $c_{\max}^G < c_{\max}^H$, the transformation can be considered as consistency increasing, since H satisfies more layers of the constraint than G . The notion of consistency increasing should also be able to detect the smallest transformations that lead to a increase of consistency, namely the inserting of a single edge or node of an existentially bound graph. To remedy this issue, we introduce *partial conditions*. Intuitively, given a constraint c in EANF ending with $\text{sco}_c(k) = \exists(a_k : C_k \hookrightarrow C_{k+1}, d)$, the condition $\text{sco}_c(k)$ is replaced by $\exists(a_k^p : C_k \hookrightarrow C', \text{true})$ with $C' \in \mathcal{U}(C_k, C_{k+1})$.

The construction of partial conditions is designed to only replace graphs in existentially bound layers, since the replacement in an universally bound layer would lead to a more restrictive constraint than the original condition up to layer. That means, given the condition $c = \forall(a_0 : C_0 \hookrightarrow C_1, \text{false})$ and let $C' \in \mathcal{U}(C_0, C_1)$. If the condition $c' = \forall(a_0^p : C_0 \hookrightarrow C_1, \text{false})$ is satisfied this implies that c is also satisfied but the backwards implication does not hold.

Definition 2.11 (partial condition). Let c be a condition in EANF.

The partial condition of c at layer $k < \text{nl}(c)$ with with

$$C' \in \begin{cases} \mathcal{U}(C_k, C_{k+1}) & \text{if } k \text{ is even} \\ \mathcal{U}(C_{k+1}, C_{k+2}) & \text{if } k \text{ is odd,} \end{cases}$$

$\text{part}(k, c, C')$, is defined as:

1. If k is odd, let $\text{sco}_c(k+1) = \exists(a : C_{k+1} \hookrightarrow C_{k+2}, f)$:

$$\text{part}(k, c, C') := \text{sub}(k+1, c, \exists(a^p : C_{k+1} \hookrightarrow C', \text{true}))$$

2. If k is even, let $\text{sco}_c(k+1) = \exists(a : C_k \hookrightarrow C_{k+1}, f)$:

$$\text{part}(k, c, C') := \text{sub}(k, c, \exists(a^p : C_{k+1} \hookrightarrow C', \text{true}))$$

Example 2.4.

If a graph G does not satisfy the condition up to layer $k := c_{\max}^G + 2$ of a given constraint c , with k being even, there does exists a graph $C' \in \mathcal{U}(C_{k-1}, C_k)$, such that G satisfies $\text{part}(k, c, C')$, note that G always satisfies $\text{part}(k, c, C_{k-1})$. For the biggest of these graphs, we define the notions of *biggest partially satisfying conditions and graphs*.

Definition 2.12 (biggest partially satisfying condition). Let G be a graph, c a condition in EANF over C_0 and $p : C_0 \hookrightarrow G$ a morphism with $p \models_k c$ and $p \not\models c$.

A partial condition $c = \text{part}(k+2, c, C')$ with $p \models c$ is a biggest partially satisfying condition w.r.t c of p if there does not exist a graph C'' , such that C' is a subgraph of C'' , with $p \models \text{part}(c_{\max}, c, C'')$. The graph C' is then called a biggest partially satisfying graph w.r.t c of p .

Given a constraint c , the set of biggest partially satisfying conditions w.r.t c of $p : \emptyset \inf G$ is denoted by \mathcal{P}_c^G .

The set of all biggest partially satisfying graphs w.r.t c of p is denoted by \mathcal{G}_c^G .

2.3 consistency increasing transformations and rules

With the results above, we are now ready to define the notions of *consistency increasing* and *direct consistency increasing*, with direct increasing being a more restrictive version of increasing, yielding the advantage that second-order logic formulas can be used in order to determine whether a transformation is (direct) consistency increasing or not.

These notions are designed to only detect transformations that increase the consistency of the first two unsatisfied layer of a constraint c . That means, given a graph G and a constraint c , let $k = c_{\max}^G$ and $\forall(a_{k+1} : C_k \hookrightarrow C_{k+1}, \exists(a_{k+2} : C_{k+1} \hookrightarrow C_{k+2}, d)) =: \text{sco}_c(k+1)$. A transformation $t : G \implies H$ is considered as (direct) consistency increasing if $c_{\max}^G \leq c_{\max}^H$, i.e. the satisfaction up to layer k is not decreased, and more increasing insertions or deletions have been performed than decreasing ones. An increasing deletion is the deletion of an occurrence of C_{k+1} that does not satisfy $\exists(a_{k+2} : C_{k+1} \hookrightarrow C_{k+2}, \text{true})$, an increasing insertion is the insertion of elements of C_{k+2} , such that for at least one occurrence p of C_{k+1} it holds that $p \not\models \exists(a_{k+2}^p : C_{k+1} \hookrightarrow C', \text{true})$ and $\text{tr}_t \circ p \models \exists(a_{k+2}^p : C_{k+1} \hookrightarrow C', \text{true})$ for an $C' \in \mathcal{U}(C_{k+1}, C_{k+2})$. A decreasing insertion is the creation of an occurrence of C_{k+1} not satisfying $\exists(a_{k+2} : C_{k+1} \hookrightarrow C_{k+2}, \text{true})$ and a decreasing deletion is the deletion of elements of C_{k+2} such that for an occurrence p of C_{k+1} with $p \models \exists(a_{k+2}^p : C_{k+1} \hookrightarrow C', \text{true})$ it holds that $\text{tr}_t \circ p \not\models \exists(a_{k+2}^p : C_{k+1} \hookrightarrow C', \text{true})$.

To evaluate this, we define the *number of violations*. Intuitively, for all occurrences p of C_{k+1} the number of graphs $C' \in \mathcal{U}(C_{k+1}, C_{k+2})$ with $p \not\models \exists C'$ is add up and it can be determined whether more increasing or decreasing actions have been performed by a transformation.

Note, that the number of violations is defined for each layer of the constraint, but only for the first unsatisfied layer the sum is calculated as described above. For all layer k with $k \leq c_{\max}^G$ it is set to 0 and for all layer k with $k > c_{\max}^G + 1$ it is set to ∞ . Through this, a transformation $tG \implies H$ that increases the partial consistency can easily detected since the number of violations in H at layer $c_{\max}^G + 1$ will be set to 0.

Definition 2.13 (number of violations). Let G be a graph and c a constraint in EANF. The number of violations $\text{nvc}(j, G)$ at layer j in G is defined as:

1. If $j < c_{\max}^G + 1$:

$$\text{nvc}(j, G) := 0$$

2. If $j = c_{\max}^G + 1$, let $d = \forall(a_k : C_j \hookrightarrow C_{j+1}, e)$ be the subcondition of c at layer $j + 1$.

$$\text{nvc}(j, G) := \begin{cases} \sum_{C' \in \mathcal{U}(C_j, C_{j+1})} |\{q \mid q : C_{j+1} \hookrightarrow G \wedge q \not\models \text{part}(1, e, C')\}| & \text{if } e \neq \text{false} \\ |\{q \mid q : C_{j+1} \hookrightarrow G\}| & \text{if } e = \text{false} \end{cases}$$

3. If $j > c_{\max}^G + 1$:

$$\text{nvc}(j, G) := \infty$$

Via the number of violations, we now define *consistency increasing* transformations and rules, by checking whether the number of violations has decreased for any layer of the constraint.

Definition 2.14 (consistency increasing). Let a graph G , a rule ρ and a constraint c in EANF be given.

A transformation $G \Rightarrow_{\rho,m} H$ is called consistency increasing w.r.t c , if

$$\text{nvc}(k, H) < \text{nvc}(k, G)$$

for any $0 \leq k \leq \text{nl}(c)$. A rule r is called consistency increasing w.r.t c , if all of its applications are.

Note that if $G \models c$ there does not exist a consistency increasing transformation w.r.t c , since $\text{nvc}(j, G) = 0$ for all $1 \leq j \leq \text{nl}(c)$. Also, no plain rule ρ is consistency increasing w.r.t c , since a graph G satisfying c , such that a transformation $t : G \Rightarrow_{\rho,m} H$ exists can always be constructed. Therefore, each consistency increasing rule has to be equipped with at least one application condition.

As mentioned above, a transformation should be detected as consistency increasing if it increases the partial consistency, which is shown by the following theorem.

Theorem 2.1. Let a graph G , a rule ρ and a constraint c in EANF be given. A transformation $t : G \Rightarrow_{\rho,m} H$ is consistency increasing if $c_{\max}^G < c_{\max}^H$.

Proof. No $\ell > c_{\max}^G$ with $G \models_{\ell} c$ exists. Hence, $\text{nvc}(c_{\max}^G + 1, G) > 0$ and $\text{nvc}(c_{\max}^G + 1, H) \neq \infty$. Since $c_{\max}^H > c_{\max}^G$, $\text{nvc}(k + 1, H) = 0$ and it follows immediately that t is consistency increasing. \square

Since no consistency increasing transformation originating in consistent graphs exist, there do not exist infinite long sequences of consistency increasing transformations. Additionally, if a set of rules \mathcal{R} is given, and a sequence of consistency increasing transformations with rules of \mathcal{R} ends with a graph G , then either G satisfies the constraint or there do not exist any consistency increasing transformations $G \Rightarrow_{\rho,m} H$ with $\rho \in \mathcal{R}$.

Theorem 2.2. Let a constraint c in EANF and a set of rules \mathcal{R} be given. Every sequence of minimal consistency improving transformation w.r.t c with rules in \mathcal{R} is finite.

Proof. Let G_0 be a graph and

$$G_0 \Rightarrow_{\rho_1} G_1 \Rightarrow_{\rho_2} G_2 \Rightarrow_{\rho_3} \dots$$

be a sequence of minimal consistency improving transformations w.r.t c with $\rho_i \in \mathcal{R}$. We assume that $c_{\max}^{G_0} < \text{nl}(c)$, otherwise $\text{nvc}(j, G_0) = 0$ for all $j \in \{0, \dots, \text{nl}(c)\}$ and no transformation $G_0 \Rightarrow H$ is consistency increasing.

We show that after at most $j := \text{nvc}(c_{\max}^{G_0} + 1, G_0)$ transformations $G_j \models_{c_{\max}^{G_0} + 2} c$ holds. Note that j has to be finite, since G_0 contains only a finite number of occurrences of C_{j+1} . After each transformation it holds that $\text{nvc}(c_{\max}^{G_{i+1}} + 1, G_{i+1}) \leq \text{nvc}(c_{\max}^{G_i} + 1, G_i) -$

1. Therefore, after j transformations, $\text{nvc}(c_{\max}^{G_0} + 1, G_j) \leq \text{nvc}(c_{\max}^{G_0} + 1, G_0) - j = 0$ holds and with that $G_j \models_{c_{\max}^{G_0} + 2} c$. By iteratively applying this, it follows that after a finite number of transformations a graph G_k with $G_k \models c$ has to exist. Since no consistency increasing transformation $G_k \implies G_{k+1}$ exists, the sequence has to be finite. Also, for some $G_{k'}$ there may not exist a consistency increasing transformation $G_{k'} \implies_{\rho} H$ with $\rho \in \mathcal{R}$ and therefore the sequence is also finite, but does not end with a consistent graph. \square

Now, we define *direct consistency increasing* transformations, a stricter version of consistency increasing.

Definition 2.15 (direct minimal consistency improving). Let G be a graph, ρ a plain rule and c a constraint in EANF. Let $d = \forall(a_k : C_{k-1} \hookrightarrow C_k, e)$ be the condition at layer $k = c_{\max}^G + 1 \leq \text{nl}(c)$ of c . A transformation $t : G \implies_{\rho, m} H$ is called direct minimal consistency improving if the following equations hold.

Every occurrence of C_k in G that satisfies $\text{part}(1, e, C')$ for any $C' \in \mathcal{U}(C_k, C_{k+1})$ still satisfies $\text{part}(1, e, C')$ in H .

$$\begin{aligned} \forall p : C_k \hookrightarrow G \left(\bigwedge_{C' \in \mathcal{U}(C_k, C_{k+1})} (p \models \text{part}(1, e, C') \wedge \text{tr}_t \circ p \text{ is total}) \right. \\ \left. \implies \text{tr}_t \circ p \models \text{part}(1, e, C') \right) \end{aligned} \quad (2.1)$$

Every new inserted occurrence of C_k by t satisfies $\text{part}(1, e, C_{k+1})$.

$$\forall p' : C_k \hookrightarrow H (\neg \exists p : C_k \hookrightarrow G (p' = \text{tr}_t \circ p) \implies p' \models \text{part}(1, e, C_{k+1})) \quad (2.2)$$

At least one occurrence of C_k in G that does not satisfy $\text{part}(1, e, C')$, for any $C' \in \mathcal{U}(C_k, C_{k+1})$, either has been destroyed by t or satisfies $\text{part}(1, e, C')$ in H .

$$\begin{aligned} \exists p : C_k \hookrightarrow G \left(\bigvee_{C' \in \mathcal{U}(C_k, C_{k+1})} (p \not\models \text{part}(1, e, C') \wedge (\text{tr}_t \circ p \text{ is not total} \right. \\ \left. \vee (\text{tr}_t \circ p \text{ is total} \wedge \text{tr}_t \circ p \models \text{part}(1, e, C')))) \right) \end{aligned} \quad (2.3)$$

No occurrence of a universally bound graph C_j with $j < k$ gets inserted.

$$\bigwedge_{\substack{i < k \\ i \text{ even}}} \forall p : C_i \hookrightarrow H (\exists p' : C_i \hookrightarrow G (p' = \text{tr}_t \circ p)) \quad (2.4)$$

No occurrence of an existentially bound graph C_j with $j < k$ gets deleted.

$$\bigwedge_{\substack{i < k \\ i \text{ odd}}} \forall p : C_i \hookrightarrow G (\text{tr}_t \circ p \text{ is total}) \quad (2.5)$$

Lemma 2.16. Let a transformation $t : G \implies H$ and a constraint c in EANF be given, such that (2.4) and (2.5) of definition 2.15 are satisfied. Then,

$$H \models_{c_{\max}^G} c.$$

Proof. Assume that $G \not\models_{c_{\max}^G} c$. Then, either a new occurrence of an universally bound graph of c has been inserted or an occurrence of an existentially bound graph of c has been destroyed. Therefore, the following holds:

$$\exists p : C_i \hookrightarrow G (\neg \exists p' : C_i \hookrightarrow (p' = \text{tr}_t \circ p) \vee \exists p : C_j \hookrightarrow G (\text{tr}_t \circ p \text{ is not total}))$$

with $i, j < k$, i being odd and j being even. It follows immediately that either (2.4) or (2.5) is not satisfied. This is a contradiction. \square

Lemma 2.17. Let a graph G , a constraint c in EANF and a direct minimal improving transformation $t : G \Rightarrow_{r,m} H$ w.r.t. c be given. Then, t is also a minimal improving transformation.

Proof. Let G be a graph with $k = c_{\max}^G + 1$ and $G \models \text{part}(k, c, C)$ with $\text{part}(k, c, C) \in \mathcal{P}_c^G$. Let d be the subcondition of c at layer k . With lemma 2.16 follows that $c_{\max}^H \geq c_{\max}^G$ and with that $\text{nvc}(k, H) \neq \infty$.

1. We show that equations (2.1) and (2.2) imply that $\text{nvc}(k, H) \leq \text{nvc}(k, G)$. Assume that $\text{nvc}(k, H) > \text{nvc}(k, G)$. Therefore, a morphism $p : C_k \hookrightarrow H$ with $p \not\models \text{part}(1, d, C')$ for any $C' \in \mathcal{U}(C, C_{k+1})$ exists, such that either 1a or 1b is satisfied.
 - (a) There does exist a morphism $q' : C_k \hookrightarrow G$ with $q' \models \text{part}(1, d, C')$ and $p = \text{tr}_t \circ q'$.
 - (b) There does not exist a morphism $q : C_k \hookrightarrow G$, such that $p = \text{tr}_t \circ q$.

This is a contradiction, if 1a is satisfied, q' does not satisfy equation (2.1) and if 1b is satisfied q does not satisfy equation (2.2). It also follows that

$$|\{q \mid q : C_k \hookrightarrow G \wedge q \not\models \text{part}(1, e, C')\}| \leq |\{q \mid q : C_k \hookrightarrow H \wedge q \not\models \text{part}(1, e, C')\}|$$

for all $C' \in \mathcal{U}(C_k, C_{k+1})$.

2. Since (2.3) is satisfied, a morphism $p : C_k \hookrightarrow G$ with $p \not\models \text{part}(1, d, C')$, such that either $\text{tr} \circ p$ is total and $\text{tr}_t \circ p \models \text{part}(1, d, C')$ or $\text{tr} \circ p$ is not total exists, for any $C' \in \mathcal{U}(C_k, C_{k+1})$. In both cases the following holds

$$\begin{aligned} p \in \{q \mid q : C_k \hookrightarrow G \wedge q \not\models \text{part}(1, e, C')\} \wedge \\ \text{tr} \circ p \notin \{q \mid q : C_k \hookrightarrow H \wedge q \not\models \text{part}(1, e, C')\}. \end{aligned}$$

With that and 1. it follows that

$$|\{q \mid q : C_k \hookrightarrow G \wedge q \not\models \text{part}(1, e, C')\}| < |\{q \mid q : C_k \hookrightarrow H \wedge q \not\models \text{part}(1, e, C')\}|.$$

In total $\text{nvc}(k, H) < \text{nvc}(k, G)$ follows and t is a minimal improving transformation. \square

2.4 Comparison with other consistency concepts

Lemma 2.18. Let a constraint c in ANF, graphs G and H with $G \not\models c$, and a transformation $t : G \implies H$ be given. Then,

$$t \text{ is } c\text{-guaranteeing} \implies t \text{ is minimal consistency improving w.r.t } c$$

$$\begin{aligned} & , \\ t \text{ is } c\text{-guaranteeing} & \not\implies t \text{ is direct minimal consistency improving w.r.t } c \end{aligned}$$

and

$$t \text{ is minimal consistency improving w.r.t } c \not\implies t \text{ is } c\text{-guaranteeing}$$

Proof. Let c' be the equivalent constraint in EANF.

1. Let t be c -guaranteeing transformation, then $H \models c$. Since $G \not\models c$, $G \not\models c'$ follows and $\text{nvc}(c_{\max}^G + 1, G) > 0$ and $\text{nvc}(c_{\max}^G + 1, H) = 0$. Therefore t is minimal consistency improving.
2. Let a constraint $c = \exists(a_0 : \emptyset \hookrightarrow C_0, \forall(a_1 : C_0 \hookrightarrow C_1, \exists(a_2 : C_1 \hookrightarrow C_2, \text{true})))$ and a graph $G = C' \dot{\cup} C'$ with $C' = C_2 \setminus \{e\}$ for an $e \in E_{C_2 \setminus C_1}$ and the occurrences $p_1 : C_1 \hookrightarrow G$ and $p_2 : C_1 \hookrightarrow G$ be given. Let $t : G \implies H$ be a transformation with $H = C_2 \dot{\cup} C''$ and $C'' = C' \setminus \{e'\}$ for an $e' \in E_{C_2 \setminus C_1}$, such that $\text{tr}_t \circ p_1$ and $\text{tr}_t \circ p_2$ are total. Then, t is c -guaranteeing and not direct minimal consistency improving, since $p_i \models \exists(a'_2 : C_1 \hookrightarrow C', \text{true})$ for $i = 1, 2$ and either $\text{tr}_t \circ p_1 \not\models \exists(a'_2 : C_1 \hookrightarrow C', \text{true})$ or $\text{tr}_t \circ p_2 \not\models \exists(a'_2 : C_1 \hookrightarrow C', \text{true})$.
3. Let t be a minimal consistency improving transformation w.r.t c' , such that $\text{nvc}(c_{\max}^G + 1, H) > \text{nvc}(c_{\max}^G + 1, H) > 1$ and $H \not\models c_{\max+2}^G c'$. Then, $H \not\models c'$ and t is not a guaranteeing transformation.

□

Corollary 2.19. Let c be an existentially bound constraint in ANF. Then, a transformation t is c -guaranteeing if and only if it is consistency improving. With lemma 2.18 follows:

$$t \text{ is consistency guaranteeing w.r.t } c \implies t \text{ is minimal consistency improving w.r.t } c$$

and

$$t \text{ is minimal consistency improving w.r.t } c \not\implies t \text{ is consistency improving w.r.t } c$$

Lemma 2.20. let a universally bound constraint c with $\text{nl}(c) = 1$ in ANF, graphs G and H with $G \not\models c$, and a transformation $t : G \implies H$ be given. Then,

$$t \text{ is consistency improving} \iff t \text{ is minimal consistency improving}$$

Proof. Let $c = \forall(a : \emptyset \hookrightarrow C, \text{false})$. Then, the equivalent constraint in EANF is $c' = \forall(a : \emptyset \hookrightarrow C, \exists(\text{id} : C \hookrightarrow C, \text{false}))$. Since $\mathcal{U}(C, C) = \{C\}$, $\text{nvc}(1, G)$ is the number of occurrences of C in G . This is exactly the definition of the number of violations for consistency improving transformations and

$$t \text{ is minimal consistency improving} \iff t \text{ is consistency improving}$$

follows immediately. \square

Lemma 2.21. *Let a universally bound constraint c with $\text{nl}(c) \geq 2$ in ANF, graphs G and H with $G \not\models c$ and a transformation $t : G \Rightarrow H$ be given. Then,*

$$t \text{ is direct consistency improving} \Rightarrow t \text{ is minimal consistency improving}$$

and

$$t \text{ is direct minimal consistency improving} \Rightarrow t \text{ is consistency sustaining}$$

Proof. 1. Let $c = \forall(a_0 : \emptyset \hookrightarrow C_0, \exists(a_1 : C_0 \hookrightarrow C_1, \text{true}))$ be a constraint. Let $V_{C_0} = V_{C_1}$ and $|E_{C_1}| - |E_{C_0}| = 2$. Let $G = C' \dot{\cup} C'$ with $C' = C_1 \setminus \{e\}$ with $e \in E_{C_1} \setminus E_{C_0}$ and the occurrences $p_1 : C_0 \hookrightarrow G$ and $p_2 : C_0 \hookrightarrow G$ be given. It follows that $\text{nvc}(1, G) = 2$. Let $t : G \Rightarrow H$ be a transformation, such that $H = C_0$. Then, t is a direct consistency improving transformation, since H contains only one occurrence of C_0 not satisfying $\exists(a_1 : C_0 \hookrightarrow C_1, \text{true})$. But, t is not minimal consistency improving since $\text{nvc}(1, H) = 3$.

2. Let $c := \forall(a_0 : \emptyset \hookrightarrow C_0, \exists(a_1 : C_0 \hookrightarrow C_1, \forall(a_2 : C_1 \hookrightarrow C_2, \exists(a_3 : C_2 \hookrightarrow C_3, \text{true}))))$ be a constraint.

Let a graph $G = C_0$ with the morphism $q : C_0 \hookrightarrow G$ and a transformation $t : G \Rightarrow H$ with $H := C_2 \dot{\cup} C_2$ be given, such that $\text{tr}_t \circ q$ is total. Then t is a direct minimal consistency improving transformation but not a consistency sustaining one, since H contains more occurrences of C_0 not satisfying $\exists(a_1 : C_0 \hookrightarrow C_1, \forall(a_2 : C_1 \hookrightarrow C_2, \exists(a_3 : C_2 \hookrightarrow C_3, \text{true})))$ than G . \square

Lemma 2.22. *Let a universally bound constraint c with $\text{nl}(c) \geq 2$ in ANF, graphs G and H with $G \not\models c$ and a transformation $t : G \Rightarrow H$ be given. If t satisfies (2.1) and (2.2), $H \models_{c_{\max}^G} c$ and*

$$\exists p : C_0 \hookrightarrow G(p \not\models \text{part}(c_{\max}^G + 2, c, \text{true}) \wedge \text{tr}_t \circ p \models c).$$

Then,

$$t \text{ is consistency improving} \Rightarrow t \text{ is minimal consistency improving.}$$

Proof. \square

3 application condition

Definition 3.1 (extended overlaps). Let G and $C_0 \subseteq C_1$ with the inclusion $i : C_0 \hookrightarrow C_1$ be graphs. Let C be an overlap of C_0 and G with the inclusion $q : C_0 \hookrightarrow C$. The set of extended overlaps of C with i , $\text{eol}(C, i)$, is the set of all overlaps C' of G and C_1 , such that $C \subseteq C'$ and $q \models \exists(i : C_0 \hookrightarrow C_1, \text{true})$.

Definition 3.2 (overlap shift). Let $\rho = L \hookleftarrow K \hookrightarrow R$ be a plain rule, C a graph and C' an overlap of C and L with morphisms $p : L \hookrightarrow C'$, $k : K \hookrightarrow C'$, $c : C \hookrightarrow C'$ and the partial morphism $q : R \hookrightarrow C'$. We define

$$\begin{aligned} D := \{e \in C' \mid & (\exists e' \in L : p(e') = e \\ & \vee \exists e' \in R : q(e') = e) \\ & \wedge \exists e' \in C : c(e') = e\} \end{aligned} \quad (3.1)$$

Let $r = L \hookleftarrow K' \hookrightarrow R$ be the rule with

$$K' := K \cup D$$

The graph H derived by the transformation $G \Rightarrow_{r,p} H$ is called the overlap shifted graph of C' with C and ρ . The overlap shifted graph of an graph C is denoted by $\text{ols}_\rho(C, C')$.

Definition 3.3 (application condition). Let $\rho = L \hookleftarrow K \hookrightarrow R$ be a plain rule and c a constraint in EANF. Let $d = \forall(a_k : C_{k-1} \hookrightarrow C_k, \exists(a_{k+1} : C_k \hookrightarrow C_{k+1}, e))$ be the subcondition of c at layer k with k being and even number.

The application condition ap_k of c at layer k with $C' \in \mathcal{U}(C_k, C_{k+1})$ and $i : C_k \hookrightarrow C'$ is defined as:

$$\text{ap}(k, C') := \left(\bigvee_{P \in \text{ol}(L, C_k)} \text{nex}(P, C') \wedge \text{rep}(P, C') \right) \wedge \text{nin} \wedge \text{rem} \wedge \text{nwo} \quad (3.2)$$

with

1.

$$\text{nex}(P, C') := \begin{cases} \exists(i_L : L \hookrightarrow P, \text{true}) & , \text{ if } C_k = C_{k+1} \\ \bigwedge_{Q \in \text{eol}(P, i)} \exists(i_L : L \hookrightarrow P, \neg \exists(i_P : P \hookrightarrow Q, \text{true})) & , \text{ otherwise} \end{cases}$$

2. Let \mathbf{P} be the set of all overlaps of R and C' , such that $i_R(R \setminus K) \cap i_{C'}(C') \neq \emptyset$:

(a) If $C_k = C_{k+1}$:

$$\text{rep}(P, C') := \begin{cases} \text{true} & , \text{ if } i_L(L \setminus K) \cap i_{C_k}(C_k) \neq \emptyset \\ \text{false} & , \text{ otherwise} \end{cases}$$

(b) Otherwise:

$$\text{rep}(P, C') := \begin{cases} \text{true} & , \text{ if } i_L(L \setminus K) \cap i_{C_k}(C_k) \neq \emptyset \\ \bigvee_{P' \in \mathbf{P}} \text{Left}(\exists(i_R : R \hookrightarrow P', \text{true}), \rho) & , \text{ otherwise} \end{cases}$$

3. Let E be the set of all graphs C_j of c with $j \leq k$ and j being odd and let \mathbf{P}_{C_j} be the set all overlaps of L and C_j with $i_L(L \setminus K) \cap i_{C_j}(C_j) \neq \emptyset$.

$$\text{rem} := \bigwedge_{C \in E} \bigwedge_{C' \in \mathbf{P}_{C_j}} \neg \exists(i_L : L \hookrightarrow C', \text{true})$$

4. Let U be the set of all graphs C_j of c with $j \leq k$ and j being even and let \mathbf{P}_{C_j} be the set of all overlaps of R and C_j with $i_R(R \setminus K) \cap i_{C_j}(C_j) \neq \emptyset$.

$$\text{ninf} := \bigwedge_{C \in U} \bigwedge_{C' \in \mathbf{P}_{C_j}} \text{Left}(\neg \exists(i_R : R \hookrightarrow C', \text{true}), \rho)$$

5. Let E be the set of all overlaps of L and C_k , such that each $P \in E$ is also an overlap of L and $C'' \in \mathcal{U}(C_k, C_{k+1})$, $i_{C_k} \models \exists(a'_k : C_k \hookrightarrow C'', \text{true})$ and $i_L(L \setminus K) \cap i_{C''}(C'' \setminus C_k) \neq \emptyset$.

$$\text{nwo} := \bigwedge_{P \in E} \neg \exists(i_L : L \hookrightarrow P, \text{true})$$

Lemma 3.4. Let a graph G , a constraint c in EANF, with $G \not\models c$, and a plain rule ρ be given. Then, the rule $\rho'(\rho, \text{ap}(c_{\max}^G, C))$ for a graph $C \in \mathcal{U}(C_{c_{\max}^G}, C_{c_{\max}^G + 1})$ is a minimal consistency improving rule.

Proof. Let $t : G \implies_{\rho', m} H$ be a transformation and $k = c_{\max}^G + 1$. We show, by contradiction, that this transformation is direct minimal consistency improving by showing that (2.1), (2.2), (2.3), (2.4) and (2.5) are satisfied. With that, it follows that ρ' is a minimal consistency improving rule. Let $d = \forall(a_k : C_k \hookrightarrow C_{k+1}, e)$ be the subcondition of c at layer k .

1. Assume that (2.1) does not hold. Then, there does exist an morphism $p : C_k \hookrightarrow G$, such that $p \models \text{part}(1, e, C')$, $\text{tr}_t \circ p$ is total and $\text{tr}_t \circ p \not\models \text{part}(1, e, C')$ for a graph $C' \in \mathcal{U}(C_k, C_{k+1})$. There has to exist an overlap P of L and C' such that $i_{C_k} \models \exists(a_k^p : C_k \hookrightarrow C', \text{true})$ and $m \models \exists(i_L : L \hookrightarrow P, \text{true})$. Then, nwo and with that $\text{ap}(c_{\max}^G, C)$ is not satisfied.
2. Assume that (2.2) does not hold. Then, a morphism $p' : C_k \hookrightarrow H$ with $p' \not\models \text{part}(1, e, C_{k+1})$ exists, such that there does not exist a morphism $p : C_k \hookrightarrow G$ with $\text{tr}_t \circ p = p'$. Then, an overlap P of R and C_k with $i_R(R \setminus K) \cap i_R(C_k) \neq \emptyset$ exists, such that $m \models \text{Left}(\exists(i_R : R \hookrightarrow P, \text{true}), \rho)$. Then, $m \not\models \text{ap}(c_{\max}^G, C)$.

3. Assume that (2.3) does not hold. Then, there does not exist a morphism $p : C_k \hookrightarrow G$ with $p \not\models \text{part}(1, e, C)$, such that $\text{tr}_t \circ p$ is not total or $\text{tr}_t \circ p \models \text{part}(1, e, C)$ and $\text{tr}_t \circ p$ is total. Then, no overlap P of L and C_k with $i_L(L \setminus K) \cap i_{C_k}(C_k) \neq \emptyset$ exists, such that $m \models \text{nex}(P, C)$. Also, no overlap P of C and R with $i_R(R \setminus K) \cap i_C(C) \neq \emptyset$ and $m \models \text{nex}(P, C)$ exists, such that $m \models \text{Left}(\exists(i_R : R \hookrightarrow P, \text{true}), \rho)$ and therefore $\text{rep}(P, C) = \text{false}$. It follows that for all $P \in \text{ol}(L, C_k)$ it holds that $\text{nex}(P, C) \wedge \text{rep}(P, C) = \text{false}$ and with that $m \not\models \text{ap}(c_{\max}^G, C)$.
4. Assume that (2.4) does not hold. Then, there does exist a morphism $p : C_j \hookrightarrow G$ with $j < k$ and i being even, such that no morphism $p' : C_j \hookrightarrow G$ with $\text{tr}_t \circ p' = p$ exists. Then, an overlap P of C_j and R with $i_R(R \setminus K) \cap i_{C_j}(C_j) \neq \emptyset$ exists, such that $m \models \text{Left}(\exists(i_R : R \hookrightarrow P, \text{true}), \rho)$. It follows that $m \not\models \text{rem}$ and with that $m \not\models \text{ap}(c_{\max}^G, C)$.
5. Assume that (2.5) does not hold. Then, there does exist and morphism $p : C_j \hookrightarrow G$ with $j < k$ and j being odd, such that $\text{tr}_t \circ p$ is not total. Then, an overlap P of C_j and L with $i_L(L \setminus K) \cap i_{C_j}(C_j) \neq \emptyset$ exists, such that $m \models \exists(i_L : L \hookrightarrow P, \text{true})$. It follows that $m \not\models \text{rem}$ and with that $m \not\models \text{ap}(c_{\max}^G, C)$.

In total follows that if $m \models \text{ap}(c_{\max}^G, C)$, then t is a direct minimal improving transformation. \square

Lemma 3.5. *Let G be a graph, c a constraint in EANF, with $c_{\max}^G < \text{nl}(c)$, and $\rho = L \leftarrow K \hookrightarrow R$ a plain rule. Let $d = \forall(a_k : C_{k-1} \hookrightarrow C_k, \exists(b : C_k \hookrightarrow C_{k+1}, e))$ be the subcondition of c at layer $c_{\max} + 1$ and $\text{ap}(c_{\max}, C')$ the application condition constructed by definition 3.3 with $C' \in \mathcal{U}(C, C_{k+1})$ for any $C \in \mathcal{G}_c^G$. If*

$$((R \setminus K) \cap C_{k+1}) \cup ((L \setminus K) \cap C_{k+1}) = \emptyset$$

$\text{ap}(c_{\max}, C')$ can be replaced by false .

Proof. There does not exist an overlap P of C_k and L with $i_L(L \setminus K) \cap i_{C_k}(C_k) \neq \emptyset$ and $\text{rep}(P, C')$ will be equal to false , if $C_k = C_{k+1}$, or equal to $\bigvee_{P' \in \mathbf{P}} \text{Left}(\exists(i_R : R \hookrightarrow P', \text{true}), \rho)$. Since the set \mathbf{P} has to be empty, this expression can be replaced by false . It follows that $\text{rep}(P, C') = \text{false}$ for all $P \in \text{ol}(L, C_k)$ and therefore $\text{ap}(c_{\max}, C')$ will always be evaluated to false . \square

3.1 potentially minimal improving rules

Definition 3.6 (basic improving rule). *Let a constraint c and a plain rule $\rho = L \xleftarrow{l} K \xrightarrow{r} R$ be given. The rule ρ is called basic improving w.r.t c at layer k with $C_k \subset P \subseteq C_{k+1}$, $p : C_k \hookrightarrow P$ being the inclusion, and $k \in \{1, 3, \dots, \text{nl}(c)\}$, if*

$$(L \setminus K) \cap (C_{k-1} \cup (C_{k+1} \setminus C_k)) = \emptyset \tag{3.3}$$

and

$$(R \setminus K) \cap C_k = \emptyset \tag{3.4}$$

and either 1. or 2. applies.

1. The rule ρ deletes elements of $C_k \setminus C_{k-1}$:

$$L \subseteq C_k \wedge L \setminus K \neq \emptyset \quad (3.5)$$

Then, ρ is called a deleting basic improving rule.

2. The rule ρ creates an instance of P :

$$L = C_k \wedge P \subseteq R \quad (3.6)$$

and p is a partial morphism of r . Then, ρ is called an inserting basic improving rule.

Definition 3.7 (application conditions for basic improving rules). Let a constraint c in EANF and a basic improving rule $\rho = L \xleftarrow{l} K \xrightarrow{r} R$ w.r.t c at layer k with $C_k \subseteq P \subseteq C_{k+1}$ be given. We define the application condition for r as:

1. If ρ is a deleting potentially minimal improving rule:

$$\text{ap}_{\text{pi}}(j, P) := \begin{cases} \bigvee_{P \in \text{ol}(L, C_k)} \bigwedge_{P' \in \text{ol}(P, a_k)} \exists(i_L : L \hookrightarrow P, \neg \exists(i_P : P \hookrightarrow P', \text{true})) & , \text{ if } j = k \\ \text{false} & , \text{ if } j \neq k. \end{cases}$$

2. If ρ is an inserting potentially minimal improving rule:

$$\text{ap}_{\text{pi}}(j, P) := \begin{cases} \neg \exists(a_k^p : L \hookrightarrow P, \text{true}) & , \text{ if } j = k \\ \text{false} & , \text{ if } j \neq k \end{cases}$$

Theorem 3.1. Let a graph G , a constraint c in EANF, with $G \models \text{part}(c_{\max}^G, c, C)$ and $\text{part}(c_{\max}^G, c, C) \in \mathcal{P}_c^G$, and a potentially minimal improving rule $\rho = L \xleftarrow{l} K \xrightarrow{r} R$ at layer c_{\max}^G with $C_{c_{\max}+1} \subseteq P \subseteq C_{c_{\max}+2}$ be given. Then, $\rho' = (\rho, \text{ap}_{\text{pi}}(c_{\max}^G + 1, P))$ is a direct minimal consistency improving rule.

Proof. Let $t : G \Rightarrow_{\rho', m} H$ be a transformation, $k = c_{\max}^G + 1$ and e be the subcondition of c at layer k . We show that t is a direct minimal consistency improving transformation. First, we show that equation (2.1) is satisfied. Let $p : C_k \hookrightarrow G$ be a morphism. If ρ is a deleting potentially minimal improving rule, either 1. or 2. applies, if ρ is an inserting and not a deleting potentially minimal improving rule, only 2. applies, because ρ cannot destroy any occurrences of C_k in G .

1. If $p(C_k) \cap m(L \setminus K) \neq \emptyset$, $\text{tr}_t \circ p$ is not total, since at least one element of $p(C_k)$ has been deleted by t and p does satisfy $\bigwedge_{C' \in \mathcal{U}(C_k, C_{k+1})} (p \models \text{part}(1, e, C') \wedge \text{tr}_t \circ p \text{ is total}) \implies \text{tr}_t \circ p \models \text{part}(1, e, C')$.
2. If $p(C_k) \cap m(L \setminus K) = \emptyset$, $\text{tr}_t \circ p$ is total, since no element of $p(C_k)$ has been deleted by t . Because (3.3) holds, t does not delete any elements of $C_{k+1} \setminus C_k$ and $p \models \text{part}(1, e, C') \implies \text{tr}_t \circ p \models \text{part}(1, e, C')$ for all $C' \in \mathcal{U}(C_k, C_{k+1})$.

With 1. and 2. follows that (2.1) is satisfied.

Second, we show that equation (2.2) is satisfied. Let $p' : C_k \hookrightarrow H$ be a morphism. Because (3.4) is satisfied, t does not create any elements of C_k and there must exist a morphism $p : C_k \hookrightarrow G$ with $\text{tr}_t \circ p = p'$. It follows that (2.2) is satisfied.

Third, we show that (2.3) is satisfied. We consider the cases that firstly, ρ is a deleting minimal potentially improving rule and secondly, that ρ is an inserting and not a deleting minimal potentially improving rule.

1. If ρ is a deleting potentially minimal improving rule, the condition $\text{ap}_{\text{pi}}(k, P) = \exists(b : L \hookrightarrow C_k, \neg \exists(a_{k+2} : C_k \hookrightarrow C_{k+1}, \text{true}))$ is satisfied by m . Therefore a morphism $p : C_k \hookrightarrow G$ with $p \not\models \neg \exists(a_{k+1} : C_k \hookrightarrow C_{k+1}, \text{true}) = \text{part}(1, e, C_{k+1})$ and $m = p \circ b$ must exist. Since ρ is a deleting rule, at least one element of $p(C_k)$ has been deleted by t and $\text{tr}_t \circ p$ is not total. It follows that (2.3) is satisfied.
2. If ρ is an inserting and not a deleting potentially minimal improving rule, $\text{ap}_{\text{pi}}(k, P) = \neg \exists(b : L \hookrightarrow P, \text{true})$ is satisfied by m . Because $L = C_k$, $m \models \neg \exists(b : C_k \hookrightarrow P, \text{true}) = \text{part}(1, e, P)$. Since (3.6) is satisfied, $\text{tr} \circ p$ is total and $\text{tr}_t \circ p \models \text{part}(1, e, P)$. Therefore, (2.3) is satisfied.

Last, since (3.4) and (3.3) are satisfied, ρ cannot create any occurrence of C_j with $j \leq k$ and j being even and ρ cannot delete any occurrences of C_j with $j \leq k$ and j being odd. Therefore, (2.4) and (2.5) are satisfied.

In total follows that t is a direct minimal consistency improving transformation and ρ' is a direct minimal consistency improving rule. \square

Theorem 3.2. *Let a constraint c in EANF, a basic increasing rule ρ with P and a consistency increasing transformation $t : G \implies_{\rho, m} H$ w.r.t c be given. Then, $m \models \text{ap}_{\text{pi}}(c_{\max}^G, P)$.*

Proof. Let $k = c_{\max}^G$.

1. If c is a inserting rule, then $\text{ap}_{\text{pi}}(c_{\max}^G, P) = \neg \exists(a_k^P : L \hookrightarrow P, \text{true})$. Let $n : R \hookrightarrow H$ be the comatch of t . Since t is a consistency increasing transformation, there does exist a morphism $p : C_k \hookrightarrow G$ with $p \not\models \exists(a_k^P : C_k \hookrightarrow P, \text{true})$, $\text{tr}_t \circ p \models \exists(a_k^P : C_k \hookrightarrow P, \text{true})$ such that $p'(P) \cap n(R \setminus K)$ for the morphism $p' : P \hookrightarrow H$ with $p = p \circ a_k^P$. Assume that $m \models \exists(a_k^P : L \hookrightarrow P, \text{true})$. Then, every morphism $p : C_k \hookrightarrow G$ with $\text{tr}_t \circ p \models \exists(a_k^P : C_k \hookrightarrow P, \text{true})$ with $p'(P) \cap n(R \setminus K)$ for the morphism $p' : P \hookrightarrow H$ with $p = p' \circ a_k^P$ already satisfies $\exists(a_k^P : C_k \hookrightarrow P, \text{true})$. This is a contradiction.
2. If ρ is a deleting rule, $\text{ap}_{\text{pi}}(k, P) = \bigvee_{P \in \text{ol}(L, C_k)} \bigwedge_{P' \in \text{eol}(P, a_k)} \exists(i_L : L \hookrightarrow P, \neg \exists(i_P : P \hookrightarrow P', \text{true}))$. Since t is a consistency increasing transformation, there does exist a morphism $p : C_k \hookrightarrow G$, such that $p \not\models \exists(a_k : C_k \hookrightarrow C_{k+1}, \text{true})$ and $\text{tr}_t \circ p$ is not total. Assume that $m \not\models \text{ap}_{\text{pi}}(k, P)$. Then, for each overlap Q of L and C_k , the inclusion i_{C_k} does satisfy $\exists(a_k : C_k \hookrightarrow C_{k+1}, \text{true})$. This is a contradiction, because $p(C_k) \cap m(L \setminus K) \neq \emptyset$ has to hold.

□

Lemma 3.8. *Let a constraint c in EANF and inserting basic increasing rules ρ_1, ρ_2 at layer k with P_1 and P_2 , respectively, be given. Then, the rules $\rho'_1 = (\rho_1, \text{ap}_{\text{pi}}(k, P_1))$ and $\rho'_2 = (\rho_2, \text{ap}_{\text{pi}}(k, P_2))$ are sequentially independent.*

Proof. Let $G_1 \Rightarrow_{\rho'_1, m_1} G_2 \Rightarrow_{\rho'_2, m_2} G_3$ be a sequence of transformations. First, note that this sequence can only exist if $P_1 \not\subseteq P_2$ and $P_2 \not\subseteq P_1$. Since both rules are do not insert elements of C_k , a morphism $d_1 : L_2 \hookrightarrow D_1$ with $m_2 = h_1 \circ d_1$ exists. Because both rules do not delete any elements a morphism $d_2 : R_1 \hookrightarrow D_2$ with $m_1^* = g_2 \circ d_2$ exists. Because ρ_2 is not a basic increasing rule with P_1 , $h_2 \circ d_2 \models R(\text{ap}_{\text{pi}})$ and because ρ_1 does not delete any elements of P_2 it holds that $g_1 \circ d_1 \models R(\text{ap}_{\text{pi}})$. □

Definition 3.9 (repairing rule set). *Let a constraint c in EANF and a set of rules \mathcal{R} be given. Then, \mathcal{R} is called a repairing rule set for c at layer k if for all graphs G with $k = c_{\max}$ a sequence*

$$G = G_0 \Rightarrow_{r_0} \dots \Rightarrow_{r_{n-1}} G_n = H$$

exists, such that $r_j \in \mathcal{R}$ for all $j \in \{0, \dots, n-1\}$ and $H \models_{k+2} c$.

Corollary 3.10. *Let a constraint c in EANF and a set of rules \mathcal{R} be given. If \mathcal{R} is a repairing rule set for c at layer k , \mathcal{R} is a repairing rule set w.r.t c at layer j for all $k < j \leq \text{nl}(c)$.*

Corollary 3.11. *Let a constraint c in EANF and a repairing rule set \mathcal{R} for c at layer k , for all $k \in \{1, 3, \dots, \text{nl}(c)\}$, be given. Then, for all graphs G , a sequence*

$$G = G_0 \Rightarrow_{r_0} \dots \Rightarrow_{r_{n-1}} G_n = H$$

exists, such that $r_j \in \mathcal{R}$ for all $j \in \{0, \dots, n-1\}$ and $H \models c$.

Definition 3.12 (decomposition of a graph). *Let graphs $G_0 \subset G_1$ be given.*

A decomposition of G_1 with G_0 is a minimal set

$$\mathbf{P} \subseteq \{P_v \mid v \in V_{C_1 \setminus C_0}\}$$

of subgraphs of G_1 , such that every element of G_1 is contained in at least one $P \in \mathbf{P}$ and every P_v is constructed in the following way: $G_0 \subset P_v$, $v \in P$ and for all nodes $v' \in P \setminus C_k$ it holds that P contains all edges $e \in E_{G_1 \setminus G_0}$ and all nodes $u \in V_{G_1}$ such that either $\text{tar}(e) = v' \wedge \text{src}(e) = u$ or $\text{src}(e) = u \wedge \text{tar}(e) = v'$ holds.

Lemma 3.13. *Let graphs $G_0 \subset G_1$ and a decomposition \mathbf{P} of G_1 with G_0 be given. Then, for each pair $P, P' \in \mathbf{P}$ with $P \neq P'$ the following holds:*

$$(P \setminus C_k) \cap (P' \setminus C_k) = \emptyset$$

Proof. Assume that $(P \setminus C_k) \cap (P' \setminus C_k) \neq \emptyset$, therefore a node $v \in G_1 \setminus G_0$ with $v \in P \cap P'$ exists. By the construction of P and P' it follows that $P = P_v$ and $P' = P_v$ and therefore $P = P'$. This is a contradiction. □

Lemma 3.14. Let graphs $G_0 \subset G_1$ and a decomposition \mathbf{P} of G_1 with G_0 be given. Then,

$$G_1 = \bigcup_{P \in \mathbf{P}} P.$$

Proof. Let $H := \bigcup_{P \in \mathbf{P}} P$. Firstly, we show that $H \subseteq G_1$. Since every $P \in \mathbf{P}$ is a subgraph of G_1 it follows that $V_H \subseteq V_{G_1}$ and $E_H \subseteq E_{G_1}$.

Secondly, we show that $G_1 \subseteq H$. Let $u \in V_{G_1}$ be a node, if $u \in V_{G_0}$, then u is contained in every $P \in \mathbf{P}$ and therefore $u \in V_H$. Otherwise, if $u \notin V_{G_0}$, then u has to be, by the definition of \mathbf{P} , contained in at least one $P \in \mathbf{P}$ and $V_{G_1} \subseteq V_H$ follows. Let $e \in E_{G_1}$ be an edge. If $e \in E_{G_0}$, then e is contained in every $P \in \mathbf{P}$ and $e \in E_H$. Otherwise, if $e \notin E_{G_0}$, by the definition of \mathbf{P} , e has to be contained in at least one $P \in \mathbf{P}$. It follows that $e \in E_H$ and with that $E_{G_1} \subseteq E_H$. \square

Theorem 3.3. Let a constraint c in EANF and a set of rules \mathcal{R} be given. Then, \mathcal{R} is a repairing set of c at layer $k \leq \text{nl}(c)$ if either 1 or 2 applies.

1. For any universally bound graph C_j at layer $j \leq k$ of c , $(\rho, \text{appi}(j, C_{j+1})) \in \mathcal{R}$ and ρ is a deleting potentially minimal improving rule at layer j with C_{j+1} , such that ρ only deletes edges of C_j .
2. A decomposition \mathbf{P} of C_k with C_{k-1} exists, such that for each $P \in \mathbf{P}$ a rule $(\rho, \text{appi}(k, P)) \in \mathcal{R}$ exists, such that ρ is an inserting basic improving rule at layer k with P .

Proof. Let a constraint c in EANF, a rule set \mathcal{R} and a graph G with $k = c_{\max}$ and $c_{\max} < \text{nl}(c)$ be given. We show that a sequence $G = C'_0 \Rightarrow \dots \Rightarrow C'_n = H$ with rules of \mathcal{R} exists, such that $H \models_{k+2} c$ if 1. or 2. of theorem 3.3 is satisfied.

1. Assume that 1. of theorem 3.3 holds. Let $(\rho, \text{appi}(j, C_{j+1})) \in \mathcal{R}$, such that $\rho = L \xleftarrow{l} K \xrightarrow{r} R$ is a deleting potentially minimal improving rule at layer $j \leq k$ with C_{j+1} and C_j is a universally bound graph of c . Then, $\text{appi}(j, C_{j+1}) = \exists(b : L \hookrightarrow C_j, \neg \exists(a_j : C_j \hookrightarrow C_{j+1}, \text{true}))$. Let $q : C_j \hookrightarrow G$ be a morphism such that $q \not\models \exists(a_j : C_j \hookrightarrow C_{j+1}, \text{true})$. Since $L \subseteq C_j$, we can construct a morphism $m_1 : L \hookrightarrow G$ with $m_1 = q \circ b$ and therefore $m_1 \models \text{appi}(j, C_{j+1})$. Since r only deletes edges, a transformation $t : G = G_0 \Rightarrow_{r, m_1} G_1$ exists and $\text{tr}_t \circ p$ is not total. Because r does not insert any elements of C_j :

$$|\{q : C_k \hookrightarrow G_0 \mid q \not\models d\}| < |\{q : C_k \hookrightarrow G_1 \mid q \not\models d\}|$$

with $d = \exists(a_j : C_j \hookrightarrow C_{j+1}, \text{true})$. By iteratively applying this construction, we can generate a finite sequence of transformations

$$G = G_0 \Rightarrow_{r, m_1} G_1 \Rightarrow_{r, m_2} \dots \Rightarrow_{r, m_n} G_n = H$$

such that $|\{q : C_k \hookrightarrow G_n \mid q \not\models d\}| = 0$ and therefore $H \models_j c$. With lemma 2.9, $H \models_{k+2} c$ and $H \models c$ follows.

2. Assume that 2. of theorem 3.3 holds. Let $\rho_0 = (\rho, \text{app}_\text{pi}(k, P)) \in \mathcal{R}$, such that ρ is an inserting basic improving rule of c at layer k with $P \in \mathbf{P}$. Then, $\text{app}_\text{pi}(k, P) = \neg \exists(b : L \hookrightarrow P, \text{true})$. Let $q_0 : C_k \hookrightarrow G$ be a morphism, such that $q_0 \not\models \exists(a'_k : C_k \hookrightarrow P, \text{true})$ with a'_k being a partial morphism of a_k . Since $L = C_k$, we set $m_0 : C_k \hookrightarrow G$ with $m_0 = q_0$. It follows that $m_0 \models \neg \exists(a'_k : C_k \hookrightarrow P, \text{true}) = \text{app}_\text{pi}(k, P)$. Because r does not delete any elements, a transformation $t_0 : G \Rightarrow_{r_0, m_0} G_1$ exists and $\text{tr}_t \circ q \models \exists(a'_k : C_k \hookrightarrow P, \text{true})$. We set $q_1 = \text{tr}_{t_0} \circ q_0$ and apply the same method to q_1 .

By iteratively applying this, we can construct a finite sequence of transformations

$$G \Rightarrow_{r_0, m_0} G_0 \Rightarrow_{r_1, m_1} \dots \Rightarrow_{r_n, m_n} G_n$$

such that $m_i = \text{tr}_{t_{i-1}} \circ \dots \circ \text{tr}_{t_0} \circ m_0$ and $q \models \exists(b_i : C_k \hookrightarrow P_i, \text{true})$ for all $P_i \in \mathbf{P}$ with $q = \text{tr}_{t_n} \circ q_n$. Let $p_i : P_i \hookrightarrow G_n$ be the morphism, such that $q = p_i \circ b_i$.

Now, we can construct a morphism $p : C_{k+1} \hookrightarrow G$ with

$$p(e) := \begin{cases} p_1(e) & \text{if } e \in P_1 \\ & \vdots \\ p_j(e) & \text{if } e \in P_j. \end{cases}$$

Let $e \in C_k$, because $q(e) = p_i \circ b_i(e)$ and $q(e) = p_\ell \circ b_\ell(e)$ and b_i and b_ℓ are both partial morphisms of a_k , it follows that $b_i(e) = b_\ell(e)$ and therefore $p_i(e) = p_\ell(e)$. Because $(P_i \cap P_\ell) \setminus C_k = \emptyset$ for all $i \neq \ell$, p is a morphism and by the definition of p it follows that $q = p \circ a_k$ and therefore $q \models \exists(a_k : C_k \hookrightarrow C_{k+1}, \text{true})$.

By iteratively applying this whole construction to all occurrences of C_k that do not satisfy $\exists(a_k : C_k \hookrightarrow C_{k+1}, \text{true})$ the derived graph H does not contain any occurrences of C_k not satisfying $\exists(a_k : C_k \hookrightarrow C_{k+1}, \text{true})$ and therefore $H \models_{k+2} c$.

□

Corollary 3.15. *If a set of rule \mathcal{R} is a repairing set of c at layer $k \leq \text{nl}(c)$ and 1. of theorem 3.3 applies, then \mathcal{R} is a repairing set of c at layer j for all $k \leq j \leq \text{nl}(c)$.*

Lemma 3.16. *Let a graph G_0 , a constraint c in EANF and a repairing set \mathcal{R} at layer $c_{\max}^{G_0} + 1$ be given, such that each rule in \mathcal{R} is a basic increasing. Then, for every sequence*

$$G_0 \Rightarrow_{\rho_0, m_0} G_1 \Rightarrow_{\rho_1, m_1} \dots \Rightarrow_{\rho_n, m_n} G_n$$

such that $\rho_i = (\rho, \text{app}_\text{pi}(c_{\max}^{G_0} + 1, P))$ with $\rho \in \mathcal{R}$ being a consistency increasing rule at layer $c_{\max}^{G_0} + 1$ with P , it holds that

References

- [1] C. Sandmann and A. Habel. Rule-based graph repair. *arXiv preprint arXiv:1912.09610*, 2019.