

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

Alexander Lauer

January 4, 2023

## **Abstract**

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Related Work</b>	<b>3</b>
<b>3</b>	<b>Preliminaries</b>	<b>3</b>
3.1	Graphs and Graph morphisms . . . . .	3
3.2	Nested Graph Conditions and Constraints . . . . .	4
3.3	Rules and Graph Transformations . . . . .	6
<b>4</b>	<b>Consistency Increase- and Maintenance</b>	<b>7</b>
4.1	Universally quantified ANF . . . . .	8
4.2	Conditions up to Layer . . . . .	9
4.3	Consistency increasing transformations and rules . . . . .	13
4.4	Comparison with other consistency concepts . . . . .	18
<b>5</b>	<b>general application conditions</b>	<b>21</b>
5.1	Conflicts within and between conditions . . . . .	25
5.2	Basic increasing rules . . . . .	27
5.3	Repairing rule Sets . . . . .	34
5.3.1	Construction of repairing sets . . . . .	35
<b>6</b>	<b>Rule-based Graph Repair</b>	<b>36</b>
6.1	Rule based Graph repair for conflict free constraints . . . . .	36
6.2	Rule based Graph repair for non-conflict free constraints . . . . .	37
6.3	Rule based repair for non-repairing sets . . . . .	37
<b>7</b>	<b>Conclusion</b>	<b>37</b>

# 1 Introduction

## 2 Related Work

## 3 Preliminaries

Our graph repair process is based on the concept of the double-pushout approach [1]. In this chapter we will introduce some formal prerequisites like graphs, graph morphisms, nested graph conditions and constraints and graph transformations.

### 3.1 Graphs and Graph morphisms

We start by introducing the notions of graphs and graph morphisms according to [1].

**Definition 3.1 (graph).** A graph  $G = (V, E, \text{src}, \text{tar})$  consists of a set of vertices  $V$ , a set of edges  $E$  and two mappings  $\text{src}, \text{tar} : E \rightarrow V$  that assign the source and target vertices to an edge. That means, given an edge  $e \in E$ , then  $e$  connects the vertices  $\text{tar}(e)$  and  $\text{src}(e)$ .

If a tuple as above is not given,  $V_G$ ,  $E_G$ ,  $\text{tar}_G$  and  $\text{src}_G$  denotes the sets of vertices, edges and the target and source mappings, respectively.

For the remainder of this paper we assume that all graphs are finite, that means given a graph  $G$  the sets  $V_G$  and  $E_G$  are finite.

**Definition 3.2 (graph morphism).** Let graphs  $G$  and  $H$  be given. A graph morphism  $f : G \rightarrow H$  consists of two mappings  $f_V : V_G \rightarrow V_H$  and  $f_E : E_G \rightarrow E_H$  such that the source and target functions are preserved. That means,

$$\begin{aligned} f_V \circ \text{src}_G &= \text{src}_H \circ f_E \wedge \\ f_V \circ \text{tar}_G &= \text{tar}_H \circ f_E \end{aligned}$$

holds. A graph morphism  $f$  is called injective (surjective) if  $f_E$  and  $f_V$  are injective (surjective) mappings. If  $f$  is injective, it is denoted by  $f : G \hookrightarrow H$  and we may also call it inclusion. Two morphisms  $f_1 : G_1 \rightarrow H$  and  $f_2 : G_2 \rightarrow H$  are called jointly surjective if for each element  $e$  of  $H$  either an element  $e' \in G_1$  with  $f_1(e') = e$  or an element  $e' \in G_2$  with  $f_2(e') = e$  exists.

For our newly introduced notions of consistency increase- and maintenance we also need to consider *sub-graphs*, *overlaps* of graphs and so-called *intermediate-graphs* which intuitively denotes graphs  $G'$  that lay in between two given graphs  $G$  and  $H$ . That means,  $G$  is a sub-graph of  $G'$  and  $G'$  is a sub-graph of  $H$ .

**Definition 3.3 (sub-graph).** Let  $G$  and  $H$  be graphs. Then,  $G$  is called a sub-graph of  $H$  if an injective morphism  $p : G \hookrightarrow H$  exists.

Note that with this definition every graph  $G$  is a sub-graph of itself since the morphism  $p$  is also allowed to be surjective.

**Definition 3.4 (intermediate-graph).** Let  $G$  and  $H$  be graphs such that  $G$  is a sub-graph of  $H$ . A graph  $C$  is called an intermediate-graph of  $G$  and  $H$ , if  $G$  is a sub-graph of  $C$  and  $C$  is a sub-graph of  $H$ . The set of intermediate-graphs of  $G$  and  $H$  is denoted by  $\text{IG}(G, H)$ .

**Definition 3.5 (overlap).** Let  $G_1$  and  $G_2$  be graphs. A graph  $H$  is called an overlap of  $G_1$  and  $G_2$  if morphisms  $i : G_1 \hookrightarrow H$  and  $i' : G_2 \hookrightarrow H$  exist such that  $i$  and  $i'$  are jointly surjective and  $i(G_1) \cap i'(G_2) \neq \emptyset$ . The morphisms  $i$  and  $i'$  are called overlap morphisms.

The set of all overlaps of  $G_1$  and  $G_2$  is denoted by  $\text{ol}(G_1, G_2)$ . Given an overlap  $H$  of  $G_1$  and  $G_2$ , the overlap morphisms are denoted by  $i_{G_1}^H$  and  $i_{G_2}^H$ , respectively.

Note that, graphs  $H$  such that jointly surjective morphisms  $i : G_1 \hookrightarrow H$  and  $i' : G_2 \hookrightarrow H$  such that  $i(G_1) \cap i'(G_2) = \emptyset$  can also be considered as overlap of  $G_1$  and  $G_2$ . Since, for this paper we will only consider overlaps that satisfy  $i(G_1) \cap i'(G_2) \neq \emptyset$ , we directly embedded this property into the definition.

As mentioned above, our approach also considers intermediate-graphs. Therefore a notion to restrict given graph morphisms  $p : G \rightarrow H$  is needed. For this, we introduce the notion of *restricted morphisms* which intuitively is the restriction of the domain and co-domain of  $p$  with sub-graphs of  $G$  and  $H$ , respectively.

**Definition 3.6 (restricted morphism).** Let graphs  $G$ ,  $H$  and a morphism  $f : G \rightarrow H$  be given. Then, a morphism  $f' : G' \rightarrow H'$  is called a restricted morphism of  $p$  if morphisms  $i : G' \hookrightarrow G$  and  $i' : H' \hookrightarrow H$  exists ( $G'$  is a subgraph of  $G$  and  $H'$  a sub-graph of  $H$ ) and

$$\begin{aligned} i'_E \circ f'_E &= f_E \circ i_E \wedge \\ i'_V \circ f'_V &= f_V \circ i_V \end{aligned}$$

holds. Given an morphism  $p$ , we use the notation  $p^r$  to denote that  $p^r$  is a restricted morphism of  $p$ .

Note that, given a morphism  $p : G \rightarrow H$  a restriction  $p^r : G' \rightarrow H'$  of  $p$  is uniquely determined by  $G'$  and  $H'$  meaning that, given two restrictions  $q : G' \rightarrow H'$  and  $q' : G' \rightarrow H$  of  $p$  it follows immediately that  $q = q'$ .

### 3.2 Nested Graph Conditions and Constraints

*Nested graph constraints* are useful to specify graph properties. The more general notion of *nested graph conditions* allows the specification of properties for graph morphisms and the definition of graph constraints in a recursive manner. These conditions allow the use of quantifiers and boolean operators.

**Definition 3.7 (nested graph condition[2]).** A nested graph condition over a graph  $C_0$  is recursively defined as:

1. *true* is a graph condition over every graph.

2.  $\exists(a_0 : C_0 \hookrightarrow C_1, d)$  is a graph condition over  $C_0$  if  $a_0$  is a injective graph morphism and  $d$  is a graph condition over  $C_1$ .
3.  $\neg d$ ,  $d_1 \wedge d_2$  and  $d_1 \vee d_2$  are graph conditions over  $C_0$  if  $d$ ,  $d_1$  and  $d_2$  are graph conditions over  $C_0$ .

Conditions over the empty graph  $\emptyset$  are called constraints. We use the abbreviations  $\forall(a_0 : C_0 \hookrightarrow C_1, d) := \neg\exists(a_0 : C_0 \hookrightarrow C_1, \neg d)$  and  $\text{false} = \neg \text{true}$ .

Conditions of the form  $\exists(a_0 : C_0 \hookrightarrow C_1, d)$  are called existentially bound, the graph  $C_1$  is also called existentially bound. Conditions of the form  $\forall(a_0 : C_0 \hookrightarrow C_1, d)$  are called universally bound, the graph  $C_1$  is also called universally bound.

Since these are the only types of conditions used in this paper we will only refer to it as *conditions* and *constraints*.

**Definition 3.8 (semantic of graph conditions).** Let a graph  $G$ , a condition  $c$  over  $C_0$  and a graph morphism  $p : C_0 \hookrightarrow G$  be given. Then,  $p$  satisfies  $c$ , denoted by  $p \models c$  if

1. If  $c = \text{true}$ ,  $p \models c$ .
2. If  $c = \exists(a_0 : C_0 \hookrightarrow C_1, d)$ , then  $p \models c$  if an injective morphism  $q : C_1 \hookrightarrow G$  with  $p = q \circ a$  exists, such that  $q \models d$ .
3. If  $c = \neg d$ , then  $p \models c$  if  $p \not\models d$ .
4. If  $c = d_1 \wedge d_2$ , then  $p \models c$  if  $p \models d_1$  and  $p \models d_2$ .
5. If  $c = d_1 \vee d_2$ , then  $p \models c$  if  $p \models d_1$  or  $p \models d_2$ .

A graph  $G$  satisfies a constraint  $c$ , denoted by  $G \models c$ , if the morphism  $p : \emptyset \hookrightarrow G$  satisfies  $c$ .

Our approach is designed to repair a specific type of nested constraint, namely *linear conditions* which are conditions without any boolean operators. In particular, our approach is able to repair constraints in *alternating quantifier normal form (ANF)* which are linear conditions with alternating quantifiers, since every linear condition can be transformed into an equivalent condition in ANF [6].

**Definition 3.9 (alternating quantifier normal form (ANF))**[6]. Conditions in alternating quantifier normal form (ANF) are inductively defined as

1. *true* and *false* are conditions in ANF.
2.  $\exists(a_0 : C_0 \hookrightarrow C_1, d)$  is a condition in ANF if either  $d$  is universally bound or  $d = \text{true}$ .
3.  $\forall(a_0 : C_0 \hookrightarrow C_1, d)$  is a condition in ANF if either  $d$  is existentially bound or  $d = \text{false}$ .

In both cases,  $d$  is called a sub-condition of  $\exists(a : C_0 \hookrightarrow C_1, d)$  or  $\forall(a : C_0 \hookrightarrow C_1, d)$  respectively. The nesting level  $\text{nl}(c)$  of a condition  $c$  is recursively defined as  $\text{nl}(\text{true}) = 0$  and  $\text{nl}(\exists(a : P \rightarrow Q, d)) := \text{nl}(d) + 1$ .

In literature conditions in ANF allow also conditions that end with  $\exists(a_0 : C_0 \hookrightarrow C_1, \text{false})$  or  $\forall(a_0 : C_0 \hookrightarrow C_1, \text{true})$ . We exclude these cases such that conditions in ANF are only allowed to end with conditions of the type  $\exists(a_0 : C_0 \hookrightarrow C_1, \text{true})$  or  $\forall(a_0 : C_0 \hookrightarrow C_1, \text{false})$  since it can be easily seen that every morphism  $p : C_0 \hookrightarrow G$  satisfies  $\forall(a_0 : C_0 \hookrightarrow C_1, \text{true})$  and does not satisfy  $\exists(a_0 : C_0 \hookrightarrow C_1, \text{false})$ . Therefore, these conditions can be replaced by `true` and `false` respectively.

In the following, we assume that all conditions are finite. As a direct consequence of this, the nesting level is also finite.

### 3.3 Rules and Graph Transformations

Via *rules* and *graph transformation* to manipulate given graphs by the deletion and insertion of edges and vertices. For our approach, we use the concept of the double-pushout approach for rules and transformations [1]. A rule consists of the three graphs  $L$ , called the *left-hand side*,  $K$ , called *context* and  $R$  called *right-hand side* with  $K$  being a sub-graph of  $L$  and  $R$ . During a transformation, denoted by  $G \Rightarrow H$ , elements of  $L \setminus K$  will be removed and elements of  $R \setminus K$  will be inserted such that a new morphism  $p : R \hookrightarrow H$  has been created. Additionally, the so-called *dangling edge condition* has to be satisfied, meaning that for each edge  $e \in E_H$  there do exists vertices  $u, v \in V_H$  such that  $\text{tar}(e) = u$  and  $\text{src}(e) = v$  or vice versa. Also, we define application conditions which are nested conditions over  $L$  and  $R$  which are able prevent transformation if these are not satisfied. We will use the concept of application conditions to guarantee that transformations are not able to decrease consistency. For example, application conditions that prevent a transformation if  $G \models c$  and  $H \not\models c$ , given a constraint  $c$ .

**Definition 3.10 (rules and application conditions).** A plain rule  $\rho' = L \xleftarrow{l} K \xrightarrow{r} R$  consists of graphs  $L, K, R$  and injective graph morphisms  $l : K \hookrightarrow L$  and  $r : K \hookrightarrow R$ . The rule  $\rho'^{-1} = R \xleftarrow{r} K \xrightarrow{l} L$  is called the inverse rule of  $\rho'$ .

An application condition is a nested condition over  $L$  or  $R$ , respectively. A rule  $(\text{ap}_L, \rho', \text{ap}_R)$  contains of a plain rule  $\rho'$  and application conditions  $\text{ap}_L$  over  $L$ , called left application condition, and  $\text{ap}_R$  over  $R$ , called right application condition, respectively.

**Definition 3.11 (graph transformation).** Let a rule  $\rho = (\text{ap}_L, \rho', \text{ap}_R)$ , a graph  $G$  and a morphism  $m : L \hookrightarrow G$ , called match, be given. Then, a graph transformation  $t : G \Rightarrow_{\rho, m} H$  is given in figure 1 if the squares (1) and (2) are pushouts in the sense of category theory,  $m \models \text{ap}_L$  and the morphism  $n : L \rightarrow H$ , called the co-match of  $t$ , satisfies  $\text{ap}_R$ .

The presence of right applications conditions leads to unpleasant side-effects. The satisfaction of a right application condition can only be checked after the transformation. Therefore, in case that this application condition is not satisfied, the transformation must

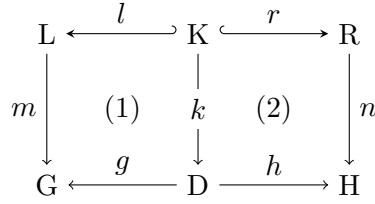


Figure 1: Diagram of a transformation in the double-pushout approach.

be reversed. To avoid this, we introduce the *shift over rule* construction which is able to transform a right into a left application condition [2]. This notion is based on the notion of *shift over morphism* which allows to shift a condition over  $C_0$  over a morphism  $p : C \hookrightarrow C_0$ .

**Definition 3.12 (shift over rule).** Let a plain rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  and a right application condition  $ap$  of the form  $\neg\exists(m : R \hookrightarrow C, \text{true})$  be given. Then,  $ap$  can be shifted into a left application condition with the transformation  $t : C \Rightarrow_{\rho^{-1}, m} C'$ . The shifted condition over  $\rho$ ,  $\text{Left}(ap, \rho)$ , is given by  $\neg\exists(n : L \hookrightarrow C', \text{true})$  with  $n$  being the co-match of  $t$ .

Shift over rule produces an equivalent left application condition, meaning that, given a right application condition  $ap$  and a plain rule  $\rho$ , a match of a transformation satisfies  $\text{Left}(ap, \rho)$  if and only if the co-match satisfies  $ap$  [2]. In general, every right application condition can be shifted into an equivalent left application condition and vice versa. Since we only need to shift conditions of the form  $\neg\exists(a : R \hookrightarrow C, \text{true})$  we introduced the shift only for these kind of application conditions and assume that each rule only contains left application conditions, denoted by  $(ap, \rho)$ .

Via the *track morphism* it is possible to track elements across a transformation [5].

**Definition 3.13 (track morphism).** Consider the transformation  $t$  shown in figure 1. The track morphism,  $\text{tr}_t : G \rightarrow H$ , of  $t$  is defined as

$$\text{tr}_t = \begin{cases} h(g^{-1}(e)) & \text{if } e \in g(D) \\ \text{undefined} & \text{otherwise.} \end{cases}$$

For example, given a transformation  $t : G \Rightarrow H$ , via the track morphism it can be checked whether a morphism  $p : C \hookrightarrow G$  is still present in the derived graph  $H$  by checking whether  $\text{tr}_t \circ p$  is total, or that a new morphism  $q : C \hookrightarrow H$  has been inserted by checking that no morphism  $p : C \hookrightarrow H$  with  $q = \text{tr}_t \circ p$  exists.

## 4 Consistency Increase- and Maintenance

**Definition 4.1 (layer of a sub-condition).** Let  $c$  be a condition in ANF and  $d$  a sub-condition of  $c$ . The layer of  $d$  is defined as  $\text{lay}(d) := \text{nl}(c) - \text{nl}(d)$ .

Our approach is based on the idea that the consistency of a constraint is increased layer by layer and even slight improvements like the insertion of single elements of existentially bound graphs should be detectable as increasing. To formalise this, we introduce the notions of *consistency increasing* and *consistency maintaining* transformations and rules with consistency increasing indication that the consistency has indeed been increased and consistency maintaining indicating that the consistency has not been decreased.

#### 4.1 Universally quantified ANF

The definition of consistency increase- and maintenance demands that every existentially bound sub-condition of a given condition in ANF is embedded into an universally bound condition. Otherwise, case discrimination is needed. This requirement is not met if and only if the given condition is existentially bound. Therefore, we will only consider a subset of the set of conditions in ANF, namely the set of universally quantified conditions in ANF, called *universally quantified ANF* (UANF). Additionally, we will show that these sets are expressively equivalent by showing that each condition in ANF can be transformed into an equivalent condition in UANF.

**Definition 4.2 (universally quantified alternating quantifier normal form).** A conditions  $c$  in ANF is in universally quantified ANF (UANF) if it is universally bound.

Note that, given a condition  $c$  in UANF, every sub-condition of  $c$  at layer  $0 \leq k \leq \text{nl}(c)$  is universally bound if  $k$  is an even number and existentially bound if  $k$  is an odd number. The extension of conditions that is used to show that a conditions in ANF can be transformed into an equivalent condition in UANF is an already known concept [4].

**Lemma 4.3.** Any condition in ANF can be transformed into an equivalent condition in UANF.

*Proof.* Let a graph  $G$  and a constraint  $c$  in ANF be given. If  $c$  is universally bound,  $c$  is already in UANF.

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  $q \models c$ . Then,  $p \models c'$ , 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.

□

For the rest of this thesis, given a condition  $c = \forall(a_0 : C_0 \hookrightarrow C_1, d)$  in UANF, we assume that no morphism in  $c$ , except  $a_0$ , is bijective since it can be shown that each condition in ANF can be transformed into an equivalent condition in ANF fulfilling this property by showing that  $\exists(a_0 : C_0 \hookrightarrow C_0, \forall(a_1 : C_0 \hookrightarrow C_2, d))$  is equivalent to  $\forall(a_1 \circ a_0 : C_0 \hookrightarrow C_2, d)$  and that  $\forall(a_0 : C_0 \hookrightarrow C_0, \exists(a_1 : C_0 \hookrightarrow C_2, d))$  is equivalent to  $\exists(a_1 \circ a_0 : C_0 \hookrightarrow C_2, d)$ .

## 4.2 Conditions up to Layer

As already mentioned, the goal of our approach is to increase the consistency of a constraint layer by layer. For this, we introduce a notion of partial consistency, called *satisfaction at layer* which enables to check whether a constraint is satisfied at a certain layer by checking whether the so-called *truncated condition* at this layer is satisfied.

**Definition 4.4 (sub-condition at layer).** Let  $c$  be a condition in ANF. The sub-condition at layer  $0 \leq k \leq \text{nl}(c)$ , denoted by  $\text{sub}_k(c)$ , is the sub-condition  $d$  of  $c$  with  $\text{lay}(d) = k$ .

**Example 4.1.** Consider the condition  $c = \forall(a_0 : C_0 \hookrightarrow C_1, \exists(a_1 : C_1 \hookrightarrow C_2, \forall(a_2 : C_2 \hookrightarrow C_3, \text{false})))$ . Then,  $\text{sub}_1(c) = \exists(a_1 : C_1 \hookrightarrow C_2, \forall(a_2 : C_2 \hookrightarrow C_3, \text{false}))$ .

First, we introduce an operator which allows to replace a sub-condition  $\text{sub}_k(c)$  by an arbitrary condition over  $C_k$ , called *replacement at layer*.

**Definition 4.5 (replacement at layer).** Let a condition  $c = Q(a_0 : C_0 \hookrightarrow C_1, d)$ , with  $Q \in \{\forall, \exists\}$  in ANF and a condition  $e$  over  $C_k$  in ANF be given. The replacement in  $c$  at layer  $k$  with  $e$ , denoted by  $\text{rep}_k(c, e)$ , is recursively defined as:

$$\text{rep}_k(c, e) := \begin{cases} e & \text{if } k = 0 \\ Q(a_0 : C_0 \hookrightarrow C_1, \text{rep}_{k-1}(d, e)) & \text{otherwise} \end{cases}$$

**Example 4.2.** Let the conditions  $c := \forall(a_0 : C_0 \hookrightarrow C_1, \exists(a_1 : C_1 \hookrightarrow C_2, \text{true}))$  and  $d = \exists(a'_1 : C_1 \hookrightarrow C_3, e)$  be given. Then,

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

Using replacement at layer we now define *truncated conditions*. Intuitively, a condition is cut off at a certain layer, by replacing the sub-condition at this layer by `true` or `false`, depending on the quantifier, the replaced sub-condition is bound by.

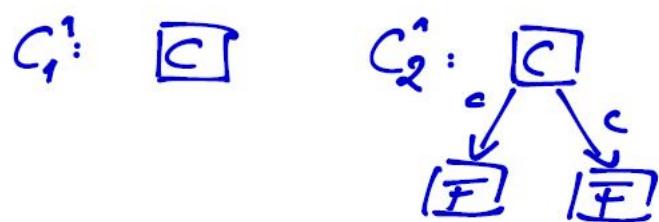
**Definition 4.6 (truncated condition).** Let  $c$  be a condition in UANF and  $d = \text{sub}_k(c)$  with  $0 \leq k \leq \text{nl}(c)$ . The truncated condition at layer  $k$  of  $c$ , denoted by  $\text{cut}_k(c)$ , is defined as

$$\text{cut}_k(c) := \begin{cases} \text{rep}_k(c, \text{false}) & \text{if } d \text{ is existentially bound, i.e. } k \text{ is odd} \\ \text{rep}_k(c, \text{true}) & \text{if } d \text{ is universally bound, i.e. } k \text{ is even.} \end{cases}$$

**Example 4.3.** Consider constraint  $c_2$  given in Figure 2. Then,  $\text{cut}_1(c_2) = \forall C_1^1 \exists C_2^2$ .

With these prerequisites we are now able to introduce *satisfaction at layer* which enables to check whether a condition is satisfied at a certain layer. A morphism or graph satisfies a condition or constraint, respectively, if it satisfies the truncated condition at this layer.

$$G = \forall C_1 \exists C_2$$



$$C_2 = \forall C_1 \exists C_2 \forall C_3 \exists C_4$$

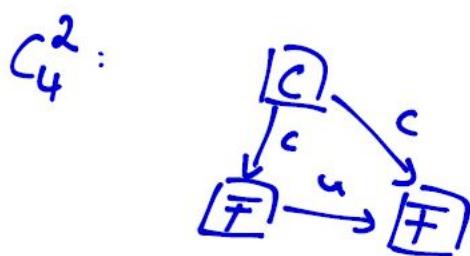
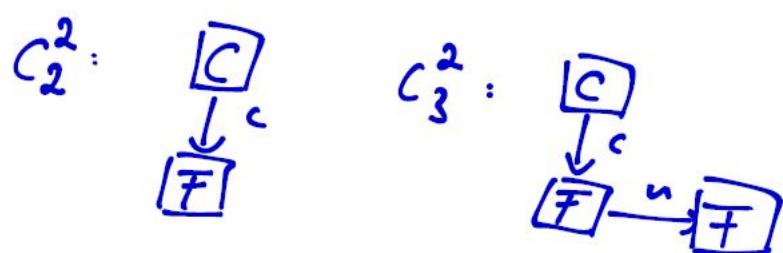


Figure 2: constraints

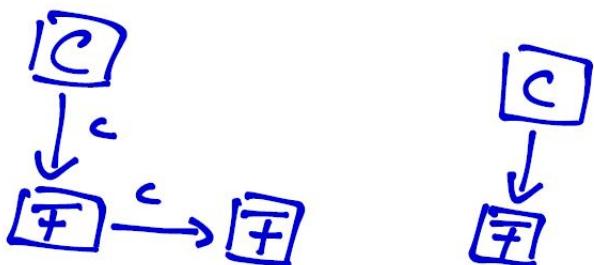


Figure 3: graph

**Definition 4.7 (satisfaction at layer).** Let a graph  $G$  and a condition  $c$  in UANF be given. A morphism  $p : C_0 \hookrightarrow G$  satisfies  $c$  at layer  $0 \leq k \leq \text{nl}(c)$ , denoted by  $p \models_k c$ , if

$$p \models \text{cut}_k(c).$$

A graph  $G$  satisfies a constraint  $c$  at layer  $0 \leq k \leq \text{nl}(c)$ , denoted by  $G \models_k c$ , if  $q : \emptyset \hookrightarrow G$  satisfies  $\text{cut}_k(c)$ . The biggest  $0 \leq k \leq \text{nl}(c)$  such that  $G \models_k c$  and no  $k < j \leq \text{nl}(c)$  with  $G \models_j c$  exists is denoted by  $\text{k}_{\max}(c, G)$ . We use the abbreviation  $\text{k}_{\max}$  when  $c$  and  $G$  are clear from the context.

Note that, given a graph  $G$  and a constraint  $c$ ,  $\text{k}_{\max}$  always exists, since the nesting level of  $c$  is finite and  $\text{cut}_0(c) = \text{true}$  is satisfied by every graph. Also, if  $p \models_{\text{nl}(c)} c$  it follows immediately that  $p \models c$ .

**Example 4.4.** Consider the graph  $G$  given in figure 3 and the constraint  $c_2$  given in figure 2. This graph does not satisfy  $c_2$ , since both occurrences of **Class** do not satisfy  $\exists C_2^2 \forall C_3^2 \exists C_4^2$ , but is does satisfy  $\text{cut}_1(c_2)$  and therefore

$$G \models_1 c_2 \text{ and } \text{k}_{\max} = 1$$

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 4.8.** Let a graph  $G$ , a condition  $c$  in UANF and a morphism  $p : C_0 \hookrightarrow G$  with  $p \models_k c$  be given. If the subcondition  $d = \text{sub}_k(c)$  is universally bound, then for any condition  $f$  over  $C_{k+1}$  it holds that

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

*Proof.* Let  $0 \leq k \leq \text{nl}(c)$  be the smallest number with  $\text{sub}_k(c) = \forall(a_k : C_k \hookrightarrow C_{k+1}, d)$  being universally bound and  $p \models_k c$ . Let  $q : C_k \hookrightarrow G$  be a morphism such that  $q \models \forall(a_k : C_k \hookrightarrow C_{k+1}, \text{false})$ . 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  $q' : C_{k+1} \hookrightarrow G$  with  $q = q' \circ a_k$ . Hence, for every condition  $f$  over  $C_{k+1}$  a morphism  $q' : C_{k+1} \hookrightarrow G$  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  $p \models \text{rep}_{k+1}(c, f)$ .

It follows that for every even  $k < j \leq \text{nl}(c)$ , such that  $p \models_j c$ , and every condition  $d$  over  $C_{j+1}$  it holds that  $p \models \text{rep}_{k+1}(c, f)$  with  $f = \text{rep}_{j-k+1}(\text{sub}_{k+1}(c), d)$ . Since  $\text{rep}_{k+1}(c, f) = \text{rep}_{j+1}(c, d)$  it follows that  $p \models \text{rep}_{j+1}(c, d)$ .  $\square$

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 4.9.** Let  $G$  be a graph,  $p : C_0 \hookrightarrow G$  a morphism and  $c$  a condition over  $C_0$  in UANF with  $p \models_k c$ . If  $\text{sub}_k(c)$  is universally bound,

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

*Proof.* Follows immediately by using lemma 4.8 and setting  $f$  to  $\text{sub}_{k+1}(c)$ .  $\square$

Additionally, with lemma 4.8 it can be shown that given an graph  $G$ , a condition in UANF over  $C_0$  and a morphism  $p : C_0 \hookrightarrow G$  such that  $p \models_k c$  and  $\text{sub}_k(c)$  is universally bound,  $p \models_j c$  for all  $k \leq j \leq \text{nl}(c)$ .

**Lemma 4.10.** *Let a graph  $G$ , a morphism  $p : C_0 \hookrightarrow G$  and a constraint  $c$  in UANF be given. Then,  $p \models_j c$  if  $j < k_{\max}$  and  $\text{sub}_j(c)$  is existentially bound.*

*Proof.* If a  $j < k_{\max}$  with  $p \models_j c$  exists such that  $\text{sub}_j(c)$  is universally bound, let  $j_1$  be the smallest of these. With lemma 4.8 follows that  $p \models_{j_2} c$  for all  $j_1 \leq j_2 < \text{nl}(c)$ .

Let  $\ell < j_1$ , such that  $\text{sub}_\ell(c)$  is existentially bound and let  $d = \text{cut}_{j_1-\ell}(\text{sub}_\ell(c)) = \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, e)$  be the condition up to layer  $j_1 - \ell$  of  $\text{sub}_\ell(c)$ . 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  $q = q' \circ a_\ell$  must exists. It follows that  $q \models \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, \text{true})$  and with that  $p \models_\ell c$ .  $\square$

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

The construction of intermediate 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  $C' \in \text{IG}(C_0, C_1)$ . 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 4.11 (Intermediate condition).** *Let a constraint  $c$  in UANF be given. Let  $0 \leq k < \text{nl}(c)$  such that either  $\text{sub}_k(c) = \exists(a_k : C_k \hookrightarrow C_{k+1}, d)$  is existentially bound or  $k = \text{nl}(c) - 1$ . The intermediate condition,  $\text{IC}_k(c, C')$ , of  $c$  at layer  $k$  with*

$$C' \in \begin{cases} \text{IG}(C_k, C_{k+1}) & \text{if } \text{sub}_k(c) \text{ is existentially bound} \\ \{C_{\text{nl}(c)}\} & \text{otherwise} \end{cases}$$

is defined as:

$$\text{IC}_k(c, C') := \begin{cases} \text{rep}_k(c, \exists(a_k^r : C_k \hookrightarrow C', \text{true})) & \text{if } \text{sub}_k(c) \text{ is existentially bound} \\ c & \text{otherwise} \end{cases}$$

**Example 4.5.** Consider constraint  $c_1$  given in figure 2. Since  $C_2^2 \in \text{IG}(C_1^1, C_1^2)$ , we can construct a intermediate condition of  $c_1$  at layer 1 with  $C_2^2$  as  $\text{IC}_1(c_1, C_2^2) = \forall C_1^1 \exists C_2^2$ . Whereas  $c_1$  checks whether each node of the type **Class** is connected to at least two nodes of the type **Feature**, the partial condition checks whether each nodes of the type **Class** is connected to at least one nodes of the type **Feature** which is trivially satisfied if  $c_1$  is satisfied.

If a graph  $G$  does not satisfy the condition up to layer  $k := k_{\max} + 2$ ,  $k < \text{nl}(c)$ , of a given constraint  $c$ , with  $\text{sub}_k(c)$  being existentially bound, there does exists a graph  $C' \in \text{IG}(C_k, C_{k+1})$ , such that  $G$  satisfies  $\text{IC}_k(c, C')$  since  $G$  always satisfies  $\text{IC}_k(c, C_k)$ .

### 4.3 Consistency increasing transformations and rules

With the results above, we are now ready to define the notions of *consistency increase-*ment and *direct consistency increase-*ment, 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  $k = k_{\max} + 1$  and  $\text{sub}_k(c) := \forall(a_k : C_k \hookrightarrow C_{k+1}, \exists(a_{k+1} : C_{k+1} \hookrightarrow C_{k+2}, d))$ . A transformation  $t : G \implies H$  is considered as (direct) consistency increasing if  $k_{\max}(c, G) \leq k_{\max}(c, H)$ , i.e. the satisfaction up to layer 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  $\exists(a_{k+1} : C_{k+1} \hookrightarrow C_{k+2}, \text{true})$ , 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  $p \not\models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C', \text{true})$  and  $\text{tr}_t op \models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C', \text{true})$  for a graph  $C' \in \text{IG}(C_{k+1}, C_{k+2})$ . A decreasing insertion is the creation of an occurrence of  $C_{k+1}$  not satisfying  $\exists(a_{k+1} : C_{k+1} \hookrightarrow C_{k+2}, \text{true})$  and a decreasing deletion is the deletion of elements of  $C_{k+2}$  such that for an occurrence  $p$  of  $C_{k+1}$  with  $p \models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C', \text{true})$  it holds that  $\text{tr}_t op \not\models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C', \text{true})$  for a graph  $C' \in \text{IG}(C_{k+1}, C_{k+2})$ .

To evaluate this, we define the *number of violations*. Intuitively, for all occurrences  $p$  of  $C_{k+1}$  the number of graphs  $C' \in \text{IG}(C_{k+1}, C_{k+2})$  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  $k \leq k_{\max}$  it is set to 0 and for all layer  $k$  with  $k > k_{\max} + 1$  it is set to  $\infty$ . Through this, a transformation  $tG \implies H$  that increases the satisfaction up to layer can easily detected since the number of violations in  $H$  at layer  $c_{\max}^G + 1$  will be set to 0.

**Definition 4.12 (number of violations).** Let a graph  $G$  and a constraint  $c$  in UANF be given. The number of violations  $\text{nvc}_j(G)$  at layer  $0 \leq j < \text{nl}(c)$  in  $G$  is defined as:

1. If  $j < k_{\max} + 1$ :

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

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

$$\text{nvc}_j(G) := \begin{cases} \sum_{C' \in \text{IG}(C_{j+2}, C_{j+3})} |\{q \mid q : C_{j+1} \hookrightarrow G \wedge q \not\models \text{IC}_0(e, C')\}| & \text{if } e \neq \text{false} \\ |\{q \mid q : C_{j+1} \hookrightarrow G\}| & \text{if } e = \text{false} \end{cases}$$

3. If  $j > k_{\max} + 1$ :

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

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 4.13 (consistency increasing transformations and rules).** Let a graph  $G$ , a rule  $\rho$  and a constraint  $c$  in UANF 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  $0 \leq k < \text{nl}(c)$ . A rule  $r$  is called consistency increasing w.r.t  $c$ , if all of its transformations 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  $0 \leq j < \text{nl}(c)$ . Also, no plain rule  $\rho$  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 satisfaction up to layer, which is shown by the following theorem.

**Theorem 4.1.** Let a graph  $G$ , a rule  $\rho$  and a constraint  $c$  in UANF be given. A transformation  $t : G \Rightarrow_{\rho, m} H$  is consistency increasing w.r.t  $c$  if  $k_{\max}(c, G) < k_{\max}(c, H)$ .

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

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 4.2.** Let a constraint  $c$  in UANF and a set of rules  $\mathcal{R}$  be given. Every sequence of consistency increasing transformations 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 consistency increasing transformations w.r.t  $c$  with  $\rho_i \in \mathcal{R}$ . We assume that  $k_{\max}(c, G_0) < nl(c)$ , otherwise  $nvc_j(G_0) = 0$  for all  $0 \leq j < nl(c)$  and no transformation  $G_0 \Rightarrow H$  is consistency increasing.

We show that after at most  $j := nvc_{k_{\max}(c, G_0)+1}(G_0)$  transformations  $G_j \models_{k_{\max}(c, G_0)+2} c$  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  $nvc_{k_{\max}(c, G_i)+1}(G_{i+1}) \leq nvc_{k_{\max}(c, G_i)+1}(G_i) - 1$ . Therefore, after at most  $j$  transformations,  $nvc_{k_{\max}(c, G_0)+1}(G_j) \leq nvc_{k_{\max}(c, G_0)+1}(G_0) - j = 0$  and  $G_j \models_{k_{\max}(c, G_0)+2} c$ . 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  $G_k \Rightarrow G_{k+1}$  exists, the sequence has to be finite. Also, for some  $G_{k'}$  with  $k' < k$  a consistency increasing transformation  $G_{k'} \Rightarrow_{\rho} H$  with  $\rho \in \mathcal{R}$  may not exist and the sequence up to  $G_{k'}$  is also finite, but does not end with a graph satisfying  $c$ .  $\square$

The replacement of a graph not satisfying a constraint  $c$  by a  $c$  satisfying graph via a transformation is a consistency increasing transformation. Therefore, the notion of consistency increasement does allow insertions or deletions that are unnecessary in order to increase consistency. That is, the deletion of occurrences of existentially bound graphs, the deletions of occurrences of universally bound graphs  $C_k$  satisfying  $\exists C_{k+1}$  or the insertion of occurrences of universally bound graphs and the insertion of existentially bound graphs  $C_k$ , such that the corresponding occurrence of  $C_{k-1}$  already satisfied  $\exists C_k$ .

*Direct consistency increasing* transformations are more restricted, in the sense, that these unnecessary deletions and insertions are leading to a transformation not being direct consistency increasing. The presence of these unnecessary actions can be checked via second-order logic formulas. Additionally, it is secured that no new violations are introduced since these can always be considered as an unnecessary insertion or deletion. With that, the removal of one violation is sufficient to state that the transformation is (direct) consistency increasing, which can also be checked via a second order logic formula. Intuitively, it is checked by the first two formulas, that no new violations are introduced. The third formula secures that at least one violation has been removed and the last formulas secures that the satisfaction up to layer is not decreased.

**Definition 4.14 (direct consistency increasing).** Let  $G$  be a graph,  $\rho$  a plain rule,  $c$  a constraint in UANF and  $\text{sub}_{k_{\max}+1}(c) = \forall(a_{k-1} : C_{k-1} \rightarrow C_k, e)$ . A transformation  $t : G \Rightarrow_{\rho, m} H$  is called direct consistency increasing w.r.t  $c$  if the following equations hold. Let

$$\mathbf{G} = \begin{cases} \text{IG}(C_k, C_{k+1}) & \text{if } e \neq \text{false} \\ \{C_k\} & \text{otherwise} \end{cases}$$

Every occurrence of  $C_k$  in  $G$  that satisfies  $\text{IC}_0(e, C')$  for any  $C' \in \mathbf{G}$  still satisfies

$\text{IC}_0(e, C')$  in  $H$ .

$$\forall p : C_k \hookrightarrow G \left( \bigwedge_{C' \in \mathbf{G}} (p \models \text{IC}_0(e, C') \wedge \text{tr}_t \circ p \text{ is total}) \implies \text{tr}_t \circ p \models \text{IC}_0(e, C') \right) \quad (4.1)$$

Let  $C' = C_{k+1}$  if  $e \neq \text{false}$  and  $C' = C_k$  otherwise. Every new inserted occurrence of  $C_k$  by  $t$  satisfies  $\text{IC}_0(e, C')$ .

$$\forall p' : C_k \hookrightarrow H (\neg \exists p : C_k \hookrightarrow G(p' = \text{tr}_t \circ p) \implies p' \models \text{IC}_0(e, C')) \quad (4.2)$$

At least one occurrence of  $C_k$  in  $G$  that does not satisfy  $\text{IC}_0(e, C')$ , for any  $C' \in \mathbf{G}$ , either has been removed or satisfies  $\text{IC}_0(e, C')$  in  $H$ .

$$\exists p : C_k \hookrightarrow G \left( \bigvee_{C' \in \mathbf{G}} (p \not\models \text{IC}_0(e, C') \wedge (\text{tr}_t \circ p \text{ is not total} \vee \text{tr}_t \circ p \models \text{IC}_0(e, C'))) \right) \quad (4.3)$$

No occurrence of a universally bound graph  $C_j$  with  $j < k_{\max}$  gets inserted.

$$\bigwedge_{\substack{i < k_{\max} \\ i \text{ even}}} \forall p : C_i \hookrightarrow H (\exists p' : C_i \hookrightarrow G(p = \text{tr}_t \circ p')) \quad (4.4)$$

No occurrence of an existentially bound graph  $C_j$  with  $j \leq k_{\max}$  gets deleted.

$$\bigwedge_{\substack{i \leq k_{\max} \\ i \text{ odd}}} \forall p : C_i \hookrightarrow G (\text{tr}_t \circ p \text{ is total}) \quad (4.5)$$

Note that (4.4) and (4.5) not only secure that the satisfaction up to layer does not decrease, as shown in the following lemma, but also prevent further unnecessary insertions and deletions, since the insertion of a universally and the deletion of a existentially bound graph will never lead to an increase of consistency.

**Lemma 4.15.** *Let a transformation  $t : G \implies H$  and a constraint  $c$  in UANF be given, such that (4.4) and (4.5) of definition 4.14 are satisfied. Then,*

$$H \models_{k_{\max}(c, G)} c.$$

*Proof.* Assume that  $H \not\models_{k_{\max}(c, G)} c$ . Then, either a new occurrence of an universally bound graph  $C_k$  with  $k < k_{\max}(c, G)$  has been inserted or an occurrence of an existentially bound graph  $C_k$  with  $k \leq k_{\max}(c, G)$  has been destroyed. Therefore, the following holds:

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

with  $i, j \leq k_{\max}(c, G)$ ,  $i$  being even and  $j$  being odd. It follows immediately that either (4.4) or (4.5) is not satisfied. This is a contradiction.  $\square$

Now, we will show the already indicated relationship between direct consistency increasing and consistency increasing, namely that a direct consistency increasing transformation is also consistency increasing. Counterexamples, that the inversion of the implication does not hold can easily be constructed, showing that these notions are not identical, but related.

**Lemma 4.16.** *Let a graph  $G$ , a constraint  $c$  in UANF, a rule  $\rho$  and a direct consistency increasing transformation  $t : G \Rightarrow_{\rho, m} H$  w.r.t.  $c$  be given. Then,  $t$  is also a consistency increasing transformation.*

*Proof.* Let  $k = \text{k}_{\max}(c, G) + 2$  and  $d = \text{sub}_k(c)$  with  $d \neq \text{false}$ . With lemma 4.15 follows that  $\text{k}_{\max}(c, G) \leq \text{k}_{\max}(c, H)$  and with that  $\text{nvc}_k(H) \neq \infty$ . Let

1. We show that equations (4.1) and (4.2) 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  $p \not\models \text{IC}_0(d, C')$  for any  $C' \in \text{IG}(C_k, C_{k+1})$  exists, such that either 1a or 1b is satisfied.
  - (a) There does exist a morphism  $q' : C_k \hookrightarrow G$  with  $q' \models \text{IC}_0(d, C')$  and  $p = \text{tr}_t \circ q'$ .
  - (b) There does not exist a morphism  $q : C_k \hookrightarrow G$  with  $p = \text{tr}_t \circ q$ .

This is a contradiction, if 1a is satisfied,  $q'$  does not satisfy equation (4.1) and if 1b is satisfied  $q$  does not satisfy equation (4.2). It follows that

$$|\{q \mid q : C_k \hookrightarrow G \wedge q \not\models \text{IC}_0(d, C')\}| \leq |\{q \mid q : C_k \hookrightarrow H \wedge q \not\models \text{IC}_0(d, C')\}|$$

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

2. Since (4.3) is satisfied, a morphism  $p : C_k \hookrightarrow G$  with  $p \not\models \text{IC}_0(d, C')$ , such that either  $\text{tr } op$  is total and  $\text{tr}_t \circ p \models \text{IC}_0(d, C')$  or  $\text{tr } op$  is not total exists, for a graph  $C' \in \text{IG}(C_k, C_{k+1})$ . In both cases the following holds:

$$\begin{aligned} p \in \{q \mid q : C_k \hookrightarrow G \wedge q \not\models \text{IC}_0(d, C')\} \wedge \\ \text{tr } op \notin \{q \mid q : C_k \hookrightarrow H \wedge q \not\models \text{IC}_0(d, C')\} \end{aligned}$$

With that and 1. it follows that

$$|\{q \mid q : C_k \hookrightarrow G \wedge q \not\models \text{IC}_0(d, C')\}| < |\{q \mid q : C_k \hookrightarrow H \wedge q \not\models \text{IC}_0(d, C')\}|.$$

In total,  $\text{nvc}_k(G) < \text{nvc}_k(H)$  and  $t$  is a consistency increasing transformation.

For the special case that  $k = \text{nl}(c)$  and  $c$  ends with  $\forall(a_k : C_{k-1} \hookrightarrow C_k, \text{false})$  it can be shown in a similar way that (4.1) and (4.2) imply that

$$|\{q \mid q : C_k \hookrightarrow G\}| \leq |\{q \mid q : C_k \hookrightarrow H\}|$$

and that (4.3) implies

$$|\{q \mid q : C_k \hookrightarrow G\}| < |\{q \mid q : C_k \hookrightarrow H\}|.$$

It follows that  $t$  is a consistency increasing transformation.  $\square$

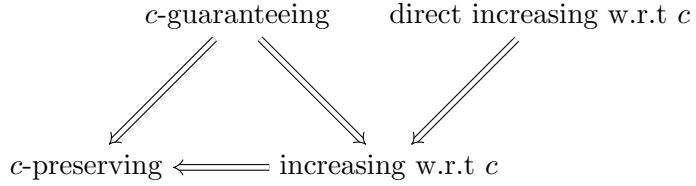


Figure 4: General relations of consistency notions.

#### 4.4 Comparison with other consistency concepts

In this chapter, the notions of (direct) consistency increasing are compared to the already known concepts of consistency guaranteeing, consistency preserving [2], (direct) consistency increasing and sustaining [3] in order to reveal relationships between them and secure that (direct) consistency increasing is indeed a new notion of consistency. These relations are displayed in figure 4.

First, we compare (direct) consistency increasing with the notions of consistency-guaranteeing and -preserving. As already mentioned, given a constraint  $c$  and a graph  $G$  with  $G \models c$ , there does not exist a (direct) consistency increasing transformation  $t : G \Rightarrow H$  and it follows immediately that (direct) increasing w.r.t  $c$  implies  $c$ -preserving.

Obviously, guaranteeing implies consistency increasing, since this property is embedded in the definition of consistency increasing. The inversion of this implication does not hold, since increasing is a way stricter notion, in the sense, that the number of removed violations has to be greater than the number of introduced violations. For guaranteeing transformations this is not the case, an arbitrary number of violations can be inserted, long as the derived graph satisfies the constraint and therefore guaranteeing does not imply direct increasing, since a direct increasing transformations is not allowed to introduce any new violations.

**Lemma 4.17.** *Let a constraint  $c$  in UANF, graphs  $G$  and  $H$  with  $G \not\models c$ , and a transformation  $t : G \Rightarrow H$  be given. Then,*

$$\begin{aligned} t \text{ is } c\text{-guaranteeing} &\implies t \text{ is increasing w.r.t } c && \wedge \\ t \text{ is } c\text{-guaranteeing} &\nrightarrow t \text{ is direct increasing w.r.t } c \wedge \\ t \text{ is increasing w.r.t } c &\nrightarrow t \text{ is } c\text{-guaranteeing}. \end{aligned}$$

*Proof.* 1. Let  $t$  be a  $c$ -guaranteeing transformation, then  $H \models c$ . Since  $G \not\models c$ ,  $\text{nvc}_{k_{\max}(c,G)+1}(G) > 0$  and  $\text{nvc}_{k_{\max}(c,G)+1}(H) = 0$ . Therefore,  $t$  is consistency increasing.

2. Let a constraint  $c = \exists(a_0 : \emptyset \hookrightarrow C_1, \forall(a_1 : C_1 \hookrightarrow C_2, \exists(a_2 : C_2 \hookrightarrow C_3, \text{true})))$  and a graph  $G = C' \dot{\cup} C'$  with  $C' = C_3 \setminus \{e\}$  for one edge  $e \in E_{C_3 \setminus C_2}$  and the occurrences  $p_1 : C_3 \hookrightarrow G$  and  $p_2 : C_2 \hookrightarrow G$  be given. Let  $t : G \Rightarrow H$  be a transformation with  $H = C_3 \dot{\cup} C''$  and  $C'' = C' \setminus \{e'\}$  for an  $e' \in E_{C_3 \setminus C_2}$ , 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 increasing, since  $p_i \models$

$\exists(a'_2 : C_2 \hookrightarrow C', \text{true})$  for  $i = 1, 2$  and either  $\text{tr}_t \circ p_1 \not\models \exists(a'_2 : C_2 \hookrightarrow C', \text{true})$  or  $\text{tr}_t \circ p_2 \not\models \exists(a'_2 : C_2 \hookrightarrow C', \text{true})$ .

3. Let  $t$  be a consistency increasing transformation w.r.t  $c$ , such that  $\text{nvc}_{k_{\max}(c,G)+1}(G) > \text{nvc}_{k_{\max}(c,G)+1}(H) > 1$  and  $H \not\models_{k_{\max}(c,G)+2} c$ . Then,  $H \not\models c$  and  $t$  is not a  $c$ -guaranteeing transformation.

□

The definition of consistency improving only differs from guaranteeing if the corresponding constraint is universally bound and these notions are identical for existentially bound constraint. Therefore, with lemma 4.17, we can state the following.

**Corollary 4.18.** *Let  $c$  be an existentially bound constraint in ANF. Then,*

$$\begin{aligned} t \text{ is consistency improving w.r.t } c &\implies t \text{ is consistency increasing w.r.t } c \wedge \\ t \text{ is consistency increasing w.r.t } c &\not\implies t \text{ is consistency improving w.r.t } c \end{aligned}$$

The notions of increasing and improving are equivalent for universally bound constraints with nesting level 1. Note that, with corollary 4.18, this equivalence does not hold for existentially bound constraints with nesting level 1.

**Lemma 4.19.** *let a universally bound constraint  $c$  in UANF with  $\text{nl}(c) = 1$ , a graph  $G$  with  $G \not\models c$  and a transformation  $t : G \Rightarrow H$  be given. Then,*

$$t \text{ is consistency improving w.r.t } c \iff t \text{ is consistency increasing w.r.t } c$$

*Proof.* Let  $c = \forall(a : \emptyset \hookrightarrow C, \text{false})$ . Since  $\text{sub}_1(c) = \text{false}$ ,  $\text{nvc}_1(G)$  is the number of occurrences of  $C$  in  $G$ . This is exactly the definition of the number of violations for consistency improving transformations and the statement follows immediately. □

For universally bound constraints  $c$  with  $\text{nl}(c) \geq 2$ , (direct) consistency increasing is not related to (direct) consistency improving and sustaining. As shown in [3], (direct) consistency improving implies (direct) consistency sustaining. Therefore, it is sufficient to show that direct improving does not imply increasing and that direct increasing does not imply consistency sustainment.

**Lemma 4.20.** *Let a universally bound constraint  $c$  in UANF with  $\text{nl}(c) \geq 2$ , a graph  $G$  with  $G \not\models c$  and a transformation  $t : G \Rightarrow H$  be given. Then,*

$$\begin{aligned} t \text{ is direct consistency improving w.r.t } c &\not\implies t \text{ is consistency increasing w.r.t } c \wedge \\ t \text{ is direct increasing w.r.t } c &\not\implies t \text{ is consistency sustaining w.r.t } c \end{aligned}$$

*Proof.* 1. Let  $c = \forall(a_0 : \emptyset \hookrightarrow C_1, \exists(a_1 : C_1 \hookrightarrow C_2, \text{true}))$  be a constraint. Let  $V_{C_1} = V_{C_2}$  and  $|E_{C_2}| - |E_{C_1}| = 2$ . Let  $G = C' \dot{\cup} C''$  with  $C' = C_2 \setminus \{e\}$  with  $e \in E_{C_2} \setminus E_{C_1}$  and the occurrences  $p_1 : C_1 \hookrightarrow G$  and  $p_2 : C_1 \hookrightarrow G$  be given. It follows that  $\text{nvc}_0(G) = 2$ . Let  $t : G \Rightarrow H$  be a transformation, such that  $H = C_1$ . Then,  $t$  is a direct consistency improving transformation, since  $H$  contains only one occurrence of  $C_1$  not satisfying  $\exists(a_1 : C_1 \hookrightarrow C_2 \text{ true})$ . But,  $t$  is not consistency increasing, since  $\text{nvc}_0(H) = 3$ .

- ~~delete~~
- insert

move Feature :



assign Feature :



Figure 5: rules

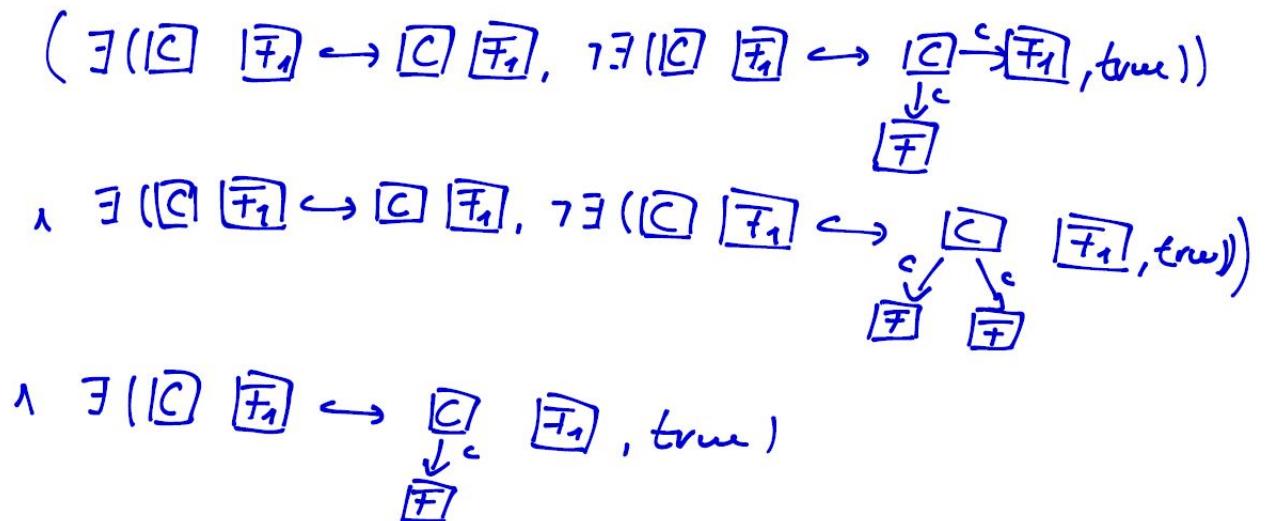


Figure 6: application condition

2. Let  $c := \forall(a_0 : \emptyset \hookrightarrow C_1, \exists(a_1 : C_1 \hookrightarrow C_2, \forall(a_2 : C_2 \hookrightarrow C_3, \exists(a_3 : C_3 \hookrightarrow C_4, \text{true}))))$  be a constraint.

Let a graph  $G = C_1$  with the morphism  $q : C_1 \hookrightarrow G$  and a transformation  $t : G \Rightarrow H$  with  $H := C_3 \dot{\cup} C_3$  be given, such that  $\text{tr}_t \circ q$  is total. Then  $t$  is a direct consistency increasing transformation but not a consistency sustaining one, since  $H$  contains more occurrences of  $C_1$  not satisfying  $\exists(a_1 : C_1 \hookrightarrow C_2, \forall(a_2 : C_2 \hookrightarrow C_3, \exists(a_3 : C_3 \hookrightarrow C_4, \text{true})))$  than  $G$ .

□

## 5 general application conditions

To guarantee that each transformation  $t : G \implies_{\rho, m} H$ , given a rule  $\rho = L \leftarrow K \hookrightarrow R$ , is (direct) consistency increasing w.r.t to a constraint  $c$ , we present applications conditions ensuring this property. It has to be ensured that at least one violation will be removed. For this, it is necessary (a) to check that an occurrence  $p$  of the universally bound graph  $C_{k_{\max}(c, G)+1}$  exists, such that  $p$  and the match  $m$  do overlap, i.e  $p(C_{k_{\max}(c, G)+1}) \cap m(L) \neq \emptyset$ , and (b) that  $p$  does not satisfy  $c' = \exists C_{k_{\max}(c, G)+2}$ . Only in this case, it is possible that the transformation does remove a violation. To ensure that  $p$  does not satisfy  $c'$ , the non-existence of all possible overlaps of  $L$  and  $C_{k_{\max}(c, G)+2}$  such that  $p \models c'$  has to be checked. For this, we introduce *extended overlaps*. Intuitively, given an overlap  $C$  of  $L$  and  $C_{k_{\max}(c, G)+1}$  with  $p : C_{k_{\max}(c, G)+1} \hookrightarrow C$ , the set of extended overlaps contains all overlaps  $C'$  of  $L$  and  $C_{k_{\max}(c, G)+2}$ , such that  $p \models c'$ .

**Definition 5.1 (extended overlaps).** Let  $G$ ,  $C_0$  and  $C_1$  with  $i_{C_0}^{C_1} : C_0 \hookrightarrow C_1$  be graphs. Let  $C$  be an overlap of  $C_0$  and  $G$  with the inclusion  $i_{C_0}^C : C_0 \hookrightarrow C$ . The set of extended overlaps of  $C$  with  $i_{C_0}^{C_1}$ ,  $\text{eol}(C, i_{C_0}^{C_1})$ , is the set of all overlaps  $C'$  of  $G$  and  $C_1$ , such that an injective morphism  $i_C^{C'} : C \hookrightarrow C'$  with  $i_C^{C'} \circ i_{C_0}^C \models \exists(i_{C_0}^{C_1} : C_0 \hookrightarrow C_1, \text{true})$  exists.

A direct consistency increasing application condition has to, as mentioned above, (a) ensure that at least one violation will be removed, (b) ensure that no new violations get inserted and (c) secure that the satisfaction up to layer is not decreased. In the definition below, `exv()` and `remv()` ensure that (a) is met, with `exv()` ensuring that a violation is present and `remv()` ensuring that this violation will be removed. Note, that the construction of these is divided in two cases. Firstly, either  $k < \text{nl}(c) - 2$  or the constraints ends with a condition of the form  $\exists(a : C \hookrightarrow C', \text{true})$  and secondly,  $k = \text{nl}(c) - 2$  and the constraint ends with a condition of the form  $\forall(a : C_0 \hookrightarrow C_1, \text{false})$ .

For `exv()`, the first case will be checked as already described above. In the second case, it is sufficient to check whether an occurrence  $p$  of  $C_k$  with  $m(L) \cap p(C_k) \neq \emptyset$  exists.

For `remv()`, in the first case, a violation can be removed by either deleting an occurrence  $p$  of  $C_k$  or inserting elements of  $C_{k+1}$ , such that  $p \not\models \exists C'$  and  $\text{tr}_t op \models \exists C'$  for a graph  $C' \in \text{IG}(C_k, C_{k+1})$ . In the second case, a violation can only be removed by deleting an occurrence  $p$  of  $C_1$ . This will only occur if  $m(L \setminus K) \cap p(C_1) \neq \emptyset$ .

The conditions constructed by `nui()` and `ndis()` ensure that (2) is met, with `nui()` checking that no new occurrence of  $C_k$  will be inserted and `ndis()` ensuring that for an occurrence  $p$  of  $C_k$  if  $p \models \exists C'$  it is secured that  $\text{tr}_t op \models \exists C'$  for any  $C' \in \text{IG}(C_k, C_{k+1})$ .

Additionally, the conditions constructed by `ned()` and `nui()` ensure (3), with `ned()` ensuring that no occurrence of existentially bound graphs  $C_j$  with  $j \leq k$  will be deleted and `nui()` ensuring that no occurrences of universally bound graphs  $C_j$  with  $j \leq k$  will be inserted. With this, it is guaranteed that the satisfaction up to layer will not be decreased.

**Definition 5.2 (general application condition).** Let a rule  $\rho = L \leftarrow K \hookrightarrow R$  and a constraint  $c$  in UANF be given. Let  $d = \text{sub}_k(c) = \forall(a_k : C_k \hookrightarrow C_{k+1}, e)$  with  $0 \leq k < \text{nl}(c)$  and  $k$  being even.

The application condition  $\text{ap}_k(C')$  of  $c$  at layer  $k$  with  $C' = C_{k+1}$ , if  $e = \text{false}$ , and  $C' \in \text{IG}(C_{k+1}, C_{k+2})$  otherwise is defined as:

$$\text{ap}_k(C') := \left( \bigvee_{P \in \text{ol}(L, C_{k+1})} \text{exv}(P, C') \wedge \text{remv}(P, C') \right) \wedge \text{nui}() \wedge \text{ned}() \wedge \text{nds}() \quad (5.1)$$

with

1. Let  $a^r : C_{k+1} \hookrightarrow C'$  be a restricted morphism of  $a_{k+1}$  and  $i_L^P$  and  $i_P^Q$  the inclusions of  $L$  in  $P$  and  $P$  in  $Q$ , respectively.

$$\text{exv}(P, C') := \begin{cases} \exists(i_L^P : L \hookrightarrow P, \text{true}) & \text{if } e = \text{false} \\ \bigwedge_{Q \in \text{ol}(P, a^r)} \exists(i_L^P : L \hookrightarrow P, \neg\exists(i_P^Q : P \hookrightarrow Q, \text{true})) & \text{otherwise} \end{cases}$$

2. If  $i_L^P(L \setminus K) \cap i_{C_{k+1}}^P(C_{k+1}) \neq \emptyset$ , we set

$$\text{remv}(P, C') := \text{true}.$$

Otherwise, let  $P'$  be the graph derived by the transformation  $P \Rightarrow_{\rho, m} P'$ . Then,  $P'$  is an overlap of  $R$  and  $C_k$ . If this transformation does not exist, we set  $\text{remv}(P, C') := \text{false}$ . Let  $a^r : C_{k+1} \hookrightarrow C'$  be a partial morphism of  $a_{k+1}$ , then

$$\text{remv}(P, C') := \begin{cases} \text{false} & \text{if } e = \text{false} \\ \bigvee_{Q \in \text{ol}(P', a^r)} \text{Left}(\exists(i_R^Q : R \hookrightarrow Q, \text{true}), \rho) & \text{otherwise} \end{cases}$$

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

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

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

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

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  $C'' \in \text{IG}(C_{k+1}, C_{k+2})$ ,  $i_{C_k}^{P'} \models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C'', \text{true})$  and  $i_L^{P'}(L \setminus K) \cap i_{C''}^{P'}(C'' \setminus C_k) \neq \emptyset$ .

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

Note that  $\text{ap}_k(C')$ , for any  $C' \in \text{IG}(C_{k+1}, C_{k+2})$ , will only be evaluated with `true` if an occurrence  $p$  of  $C_{k+1}$  with  $p \not\models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C', \text{true})$  and  $\text{tr}_t op \models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C', \text{true})$  exists. For any smaller improvements, i.e a similar improvement for a subgraph  $C'' \in \text{IG}(C_{k+1}, C_{k+2})$  of  $C'$ ,  $\text{ap}_k(C')$  would be evaluated with `false`. For any bigger improvements, i.e the same improvement for a supergraph  $C'' \in \text{IG}(C_{k+1}, C_{k+2})$  of  $C'$ ,  $\text{ap}_k(C')$  would also be evaluated with `false`, if  $p \models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C', \text{true})$ . In both cases, the application condition would prohibit the transformation, even if it would be consistency increasing. To resolve this problem, multiple application conditions could be combined by

$$\bigvee_{C' \in \text{IG}(C_{k+1}, C_{k+2})} \text{ap}_k(C').$$

This application condition will be evaluated with `true` if the cases described above do appear, with the drawback that this leads to a huge condition, even if duplicate conditions are removed. At least all duplicates of `ned()`, `nui()` and `nds()` can be removed, since they are identical for each  $\text{ap}_k(C')$  and only need to be constructed once.

In general, these application conditions are a trade-off between conditions-size and restrictiveness. They are very restrictive, since they do not allow any deletions of occurrences of existentially bound and insertions of universally bound graphs. For example, any of these application conditions with the rule `moveFeature` and constraint  $c_1$  will be equivalent to `false`; `nds()` will always be evaluated with `false` since `moveFeature` does remove elements of the existentially bound graph  $C_2^1$ . A change of the conditions constructed by `nds()` such that it is checked whether two nodes of the type `Feature` are connected to a node `Class` will yield application conditions that are satisfiable with `moveFeature`, but for a similar rule moving two nodes of type `Feature`, this newly constructed `nds()` would still be evaluated with `false`. Therefore, this only leads to a slight decrease of restrictiveness.

The conditions constructed by `ned()` and `nui()` could be changed in a similar fashion. For `ned()` and the universally bound graph  $C_j$ , by checking whether there does exist an additional occurrence  $p$  of  $C_{j+1}$  such that  $p \models \text{sub}_{j+2}(\text{cut}_{k_{\max}}(c))$  and for `nui()`, by checking whether an introduced occurrence  $p$  of  $C_j$  does satisfy  $\text{sub}_{j+1}(\text{cut}_{k_{\max}}(c))$ . The construction of these is similar to the construction of consistency guaranteeing application conditions as introduced by Habel and Pennemann [2], which is known to construct huge application conditions. Also, they do get more and more restrictive for increasing  $k$ , since the number of conditions constructed by `ned()` and `nui()` also increases.

For constraints of the form  $c := \forall C_0 \exists C_1$  it can be shown that  $\text{ap}_0(C_2)$  is not only a direct consistency increasing, but also a direct consistency improving application condition.

**Example 5.1.** Consider the rule `assignFeature` of figure 5 and constraint  $c_1$  of figure 2. The application condition of  $c_1$  at layer 1 with  $C_2^1$  for `assignFeature` constructed by definition 5.2 is shown in figure 6. The first two rows are conditions constructed by `exv(·, ·)` and the third row is the condition constructed by `remv(·, ·)`. Note that `ned()`,

`nui()` and `nwo()` did not construct any conditions since `assignFeature` does not create elements of  $C_1^1$  and does not delete elements of  $C_2^1$ .

Additionally, this application condition is also a consistency improving application condition w.r.t  $c_2$ .

Let us now show that the construction above generates consistency increasing application conditions.

**Theorem 5.1.** *Let a graph  $G$ , a constraint  $c$  in UANF with  $G \not\models c$  and a plain rule  $\rho$  be given. Then,  $\rho' = (\rho, \text{ap}_{k_{\max}+1}(C))$  with  $C \in \text{IG}(C_{k_{\max}+1}, C_{k_{\max}+2})$  is a consistency increasing rule.*

*Proof.* Let  $t : G \implies_{\rho', m} H$  be a transformation and  $k = k_{\max} + 1$ . We show, by contradiction, that this transformation is direct consistency increasing by showing that (4.1), (4.2), (4.3), (4.4) and (4.5) are satisfied. With that,  $\rho'$  is a consistency increasing rule. Let  $\text{sub}_k(c) = \forall(a_k : C_k \hookrightarrow C_{k+1}, e)$  be the subcondition of  $c$  at layer  $k$ . Let

$$\mathbf{G} := \begin{cases} \text{IG}(C_{k+1}, C_{k+2}) & \text{if } e \neq \text{false} \\ \{C_k\} & \text{otherwise.} \end{cases}$$

1. Assume that (4.1) does not hold. Then, a morphism  $p : C_{k+1} \hookrightarrow G$  exists, such that  $p \models \text{IC}_0(e, C')$ ,  $\text{tr}_t \circ p$  is total and  $\text{tr}_t \circ p \not\models \text{IC}_0(e, C')$  for a graph  $C' \in \mathbf{G}$ . Therefore, an overlap  $P$  of  $L$  and  $C'$  such that  $i_{C_{k+1}}^P \models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C', \text{true})$  with  $i_L^P(L \setminus K) \cap i_{C'}^P(C' \setminus C_{k+1}) \neq \emptyset$  exists and  $m \models \exists(i_L^P : L \hookrightarrow P, \text{true})$  holds. Thus, `nds()` and consequently also `ap` <sub>$k$</sub> ( $C$ ) cannot be satisfied.
2. Assume that (4.2) does not hold and let

$$d := \begin{cases} \text{IC}_0(e, C_{k+2}) & \text{if } e \neq \text{false} \\ \text{false} & \text{otherwise.} \end{cases}$$

Then, a morphism  $p' : C_{k+1} \hookrightarrow H$  with  $p' \not\models d$  exists, such that no morphism  $p : C_{k+1} \hookrightarrow G$  with  $\text{tr}_t \circ p = p'$  exists. Therefore, an overlap  $P$  of  $R$  and  $C_{k+1}$  with  $i_R^P(R \setminus K) \cap i_{C_{k+1}}^P(C_{k+1}) \neq \emptyset$  exists, such that  $m \models \text{Left}(\exists(i_R^P : R \hookrightarrow P, \text{true}), \rho)$  and  $m$  does not satisfy `nui()`.

3. Assume that (4.3) does not hold. Then, no morphism  $p : C_{k+1} \hookrightarrow G$  with  $p \not\models \text{IC}_0(e, C')$ , such that  $\text{tr}_t \circ p$  is not total or  $\text{tr}_t \circ p \models \text{IC}_0(e, C')$  exists, for any  $C' \in \mathbf{G}$ . Then, no overlap  $P$  of  $L$  and  $C_{k+1}$  with  $i_L^P(L \setminus K) \cap i_{C_{k+1}}^P(C_{k+1}) \neq \emptyset$  exists, such that  $m \models \text{exv}(P, C)$ .

Let  $P$  be an overlap of  $C_{k+1}$  and  $L$  with  $i_L^P(L \setminus K) \cap i_{C_{k+1}}^P(C_{k+1}) = \emptyset$ . If  $e = \text{false}$ , then `remv`( $P, C$ ) = `false`. Otherwise, if  $m \models \text{exv}(P, C) \wedge \text{remv}(P, C)$ , it holds that  $i_{C_{k+1}}^P \not\models \text{IC}_0(e, C)$  and a graph  $Q \in \text{eol}(P', a^r)$  exists, such that  $m \models \text{Left}(\exists(i_R^Q : R \hookrightarrow Q, \text{true}), \rho)$ . In this case,  $\text{tr}_t \circ i_{C_{k+1}}^P \models \text{IC}_0(e, C)$  follows and (4.3) is satisfied.

In total,  $m \not\models \text{exv}(P, C) \wedge \text{remv}(P, C)$  follows for all  $P \in \text{ol}(L, C_{k+1})$  and therefore  $m \not\models \text{ap}_k(C)$ .

4. Assume that (4.4) does not hold. Then, a morphism  $p : C_j \hookrightarrow H$  with  $C_j$  being universally bound and  $j < k$  exists, 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  $i_R^P(R \setminus K) \cap i_{C_j}^P(C_j) \neq \emptyset$  exists, such that  $m \models \text{Left}(\exists(i_R^P : R \hookrightarrow P, \text{true}), \rho)$ . Hence,  $m \not\models \text{nui}()$  and  $m \not\models \text{ap}_k(C)$ .
5. Assume that (4.5) does not hold. Then, a morphism  $p : C_j \hookrightarrow G$  with  $C_j$  being existentially bound and  $j < k$  exists, such that  $\text{tr}_t \circ p$  is not total. Then, an overlap  $P$  of  $C_j$  and  $L$  with  $i_L^P(L \setminus K) \cap i_{C_j}^P(C_j) \neq \emptyset$  exists, such that  $m \models \exists(i_L^P : L \hookrightarrow P, \text{true})$ . Hence,  $m \not\models \text{ned}()$  and  $m \not\models \text{ap}_k(C)$ .

In total follows, if  $\models \text{ap}_k(C)$ , then  $t$  is a direct consistency increasing transformation.  $\square$

## 5.1 Conflicts within and between conditions

**Definition 5.3 (Conflicts within condition).** Let a condition  $c$  in UANF be given. Then, an existentially bound graph  $C_k$  has a conflict with an universally bound graph  $C_j$  if a transformation  $t : G \xrightarrow{\rho} H$  with  $\rho = C_{k-1} \xleftarrow{\text{id}} C_{k-1} \xrightarrow{a_{k-1}} C_k$  exists such that the following holds

$$\exists p : C_j \hookrightarrow H (\neg \exists p' : C_j \hookrightarrow G (\text{tr}_t \circ p' = p)).$$

An universally bound graph  $C_k$  has a conflict with an existentially bound graph  $C_j$  if a transformation  $t : G \xrightarrow{\rho} H$  with  $\rho = C_k \xleftarrow{i_{C'}^l} C' \xrightarrow{\text{id}} C'$  and  $C' \in \text{IG}(C_{k-1}, C_k)$  exists such that the following holds.

$$\exists p : C_j \hookrightarrow G (\text{tr}_t \circ p \text{ is not total}).$$

A condition  $c$  in UANF is called conflict free if no graph  $C_k$  in  $c$  has a conflict with a graph  $C_j$  with  $0 \leq j < k \leq \text{nl}(c)$ . A graph  $C_k$  has a transitive conflict with  $C_{j'}$  a graph  $C_j$  exists such that  $C_k$  has a conflict with  $C_j$  and  $C_j$  has a (transitive) conflict with  $C_{j'}$ . A condition  $c$  does contain a circular conflict if a graph  $C_k$  exists such that  $C_k$  has a transitive conflict with itself. Otherwise,  $c$  is called circular conflict free.

**Lemma 5.4.** Let a constraint  $c$  in UANF be given.

1. Let  $C_k$  be an existentially bound graph of  $c$ . Then,  $C_k$  has a conflict with  $C_j$  if and only if an overlap  $P$  of  $C_k$  and  $C_j$  exists such that

$$i_{C_k}^P(C_k \setminus C_{k-1}) \cap i_{C_j}^P(C_j) \neq \emptyset$$

and the rule  $\rho = C_k \xleftarrow{l} C_{k-1} \xrightarrow{r} C_{k-1}$  is applicable at match  $i_{C_k}^P$ .

2. Let  $C_k$  be an universally bound graph of  $c$ . Then  $C_k$  has a conflict with  $C_j$  if and only if an overlap  $P$  of  $C_k$  and  $C_j$  exists such that

$$i_{C_k}^P(C_k \setminus C_{k-1}) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$$

and each rule  $\rho = C_k \xleftarrow{i_{C'}^{C'}} C' \xrightarrow{\text{id}} C'$  with  $C' \in \text{IG}(C_{k-1}, C_k)$  is applicable at match  $i_{C_k}^P$ .

*Proof.* Let a condition  $c$  in UANF be given.

1. “ $\implies$ ”: Let  $C_k$  be an existentially bound graph that has a conflict with an universally bound graph  $C_j$ . Then, there does exists a transformation  $t : G \implies_\rho H$  with  $\rho = C_{k-1} \xleftarrow{l} C_{k-1} \xrightarrow{r} C_k$  such that a new occurrence  $p$  of  $C_j$  has been inserted. Since only elements of  $C_k \setminus C_{k-1}$  have been inserted it holds that  $p(C_j) \cap n(C_k \setminus C_{k-1})$  with  $n$  being the comatch of  $t$ . The graph  $p(C_j) \cup n(C_k)$  is the searched for overlap and the rule  $\rho^{-1} = C_k \xleftarrow{l} C_k \xrightarrow{r} C_{k-1}$  has to be applicable at  $n$ .  
“ $\Leftarrow$ ”: Let  $C_k$  be an existentially and  $C_j$  an universally bound graph such that an overlap  $P$  of  $C_k$  and  $C_j$  with  $i_{C_k}^P(C_k \setminus C_{k-1}) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$  exists such that the rule  $\rho = C_k \xleftarrow{l} C_k \xrightarrow{r} C_{k-1}$  is applicable at match  $i_{C_k}^P$ . Then, the inverse transformation of  $t : P \implies_{\rho, i_{C_k}^P} H$  is the searched for transformation and  $C_k$  has a conflict with  $C_j$ .
2. “ $\implies$ ”: Let  $C_k$  be an universally bound graph that has a conflict with an existentially bound graph  $C_j$ . Then, a transformation  $t : G \implies_\rho H$  with  $\rho = C_k \xleftarrow{l} C' \xrightarrow{r} C'$  and  $C'$  being a subgraph of  $C_k$  exists such that a occurrence  $p$  of  $C_j$  has been deleted. Then, the graph  $p(C_j) \cup m(C_k)$  is the searched for overlap and  $i_{C_k}^P(C_k \setminus C_{k-1}) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$  has to hold since  $\rho$  only deletes elements of  $C_k \setminus C_{k-1}$ .  
“ $\Leftarrow$ ”: Let  $C_k$  be universally and  $C_j$  existentially bound such that an overlap  $P$  of  $C_k$  and  $C_j$  with  $i_{C_k}^P(C_k \setminus C_{k-1}) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$  exists such that each rule  $\rho = C_k \xleftarrow{i_{C'}^{C'}} C' \xrightarrow{\text{id}} C'$  with  $C' \in \text{IG}(C_{k-1}, C_k)$  is applicable at match  $i_{C_k}^P$ . Then, the inverse transformation of  $t : P \implies_{\rho, i_{C_k}^P} H$  is the searched for transformations and  $C_k$  has a conflict with  $C_j$ .

□

**Definition 5.5 (Conflicts between conditions).** Let conditions  $c_1$  and  $c_2$  be given. Then,  $c_1$  has a conflict with  $c_2$  if graphs  $C_k$  and  $C_j$  of  $c_1$  and  $c_2$  exist such that  $C_k$  has a conflict with  $C_j$  or vice versa. A set of conditions is called conflict free if no conditions  $c_1$  and  $c_2$  exist such that  $c_1$  has a conflict with  $c_2$  or vice versa. A condition  $c_1$  has a transitive conflict with  $c_2$  if a condition  $c'$  exists such that  $c_1$  has a conflict with  $c'$  and  $c'$  has a (transitive) conflict with  $c_2$ . A set of conditions does contain a circular conflict if a condition  $c$  has a transitive conflict with itself. Otherwise, the set is called circular conflict free.

**Definition 5.6 (conflict free graphs).** Let a graph  $G$  and a constraint  $c$  in UANF be given. Then,  $G$  is called conflict free w.r.t  $c$  if for each transformation  $t : G \Rightarrow_{\rho, m} H$  with

$$\rho = C_{k_{\max}} \xleftarrow{\text{id}} C_{k_{\max}} \xrightarrow{i_{C_{k_{\max}}}} C'$$

and  $C' \in \text{IG}(C_{k_{\max}+1}, C_{k_{\max}+2})$  or

$$\rho = C_{k_{\max}+1} \xleftarrow{1_{C'}} C' \xrightarrow{\text{id}} C'$$

and  $C' \in \text{IG}(C_{k_{\max}}, C_{k_{\max}+1})$  and (4.4) and (4.5) hold.

**Lemma 5.7.** Let a graph  $G$  and a constraint  $c$  in UANF be given. Let  $P$  be the set of all Then,  $G$  is conflict free w.r.t  $c$  if and only if either  $c$  is conflict free or

$$G \models \bigwedge_{P \in P} \forall(a : \emptyset \rightarrow P, \text{false})$$

*Proof.*

□

## 5.2 Basic increasing rules

As described above, the application conditions for general rules are very restrictive. This is because the application condition does not allow a transformation to delete occurrences of existentially bound or insert occurrences of universally bound graphs even if this would not lead to a decrease of satisfaction up to layer. Now, we will consider a special set of rules called *basic increasing rules*. The main idea is that these rules are not able to delete occurrences of existentially or insert occurrences of universally bound graphs and additionally are able to increase consistency. That means, given a basic increasing rule  $\rho$ , there does exist a transformation  $t : G \Rightarrow_{\rho} H$  such that  $t$  is a consistency increasing transformation w.r.t to a constraint  $c$  in UANF. For *basic increasing rules*, consistency increasing application conditions can be constructed that are less restrictive and smaller in size compared to application conditions for general rules.

First, we define *non-consistency decreasing rules at layer* and *non-consistency decreasing rules up to layer* where a non-consistency decreasing rule at layer  $k$  is not able to delete occurrences of  $C_{k+1}$ , if  $\text{sub}_k(c)$  is existentially bound and not able to create occurrences of  $C_{k+1}$  if  $\text{sub}_k(c)$  is universally bound. This can be checked via second order formulas by, in the first case, checking whether all occurrences of  $C_{k+1}$  still exist in  $H$  and, in the second case, checking whether all occurrences of  $C_{k+1}$  in  $H$  already existed in  $G$ . The notion of non-consistency decreasing rules up to layer  $k$  is a abbreviation, meaning that a rule  $\rho$  is non-consistency decreasing up to layer  $k$  if it is non-consistency decreasing at layer  $j$  for all  $0 \leq j \leq k$ .

**Definition 5.8 (non-consistency decreasing rule).** Let a plain rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  and a constraint  $c$  in UANF be given. Then,  $\rho$  is called a non-consistency decreasing rule w.r.t  $c$  at layer  $k$  if

1. If  $\text{sub}_k(c)$  is universally bound and for all transformations  $t : G \Rightarrow_{\rho} H$  it holds that

$$\forall p : C_{k+1} \hookrightarrow H (\exists p' : C_{k+1} \hookrightarrow G (\text{tr}_t \circ p' = p)). \quad (5.2)$$

2. If  $\text{sub}_k(c)$  is existentially bound for all transformations  $t : G \Rightarrow_{\rho} H$  it holds that

$$\forall p : C_{k+1} \hookrightarrow G (\text{tr}_t \circ p \text{ is total}). \quad (5.3)$$

The rule  $\rho$  is called non-consistency decreasing w.r.t  $c$  up to layer  $k$  if  $\rho$  is non-consistency decreasing w.r.t  $c$  at layer  $j$  for all  $0 \leq j \leq k$  and it holds that

$$\bigwedge_{C' \in \text{IG}(C_k, C_{k+1})} \forall p : C' \hookrightarrow G (\text{tr}_t \circ p \text{ is total}) \quad (5.4)$$

if  $scondkc$  is existentially bound.

Given a graph  $G$  and a constraint  $c$ , we will show that non-decreasing rules w.r.t  $c$  up to layer  $k_{\max}$  cannot decrease the satisfaction up to layer.

**Lemma 5.9.** Let a graph  $G$ , a constraint  $c$  in UANF and a non-consistency decreasing rule w.r.t  $c$  up to layer  $k_{\max}$  be given. Then, for each transformation  $t : G \Rightarrow_{\rho} H$  it holds that

$$k_{\max}(c, G) \leq k_{\max}(c, H).$$

*Proof.* Assume that  $k_{\max}(c, G) > k_{\max}(c, H)$ . Then, either 1. or 2. applies.

1. There does exists an occurrence  $p : C_j \hookrightarrow H$  with  $0 \leq j \leq k_{\max}(c, G)$  of a universally bound graph  $C_j$  such that there does not exist a occurrence  $q : C_j \hookrightarrow G$  of  $C_j$  with  $p = \text{tr}_t \circ q$ . Then  $\rho$  is not a non-consistency decreasing rule up to layer  $k_{\max}(c, G)$  since (5.2) is not satisfied.
2. There does exists an occurrence  $p : C_j \hookrightarrow G$  with  $0 \leq j \leq k_{\max}(c, G)$  of an existentially bound graph  $C_j$  such that  $\text{tr}_t \circ p$  is not total. Then,  $\rho$  is not a non-consistency decreasing rule up to layer  $k_{\max}(c, G)$  since (5.3) is not satisfied.

□

Since it is very time consuming to check whether a rule  $\rho$  is non-consistency decreasing based on the definition above, we present a characterisation for non-consistency decreasing rules which only relies on  $\rho$  itself. First, let us assume that  $\rho$  is able to create occurrences of a universally bound graph  $C_j$ . This is possible if (a)  $\rho$  does insert an edge of  $C_j \setminus C_{j-1}$  which connects already existing nodes of  $C_j$ , since it is unclear whether this would create a new occurrence of  $C_j$ . Or (b) if  $\rho$  does insert a node  $v$  of  $C_j$  such that all edges  $e \in E_{C_j}$  with  $\text{src}(e) = v$  or  $\text{tar}(e) = v$  are also inserted. If at least one of these edges is not inserted, it is guaranteed that this insertion does not create an occurrence of  $C_j$  since  $v$  is only connected to edges that have also been inserted by  $\rho$ .

Second, let us assume that  $\rho$  is able to delete occurrences of a existentially bound graph  $C_j$ . This is possible if (a)  $\rho$  does delete an edge of  $C_j \setminus C_{j-1}$  or (b)  $\rho$  does delete a node  $v$  of  $C_j \setminus C_{j-1}$  such that all edges  $e \in E_{C_j}$  with  $\text{src}(e) = v$  or  $\text{tar}(e) = v$  are also deleted. If  $\rho$  deletes a node  $c$  of  $C_j \setminus C_{j-1}$  without all its connected edges in  $C_j$  there does not exist a transformation such that  $\rho$  deletes an occurrence of  $C_j$  by deleting this node since the dangling edge condition is not satisfied.

Then, we are able to determine whether a rule  $\rho$  is non-consistency decreasing by checking whether  $\rho$  does not satisfy the just mentioned properties.

**Lemma 5.10.** *Let a plain rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  and a constraint  $c$  in EANF be given. Then,  $\rho$  is a non-consistency decreasing rule at layer  $k$  w.r.t  $c$  if either 1. or 2. applies.*

1. *If  $C_k$  is universally bound, let  $I$  be the set of all isolated nodes of  $C_k$ ,  $P$  be the set of all partial morphisms  $p : C_k \hookrightarrow R$ ,*

$$E = \bigcup_{p \in P} p(E_{C_k}) \cap (E_R \setminus E_K) \text{ and } V = \bigcup_{p \in P} p(V_{C_k}) \cap (V_R \setminus V_K).$$

*Then,  $\rho$  is called a non-consistency decreasing rule at layer  $k$  w.r.t  $c$  if  $I \cap V = \emptyset$  and*

$$\begin{aligned} & (\neg \exists e \in E : \exists v, v' \in V_K : \text{src}(e) = r(v) \wedge \text{tar}(e) = r(v')) \quad \wedge \\ & (\forall v \in V : \exists e \in E_{C_k} : \exists p \in P : \text{src}(e) = p^{-1}(v) \vee \text{tar}(e) = p^{-1}(v) \\ & \quad \wedge \neg \exists e' \in E : \exists p \in P : p^{-1}(e') = e) \end{aligned}$$

*holds.*

2. *If  $C_k$  is existentially bound, let  $I$  be the set of all isolated nodes of  $C_k$ ,  $P$  be the set of al morphisms  $p : C_k \hookrightarrow L$ ,*

$$E = \bigcup_{p \in P} p(E_{C_k}) \cap (E_L \setminus E_K) \text{ and } V = \bigcup_{p \in P} p(V_{C_k}) \cap (V_L \setminus V_K).$$

*Let  $I$  be the set of all isolated nodes of  $C_k$ . Then,  $\rho$  is called a non-consistency decreasing rule at layer  $k$  w.r.t  $c$  if  $I \cap V = \emptyset$  and*

$$\begin{aligned} & (\neg \exists e \in E_{C_k} \setminus E_{C_{k-1}} : \exists p \in P : \exists e' \in E : p(e) = e') \quad \wedge \\ & (\forall v \in V : \exists e \in E_{C_k} : \exists p \in P : \text{src}(e) = p^{-1}(v) \vee \text{tar}(e) = p^{-1}(v) \\ & \quad \wedge \neg \exists e' \in E : \exists p \in P : p^{-1}(e') = e) \end{aligned}$$

*Proof.*

□

Now we are ready to define *basic increasing rules at layer  $k$*  with the set of basic increasing rules being a subset of the set of non-decreasing rules. With the property that each basic increasing rule  $\rho$  is also a non-decreasing rule up to layer  $k$  it is satisfied

that the satisfaction up to layer is not decreased. Additionally, if  $\text{sub}_k(c)$  is existentially bound it has to be checked that the consistency for an occurrence of  $C_k$  is not decreased and that  $\rho$  removes an violation. This can be done by either deleting an element of  $C_{k+1} \setminus C_k$  if  $\text{sub}_k(c)$  is universally bound or by inserting elements of  $C_{k+1}$  such that for one occurrence  $p$  of  $C_k$  the consistency is increased if  $\text{sub}_k(c)$  is existentially bound. Additionally, the left-hand side of each basic increasing rule has to contain  $C_{k+1}$  if  $\text{sub}_k(c)$  is universally bound such that either  $C_{k+1}$  gets deleted and the left-hand side has to contain  $C_k$  if  $\text{sub}_k(c)$  is existentially bound such that the consistency of  $C_k$  is increased. This property yields the advantage that application conditions for basic increasing rules are less complex, smaller in size and during a repair process matches of these rules are easier to handle.

This, on first sight seems like a restriction of the set of basic increasing rules but the context of each rule  $\rho$  that does satisfy all properties of a basic increasing rule excluding that  $C_k$  or  $C_{k+1}$  is a subgraph of the left-hand side can be expanded such that  $\rho$  is a basic increasing rule and the semantic of  $\rho$  is not increased. Later on, a method to derive this rules will be presented.

Basic increasing rules at layer  $k$  are called *deleting basic increasing rules* if  $\text{sub}_k(c)$  is universally bound and *inserting basic increasing rules* if  $\text{sub}_k(c)$  is existentially bound. For deleting basic increasing rules we introduce the restriction that these rules are only allowed to delete edges but no nodes of  $C_{k+1}$  since otherwise it is not possible, given a rule set and a constraint to decide whether this rules set is able to repair an arbitrary graph based only on deleting basic increasing rules. For example, consider a rule that deletes a node of  $C_{k+1}$ . Then, it is unknown whether this node is connected by edges not belonging to  $C_{k+1}$  and it is unclear whether all occurrence of  $C_{k+1}$  could be destroyed by  $\rho$  since the dangling edge condition could be unsatisfied.

**Definition 5.11 (basic increasing rule).** Let a constraint  $c$  in UANF and a plain rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  be given. Then,  $\rho$  is called basic increasing w.r.t  $c$  at layer  $k$  if

1. If  $\text{sub}_k(c)$  is universally bound:  $\rho$  is a non-consistency decreasing rule up to layer  $k+1$  and there does exist a morphism  $i : C_{k+1} \hookrightarrow L$  such that

$$\begin{aligned} \exists e \in E_{C_{k+1}} \setminus E_{C_k} (\neg \exists e' \in E_K(i(e) = l(e')) \wedge \\ \forall v \in V_{C_{k+1}} \setminus V_{C_k} (\exists v' \in V_L(i(v) = l(v'))) \end{aligned}$$

holds. Then,  $\rho$  is called a deleting basic increasing rule.

2. If  $\text{sub}_k(c)$  is existentially bound:  $\rho$  is a non-consistency decreasing rule up to layer  $k$  and there does exist a morphism  $i : C_k \hookrightarrow L$  such that

$$i \not\models \exists(a_k^r : C_k \hookrightarrow C', \text{true}) \wedge r \circ l^{-1} \circ i \models \exists(a_k^r : C_k \hookrightarrow C', \text{true})$$

for any  $C' \in \text{IG}(C_k, C_{k+1})$ . Then,  $\rho$  is called a inserting basic increasing rule with  $C'$ .

Again, 5.11 relies on every transformation of a rule  $\rho$ . Therefore, we present an alternative method to determine whether  $\rho$  satisfies 5.11 or not by checking that  $\rho$  does not delete any edges or nodes of  $C_{k+1} \setminus C_k$ .

**Lemma 5.12.** *let the sets  $E$  and  $V$  be constructed as introduced in lemma 5.10 it holds that*

$$\begin{aligned}\forall v \in V : \exists v' \in V_K : l(v') = v \wedge \\ \forall e \in E : \exists e' \in E_K : l(e') = e\end{aligned}$$

*Then,  $\rho$  is called an inserting basic increasing rule and  $i$  is called the increasing morphism of  $\rho$ .*

*Proof.*

□

**Lemma 5.13.** *Let a constraint  $c$  in UANF, a basic increasing rule  $\rho$  at layer  $k_{\max} + 1$  and a transformation  $t : G \Rightarrow_{\rho} H$  be given. Then,*

$$t \text{ is consistency increasing w.r.t } c \iff t \text{ is direct consistency increasing w.r.t } c$$

*Proof.* The “ $\Leftarrow$ ” side of the equivalence holds with lemma 4.16. We need to show that the “ $\Rightarrow$ ” side of the equivalence holds. Let  $t : G \Rightarrow_{\rho} H$  be a consistency increasing transformation and  $\rho$  a basic increasing rule. Then,  $t$  satisfies (5.11) and the satisfaction of (4.1) follows immediately. Since  $\rho$  is also a non-consistency decreasing rule up to layer  $k_{\max}$ ,  $t$  satisfies (5.2) and (5.3), then the satisfaction (4.2), (4.4) and (4.5) also follows immediately.

1. If  $\rho$  is a deleting basic increasing rule. Since  $t$  is a consistency increasing rule there has to exist a morphism  $p : C_{k_{\max} + 2} \hookrightarrow G$  such that  $\text{tr}_t \circ p$  is not total and with that (4.3) is satisfied.
2. If  $\rho$  is a inserting basic increasing rule there exists an occurrence  $p : C_{k_{\max} + 2} \hookrightarrow G$  and a graph  $C \in \text{IG}(C_{k_{\max} + 2}, C_{k_{\max} + 3})$  such that  $p \not\models \text{IC}_{k_{\max} + 2}(c, C)$  and  $\text{tr}_t \circ p \not\models \text{IC}_{k_{\max} + 2}(c, C)$ . Therefore (4.3) is satisfied.

In total follows that  $t$  is a direct consistency increasing transformation w.r.t  $c$ .

□

For basic increasing rules at layer  $j < k_{\max} + 1$  this equality does not hold since these rules are allowed to destroy occurrences of universally bound graphs  $C_{j'}$  or to create occurrences of existentially bound graphs  $C_{j'}$  with  $j < j' < k_{\max} + 1$ . With this results follow that if  $t : G \Rightarrow_{\rho} H$  is a  $c$ -guaranteeing transformation and  $\rho$  is a basic increasing rule at layer  $k_{\max} + 1$ ,  $t$  is also a direct increasing rule w.r.t  $c$ .

**Definition 5.14 (application conditions for basic increasing rules).** *Let a constraint  $c$  in UANF and a basic increasing rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  w.r.t  $c$  at layer  $k$  be given. We define the application condition for  $\rho$  as:*

1. If  $\rho$  is a deleting increasing rule:

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

2. If  $\rho$  is an inserting increasing rule. Let  $a^p : C_k \hookrightarrow C$  be a partial morphism of  $a_k$ ,  $e : C_k \hookrightarrow L$  be the increasing morphism of  $c$  and  $Q$  be set of all overlaps  $P$  of  $L$  and  $C$  such that  $i_L^P \circ e \models \exists(a^p : C_k \hookrightarrow C, \text{true})$ :

$$\text{basic}_j(\rho) := \begin{cases} \bigvee_{P \in Q} \exists(i_L^P : L \hookrightarrow P, \text{true}) & , \text{ if } j = k \\ \text{false} & , \text{ if } j \neq k \end{cases}$$

**Theorem 5.2.** Let a graph  $G$ , a constraint  $c$  and a basic increasing rule  $\rho$  at layer  $k_{\max} < k \leq k_{\max} + 2$  be given. Then, each transformation

$$t : G \xrightarrow{\rho', m} H$$

with the rule  $\rho' = (\rho, \text{basic}_k(\rho))$  is a direct consistency increasing transformation.

*Proof.* Since  $\rho$  is a basic increasing rule it is also a non-consistency decreasing rule up to layer  $k_{\max} + 2$ . Therefore, (4.1) is trivially satisfied since  $\rho$  satisfies (5.4), also (4.2) is satisfied since  $\rho$  satisfies (5.2) also (4.4) and (4.5) are satisfied because  $\rho$  satisfies (5.2) and (5.3) for all  $0 \leq j \leq k + 2$ . It remains to show that (4.3) is satisfied.

1. Let  $\rho$  be a deleting basic increasing rule. Then,  $k = k_{\max} + 1$  and  $\text{basic}_k(\rho) = \bigvee_{P \in \text{ol}(L, C_k)} \bigwedge_{P' \in \text{eol}(P, a_k)} \exists(i_L^P : L \hookrightarrow P, \neg \exists(i_P^{P'} : P \hookrightarrow P', \text{true}))$ . If  $m \models \text{basic}_j(\rho)$  there does exist a occurrence  $p : C_k \hookrightarrow G$  with  $p \not\models \exists(a_{k+1} : C_k \hookrightarrow C_{k+1}, \text{true})$  and  $p(C_k \setminus C_{k-1}) \cap m(L \setminus K) \neq \emptyset$ . It follows that  $\text{tr}_t \circ p$  is not total and (4.3) is satisfied.
2. Let  $\rho$  be a inserting basic increasing rule with  $C \in \text{IG}(C_k, C_{k+1})$ . Then,  $k = k_{\max} + 1$  and  $\text{basic}_j(\rho) = \bigvee_{P \in Q} \exists(i_L^P : L \hookrightarrow P, \text{true})$ . If  $m \models \text{basic}_j(\rho)$ ,  $m \circ e \not\models \exists(a_k^r : C_k \hookrightarrow C, \text{true})$  with  $e$  being the increasing morphism of  $\rho$ . Therefore  $\text{tr}_t \circ m \circ i \models \exists(a_k^r : C_k \hookrightarrow C, \text{true})$  and (4.3) is satisfied.

In total follows that  $t$  is a direct consistency increasing transformation.  $\square$

**Theorem 5.3.** 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  $c$ ,  $(\rho, \text{basic}(j, C_{j+1})) \in \mathcal{R}$  and  $\rho$  is a deleting potentially minimal improving rule at layer  $j$  with  $C_{j+1}$ , such that  $\rho$  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  $(\rho, \text{basic}(k, P)) \in \mathcal{R}$  exists, such that  $\rho$  is an inserting basic improving rule at layer  $k$  with  $P$ .

3. There does exist a sequence of graphs

$$C_k = C'_0 \hookrightarrow C'_1 \hookrightarrow \dots \hookrightarrow C'_n = C_{k+1}$$

such that a inserting basic increasing rule  $\rho_i = L_i \xleftarrow{l_i} K_i \xrightarrow{r_i} R_i$  at layer  $k$  with  $C_i$  exists such that  $L_i \hookrightarrow R_{i-1}$  and  $e_i = i_{C_{i-1}}^{C_i} \circ e_{i-1}$  with  $e_i$  being the increasing morphism of  $\rho_i$  for all  $i \in \{1, \dots, n\}$ .

*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 5.3 is satisfied.

1. Assume that 1. of theorem 5.3 holds. Let  $(\rho, \text{basic}(j, C_{j+1})) \in \mathcal{R}$ , 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,  $\text{basic}(j, C_{j+1}) = \exists(b : L \hookrightarrow C_j, \neg \exists(a_j : C_j \hookrightarrow C_{j+1}, \text{true}))$ . 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  $m_1 \models \text{basic}(j, C_{j+1})$ . 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 4.9,  $H \models_{k+2} c$  and  $H \models c$  follows.

2. Assume that 2. of theorem 5.3 holds. Let  $\rho_0 = (\rho, \text{basic}(k, P)) \in \mathcal{R}$ , such that  $\rho$  is an inserting basic improving rule of  $c$  at layer  $k$  with  $P \in \mathbf{P}$ . Then,  $\text{basic}(k, P) = \neg \exists(b : L \hookrightarrow P, \text{true})$ . 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  $m_0 \models \neg \exists(a'_k : C_k \hookrightarrow P, \text{true}) = \text{basic}(k, P)$ . Because  $r$  does not delete any elements, a transformation  $t_0 : G \Rightarrow_{r, m_0} G_1$  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  $b_\ell$  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 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 5.15.** *If a set of rule  $\mathcal{R}$  is a repairing set of  $c$  at layer  $k \leq \text{nl}(c)$  and 1. of theorem 5.3 applies, then  $\mathcal{R}$  is a repairing set of  $c$  at layer  $j$  for all  $k \leq j \leq \text{nl}(c)$ .*

**Lemma 5.16.** *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*

$$G_0 \implies_{\rho_0, m_0} G_1 \implies_{\rho, m_0} \dots \implies_{\rho_n, m_n} G_n$$

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

### 5.3 Repairing rule Sets

**Definition 5.17 (repairing rule set).** *Let a constraint  $c$  in UANF 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  $k_{\max} = k$  a sequence*

$$G = G_0 \implies_{\rho_0} \dots \implies_{\rho_{n-1}} G_n = H$$

*exists, such that  $k_{\max}(c, H) \geq k_{\max}(c, G) + 2$  and  $\rho_j \in \mathcal{R}$  for all  $0 \leq j \leq n - 1$ .*

**Corollary 5.18.** *Let a constraint  $c$  in UANF and a rule set  $\mathcal{R}$  be given. If either (a)  $\mathcal{R}$  is a repairing rule set at layer  $k$  for all odd  $0 \leq k \leq \text{nl}(c)$  or (b)  $\mathcal{R}$  is a repairing rule set at layer  $k$  with  $k$  being even and  $\mathcal{R}$  is also a repairing rule set for all odd  $0 \leq j < k$  then for all graphs  $G$  a sequence*

$$G = G_0 \implies_{\rho_0} \dots \implies_{\rho_{n-1}} G_n$$

*with  $G_n \models c$  and  $\rho_j \in \mathcal{R}$  for all  $0 \leq j \leq n - 1$  exists.*

### 5.3.1 Construction of repairing sets

**Definition 5.19 (decomposition of a graph).** Let graphs  $G_0 \hookrightarrow G_1$  be given.

A decomposition of  $G_1$  with  $G_0$  is a minimal set

$$\mathbf{D} \subseteq \{D_v \mid v \in V_{G_1 \setminus C_0}\}$$

of subgraphs of  $G_1$ , such that every element of  $G_1$  is contained in at least one  $D \in \mathbf{D}$  and every  $D_v$  is constructed in the following way:  $G_0 \hookrightarrow D_v$ ,  $v \in D_v$  and for all nodes  $v' \in D \setminus C_0$   $D$  contains all edges  $e \in E_{G_1}$  and all nodes  $u \in V_{G_1}$  such that either  $\text{tar}(e) = v' \wedge \text{src}(e) = u$  or  $\text{src}(e) = u \wedge \text{tar}(e) = v'$ .

**Lemma 5.20.** Let graphs  $G_0 \hookrightarrow G_1$  and a decomposition  $\mathbf{D}$  of  $G_1$  with  $G_0$  be given. Then, for each pair  $D, D' \in \mathbf{D}$  with  $D \neq D'$  the following holds:

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

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

**Lemma 5.21.** Let graphs  $G_0 \hookrightarrow G_1$  and a decomposition  $\mathbf{D}$  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  $E_{G_1} \subseteq E_H$ .  $\square$

Now we are ready to provide the construction for

**Construction 5.22.** Let a constraint  $c$  in UANF be given. For each existentially bound subcondition  $\text{sub}_k(c)$  we construct a graph decomposition  $\mathbf{D}_k$  as described in definition 5.19. Then, for each element  $D \in \mathbf{D}_k$  a rule

$$\rho = C_k \xleftarrow{\text{id}} C_k \xrightarrow{i_{C_k}^D} D$$

is derived.

Additionally, for each universally bound subcondition  $\text{sub}_k(c)$  and for each edge  $e \in E_{C_{k+1} \setminus C_k}$  a rule

$$\rho = C_{k+1} \xleftarrow{i_D^{C_{k+1}}} D \xrightarrow{\text{id}} D$$

with  $D = C_{k+1} \setminus \{e\}$  is derived.

---

**Algorithm 1:** Repair for conflict free constraints and repairing set

---

**Data:** A graph  $G$ , a conflict free constraint in UANF and a repairing set  $\mathcal{R}$  for each layer  $k_{\max} < k \text{ nl}(c)$

**Result:** A graph  $H$  with  $H \models c$

```
1 for  $i = k_{\max} + 1$  to  $\text{nl}(c)$  do
2   while  $G \not\models_i c$  do
3     | Choose one occurrence  $p : C_i \hookrightarrow G$  uniformly at random ;
4     | apply a repairing sequence  $G \Rightarrow \dots \Rightarrow H$ ;
5   end
6 end
```

---

**Example 5.2.**

## 6 Rule-based Graph Repair

### 6.1 Rule based Graph repair for conflict free constraints

**Theorem 6.1.** Let a rule set  $\mathcal{R}$  and a conflict free constraint  $c$  in UANF be given. Then,  $\mathcal{R}$  is a repairing set of  $c$  at layer  $k \leq \text{nl}(c)$  if a sequence

$$C_k \Rightarrow_{\rho_0, m_0} \dots \Rightarrow_{\rho_n, m_n} H$$

of consistency increasing or non-consistency decreasing transformations  $t_0, \dots, t_n$  with  $\rho_0, \dots, \rho_n \in \mathcal{R}$  exists such that

$$\forall v \in V_{C_k} (\exists v' \in V_H (\text{tr}_{t_n} \circ \dots \circ \text{tr}_{t_0} \circ \text{id}_{C_k}(v) = v'))$$

and

1. If  $\text{sub}_k(c)$  is universally bound:

$$\text{tr}_{t_n} \circ \dots \circ \text{tr}_{t_0} \text{ is not total}$$

2. If  $\text{sub}_k(c)$  is existentially bound:

$$\text{tr } t_n \circ \dots \text{ tr } t_0 \circ \text{id}_{C_k} \models \text{sub}_k(c) \text{cut}_{k+1}(c)$$

*Proof.*

□

**Lemma 6.1.** The set rule set constructed as described in construction 5.22 is a repairing set.

*Proof.*

□

**Theorem 6.2.** Algorithm 1 is correct and does terminate.

*Proof.*

□

---

**Algorithm 2:** Repair for non conflict free constraints and repairing set

---

**Data:** A graph  $G$ , a non-conflict free constraint in UANF and a repairing set  $\mathcal{R}$   
for each layer  $k_{\max} < k \text{ nl}(c)$

**Result:** A graph  $H$  with  $H \models c$

```
1 for  $i = k_{\max} + 1$  to  $\text{nl}(c)$  do
2   while  $G \not\models_i c$  do
3     Choose one occurrence  $p : C_i \hookrightarrow G$  uniformly at random ;
4     apply a repairing sequence  $G \Rightarrow \dots \Rightarrow H$ ;
5     while  $G \not\models_{k_{\max}} c$  do
6       Choose an occurrence  $p : C_j \hookrightarrow G$  uniformly at random;
7       apply the repairing sequence.
8     end
9   end
10 end
```

---

## 6.2 Rule based Graph repair for non-conflict free constraints

**Theorem 6.3.** *Algorithm 2 terminates and returns an consistent graph.*

*Proof.* □

## 6.3 Rule based repair for non-repairing sets

## 7 Conclusion

---

**Algorithm 3:** Repair for conflict free constraints and repairing set

---

**Data:** A graph  $G$ , a conflict free constraint in UANF and a repairing set  $\mathcal{R}$  for each layer  $k_{\max} < k \text{ nl}(c)$

**Result:** A graph  $H$  with  $H \models c$

```
1  $\mathcal{G} \leftarrow \{G\}$  for  $i = 0$  to  $m$  do
2   foreach  $G \in \mathcal{G}$  do
3     foreach  $t : G \Rightarrow H$  do
4       if  $H \models c$  then
5         return  $H$ 
6       end
7       if  $t$  is non-consistency decreasing or consistency increasing then
8          $\mathcal{G} \leftarrow \mathcal{G} \cup \{H\};$ 
9       end
10      end
11    end
12 end
```

---

## References

- [1] H. Ehrig, K. Ehrig, U. Prange, and G. Taentzer. [Fundamentals of algebraic graph transformation](#). *Monographs in theoretical computer science. An EATCS series*. Springer, 2006.
- [2] A. Habel and K.-H. Pennemann. [Correctness of high-level transformation systems relative to nested conditions](#). *Mathematical Structures in Computer Science*, 19(2):245–296, 2009.
- [3] J. Kosiol, D. Strüber, G. Taentzer, and S. Zschaler. [Sustaining and improving graduated graph consistency: A static analysis of graph transformations](#). *Science of Computer Programming*, 214:102729, 2022.
- [4] K.-H. Pennemann. [Development of Correct Graph Transformation Systems](#). PhD thesis, Department für Informatik, Universität Oldenburg, Oldenburg, 2009.
- [5] D. Plump. Confluence of graph transformation revisited. In *Processes, Terms and Cycles: Steps on the Road to Infinity*, pages 280–308. Springer, 2005.
- [6] C. Sandmann and A. Habel. [Rule-based graph repair](#). *arXiv preprint arXiv:1912.09610*, 2019.