

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

Alexander Lauer

January 15, 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 and Maintaining Transformations and Rules . . . . .	14
4.4	Direct Consistency Maintaining and Increasing Transformations . . . . .	16
4.5	Comparison with other concepts of Consistency . . . . .	20
<b>5</b>	<b>Application Conditions</b>	<b>24</b>
5.1	General Application Conditions . . . . .	26
5.2	Basic Increasing and Maintaining Rules . . . . .	31
5.3	Application Conditions for Basic Rules . . . . .	37
<b>6</b>	<b>Rule-based Graph Repairs</b>	<b>39</b>
6.1	Repairing rule Sets . . . . .	39
6.1.1	Construction of repairing sets . . . . .	41
6.2	Conflicts within and between Conditions . . . . .	42
6.3	Rule based Graph repair for conflict free constraints . . . . .	44
6.4	Rule based Graph repair for non-conflict free constraints . . . . .	46
6.5	Rule based repair for non-repairing sets . . . . .	46
<b>7</b>	<b>Conclusion</b>	<b>46</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 (or nodes)  $V$ , 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

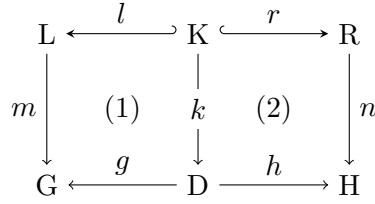


Figure 1: Diagram of a transformation in the double-pushout approach.

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.

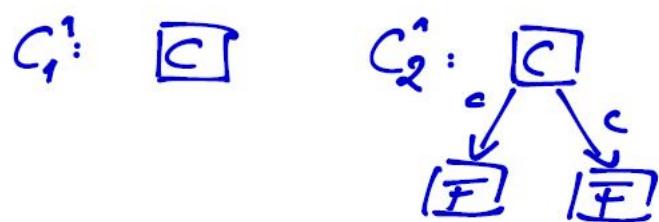
**Definition 4.6 (truncated condition).** Let  $c$  be a condition in UANF and  $d = \text{sub}_k(c)$  with  $0 \leq k \leq \text{nl}(c) - 1$ . The truncated condition at layer  $k$  of  $c$ , denoted by  $\text{cut}_k(c)$ , is defined as

$$\text{cut}_k(c) := \begin{cases} \text{rep}_{k+1}(c, \text{true}) & \text{if } d \text{ is existentially bound, i.e. } k \text{ is odd} \\ \text{rep}_{k+1}(c, \text{false}) & \text{if } d \text{ is universally bound, i.e. } k \text{ is even.} \end{cases}$$

**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.

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



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

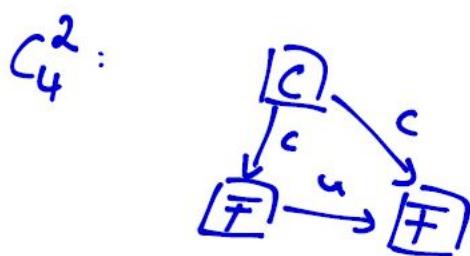
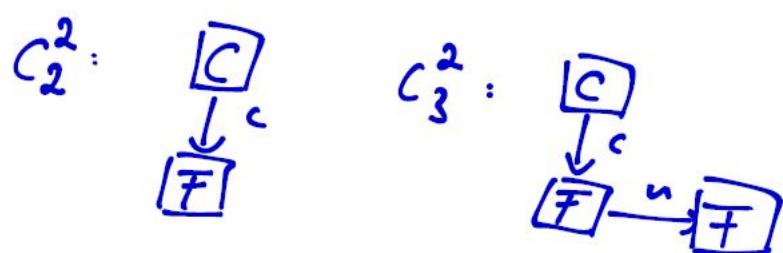


Figure 2: constraints

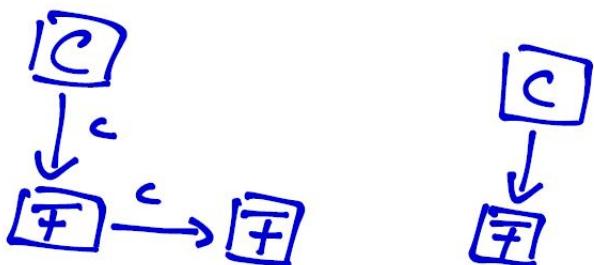


Figure 3: graph

$p \models_k c$		$p \models_{j < k} c$		$p \models_{j > k} c$		$p \models c$
		$j$ even	$j$ odd	$j$ even	$j$ odd	
$k$ even	?	✓	✓	✓	✓	✓
$k$ odd	?	✓	?	?	?	?

Table 1: Overview of the conclusions made via satisfaction at layer with “✓” indicating that  $p \models_j$  and  $p \models c$ , respectively, if  $p \models_k$ , And “?” indicating that it cannot be concluded from  $p \models_k c$  whether  $p \models_j c$  or  $p \not\models_j c$ .

**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  $0 \leq k \leq \text{nl}(c) - 1$ , 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  $\text{k}_{\max}(c, G)$ . If no such  $k$  exists, we set  $\text{k}_{\max}(c, k) = -1$ . 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  $c$ , by definition and since  $\text{nl}(c)$  is finite,  $\text{k}_{\max}(c, k)$  always exists. Also, if  $p \models_{\text{nl}(c)-1} c$  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 the second occurrence of `Class` does not satisfy  $\exists C_2^2 \forall C_3^2 \exists C_4^2$ , but it satisfies  $\text{cut}_1(c_2)$  and therefore

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

Let a graph  $G$ , a condition  $c$  and a morphism  $p : C_0 \hookrightarrow G$  be given. Assume that  $p \models_k c$  for any  $0 \leq k < \text{nl}(c)$ . Then, we are able to conclude results for the satisfaction at other layers. If  $k$  is even, that means  $\text{sub}_k(c)$  is universally bound, we can conclude that  $p \models_j c$  for all  $k < j < \text{nl}(c)$  and in particular  $p \models c$ . For any  $0 \leq k < \text{nl}(c)$  with  $p \models_k c$  we can conclude that  $p \models_j c$  for all odd  $0 \leq j < k$ . An overview of these conclusion is shown in Table 1. We show these results within the following lemmas.

We start by investigating the conclusions for satisfaction at layer  $j > k$  if  $p \models_k c$ . Our first result shows that the replacement of the sub-condition  $\text{sub}_{k+1}(c)$  by any arbitrary condition over  $C_{k+1}$  leads to a condition that is satisfied by  $p$  if  $k$  is even.

**Lemma 4.8.** Let a graph  $G$ , a condition  $c$  in UANF and a morphism  $p : C_0 \hookrightarrow G$  with  $p \models_k c$ , such that  $0 \leq k < \text{nl}(c)$  is even, be given. Then, for any condition  $f$  over  $C_{k+1}$  it holds that

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

*Proof.* Let  $0 \leq j \leq \text{nl}(c)$  be the smallest number with  $\text{sub}_j(c) = \forall(a_j : C_j \hookrightarrow C_{j+1}, d)$  being universally bound and  $p \models_j c$ . This must exist, since at least one of these exists due to the assumption. Let  $q : C_j \hookrightarrow G$  be a morphism such that  $q \models \forall(a_j : C_j \hookrightarrow C_{j+1}, \text{false})$ .

This must exist, since  $p \models_j c$  and  $j$  is the smallest even number such that  $p \models_j c$ . Therefore, there does not exist a morphism  $q' : C_{j+1} \hookrightarrow G$  with  $q = q' \circ a_j$ . Hence, for every condition  $f$  over  $C_{j+1}$  a morphism  $q' : C_{j+1} \hookrightarrow G$  with  $q \not\models f$  and  $q = q' \circ a_j$  cannot exist. It follows immediately that  $q \models \forall(a_j : C_j \hookrightarrow C_{j+1}, f)$  and with that  $p \models \text{rep}_{j+1}(c, f)$ .

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

As a direct consequence of the previous lemma, a morphism satisfying a condition at layer  $k$  with  $k$  being even also satisfies the condition at layer  $j$  for all  $j > k$ .

**Lemma 4.9.** *Let a graph  $G$ , a morphism  $p : C_0 \hookrightarrow G$  and a condition  $c$  in UANF be given. If  $0 \leq k < \text{nl}(c)$  is even, i.e.  $\text{sub}_k(c)$  is universally bound, then for all  $k < j < \text{nl}(c)$  it holds that*

$$p \models_k c \implies p \models_j c.$$

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

Since a morphism  $p$  satisfies a condition  $c$  in UANF if and only if  $p$  satisfies  $c$  at layer  $\text{nl}(c) - 1$ , because  $\text{cut}_{\text{nl}(c)-1}(c) = c$ , we can conclude the following.

**Corollary 4.10.** *Let a graph  $G$ , a morphism  $p : C_0 \hookrightarrow G$  and a condition  $c$  in UANF be given. If  $0 \leq k < \text{nl}(c)$  is even it holds that*

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

Furthermore, this allows us to make statements about the satisfaction of other conditions. Let a graph  $G$ , a morphism  $p : C_0 \hookrightarrow G$  and a condition  $c$  be given such that  $p \models_k c$  for an even  $0 \leq k < \text{nl}(c)$ . Then,  $p \models c$  and for every condition  $c'$  with  $\text{cut}_k(c) = \text{cut}_k(c')$ . With Lemma 4.8, it follows that  $p \models c'$ .

Let us now investigate the satisfaction at layer  $j$  with  $j < k_{\max}$ . If  $j$  is odd, i.e.  $\text{sub}_j(c)$  is existentially bound, we can conclude that  $p \models_j c$  as shown in Lemma 4.11. If  $j$  is even, i.e.  $\text{sub}_j(c)$  is universally bound, we are only able to make statements depending on  $k_{\max}$ . If  $k_{\max} < \text{nl}(c) - 1$ , it follows that  $p \not\models_j c$ . Otherwise  $p \models c$  and therefore  $k_{\max} = \text{nl}(c) - 1$  would follow immediately with Corollary 4.10. If  $k_{\max} = \text{nl}(c) - 1$  we can only state that at least one even  $j \leq k_{\max}$  with  $p \models_j c$  exists if  $c$  ends with a condition of the form  $\forall(a_k : C_k \hookrightarrow C_{k+1}, \text{false})$ .

**Lemma 4.11.** *Let a graph  $G$ , a morphism  $p : C_0 \hookrightarrow G$  and a constraint  $c$  in UANF be given. Then, for all odd  $0 \leq j < k_{\max}$ , i.e.  $\text{sub}_j(c)$  is existentially bound, it holds that*

$$p \models_j c.$$

*Proof.* If an even  $0 \leq j < k_{\max}$  with  $p \models_j c$  exists such that  $\text{sub}_j(c)$  is universally bound, let  $j'$  be the smallest of these. With lemma 4.8 follows that  $p \models_\ell c$  for all  $j' \leq \ell < \text{nl}(c)$ . Otherwise, if no such  $j'$  exists we set  $j' = k_{\max}$ .

Let  $\ell < j'$ , such that  $\text{sub}_\ell(c)$  is existentially bound and let  $d = \text{sub}_\ell(\text{cut}_{j'}(c)) = \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, e)$  be the sub-condition at layer  $\ell$  of the condition up to layer  $j'$  of  $c$ . Since  $\ell < j'$ , 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  $t : G \implies H$  be a transformation. If  $k_{\max}(c, G) < k_{\max}(c, H)$ , we consider the transformation as consistency increasing, since  $H$  satisfies more layers of the constraint than  $G$ . But, the notion of consistency increasement should also be able to detect the smallest changes, performed by a transformation, 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*, which will be used to recognize these types of increasements by checking whether an intermediate condition not satisfied by  $G$  is satisfied by  $H$ . Obviously, an decrease of consistency can be detected in a similar manner, by checking whether an intermediate condition satisfied by  $G$  is not satisfied by  $H$ . Intuitively, given a constraint  $c$  in UANF with  $\text{sub}_k(c) = \exists(a_k : C_k \hookrightarrow C_{k+1}, d)$  and  $0 \leq k < \text{nl}(c)$ , the condition  $\text{sub}_k(c)$  is replaced by  $\exists(a_k^r : C_k \hookrightarrow C', \text{true})$  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  $c = \forall(a_0 : C_0 \hookrightarrow C_1, \text{false})$ , let  $C' \in \text{IG}(C_0, C_1)$ . If the condition  $c' = \forall(a_0^r : C_0 \hookrightarrow C', \text{false})$  is satisfied the satisfaction of  $c$  is implied but the backwards implication does not hold.

**Definition 4.12 (intermediate condition).** Let a condition  $c$  in UANF be given. Let  $0 \leq k < \text{nl}(c)$  such that  $k$  is odd, i.e.  $\text{sub}_k(c)$  is existentially bound. The intermediate condition, denoted by  $\text{IC}_k(c, C')$ , of  $c$  at layer  $k$  with  $C' \in \text{IG}(C_k, C_{k+1})$  is defined as

$$\text{IC}_k(c, C') := \text{rep}_k(c, \exists(a_k^r : C_k \hookrightarrow C', \text{true})).$$

**Example 4.5.** Consider constraint  $c_1$  given in figure 2. Since  $C_2^2 \in \text{IG}(C_1^1, C_2^1)$ , 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 type **Class** is connected to at least two nodes of type **Feature**, the intermediate condition checks whether each node of type **Class** is connected to at least one node of type **Feature** which is trivially satisfied if  $c_1$  is satisfied.

Given a graph  $G$  and a constraint  $c$  in UANF with  $k_{\max} < \text{nl}(c) - 3$ , note that in this case  $\text{sub}_{k_{\max}}(c)$  has to be existentially bound, it holds that  $G \not\models_{k_{\max}+2} c$  and there does exist at least one graph  $C' \in \text{IG}(C_{k_{\max}+2}, C_{k_{\max}+3})$  such that  $G \models \text{IC}_{k_{\max}+2}(c, C')$  since  $G$  always satisfies  $\text{IC}_{k_{\max}+2}(c, C_{k_{\max}+2})$ .

### 4.3 Consistency Increasing and Maintaining Transformations and Rules

With the results above, we are now ready to define the notions of *consistency increasement* and *maintainment*, with increasement being a special case of maintenance. A transformation  $t$  is considered as consistency maintaining if the consistency is not decreased, whereas  $t$  is considered as consistency increasing if the consistency has been increased.

These notions are designed to only detect transformations that maintain (or 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 consistency maintaining if  $k_{\max}(c, G) \leq k_{\max}(c, H)$ , i.e. the satisfaction up to layer is not decreased, and at least the same amount of 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  $C' \in \text{IG}(C_{k+1}, C_{k+2})$ . If  $k_{\max}(c, G) < k_{\max}(c, H)$  or the number of increasing insertions and deletions is greater than the number of decreasing ones,  $t$  is considered as consistency increasing.

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 layers  $k$  with  $k \leq k_{\max}$  it is set to 0 and for all layers  $k$  with  $k > k_{\max} + 1$  it is set to  $\infty$ . Through this, a transformation  $t : G \implies H$  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 4.13 (number of violations).** Let a graph  $G$  and a constraint  $c$  in UANF be given. Let  $e = \text{sub}_{k_{\max} + 2}(c)$ . The number of violations  $\text{nvc}_j(G)$  at layer  $0 \leq j < \text{nl}(c)$  in  $G$  is defined as:

$$\text{nvc}_j(G) := \begin{cases} 0 & \text{if } j < k_{\max} + 1 \\ \sum_{C' \in \text{IG}(C_{j+1}, C_{j+2})} |\{q \mid q : C_{j+1} \hookrightarrow G \wedge q \not\models \text{IC}_0(e, C')\}| & \text{if } e \neq \text{false} \text{ and } j = k_{\max} + 1 \\ |\{q \mid q : C_{j+1} \hookrightarrow G\}| & \text{if } e = \text{false} \text{ and } j = k_{\max} + 1 \\ \infty & \text{if } j > k_{\max} + 1 \end{cases}$$

Note that the second and third case of Definition 4.13 only apply if  $G \not\models c$  and  $\text{sub}_{k_{\max}}(c)$  is existentially bound. Therefore  $e$  is also existentially bound or equal to `false`,

if  $c$  ends with  $\forall(a_{\text{nl}(c)-1} : C_{\text{nl}(c)-1} \hookrightarrow C_{\text{nl}(c)}, \text{false})$ . Via the number of violations, we now define *consistency maintaining* and *increasing* transformations and rules, by checking whether the number of violations has not been increased, or in case of consistency increasing, has not been decreased for any layer of the constraint.

**Definition 4.14 (consistency maintaining and increasing transformations and rules).** Let a graph  $G$ , a rule  $\rho$  and a constraint  $c$  in UANF be given. A transformation  $t : G \Rightarrow_{\rho,m} H$  is called *consistency maintaining* w.r.t.  $c$ , if

$$\text{nvc}_k(H) \leq \text{nvc}_k(G)$$

for all  $0 \leq k < \text{nl}(c)$ . The transformation  $t$  is called *consistency increasing* w.r.t.  $c$  this inequality is strict. A rule  $\rho$  is called *consistency maintaining/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  $G \Rightarrow H$  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 with  $G \not\models c$  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  $k_{\max}(c, H) > k_{\max}(c, G)$ ,  $\text{nvc}_{k_{\max}(c,G)+1}(H) = 0$  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.

**Theorem 4.2.** Let a constraint  $c$  in UANF be given. Every sequence of consistency increasing transformations w.r.t  $c$  is finite.

*Proof.* Let  $G_0$  be a graph and

$$G_0 \Rightarrow_{\rho_0, m_0} G_1 \Rightarrow_{\rho_1, m_1} G_2 \Rightarrow_{\rho_2} \dots$$

be a sequence of consistency increasing transformations w.r.t  $c$ . We assume that  $k_{\max}(c, G_0) < \text{nl}(c)$ , otherwise  $\text{nvc}_j(G_0) = 0$  for all  $0 \leq j < \text{nl}(c)$  and no consistency increasing transformation  $G_0 \Rightarrow H$  w.r.t.  $c$  exists.

We show that after at most  $j := \text{nvc}_{\text{k}_{\max}(c, G_0)+1}(G_0)$  transformations  $G_j \models_{\text{k}_{\max}(c, G_0)+2} c$ . Note that  $j$  has to be finite, since  $G_0$  contains only a finite number of occurrences of  $C_{j+1}$ . Since each transformation is consistency increasing w.r.t.  $c$  it holds that  $\text{nvc}_{\text{k}_{\max}(c, G_i)+1}(G_{i+1}) \leq \text{nvc}_{\text{k}_{\max}(c, G_i)+1}(G_i) - 1$  after each transformation. Therefore, after at most  $j$  transformations,  $\text{nvc}_{\text{k}_{\max}(c, G_0)+1}(G_j) \leq \text{nvc}_{\text{k}_{\max}(c, G_0)+1}(G_0) - j = 0$  and  $G_j \models_{\text{k}_{\max}(c, G_0)+2} c$ . 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  $G_k \Rightarrow_{\rho_k, m_k} G_{k+1}$  exists, the sequence has to be finite.  $\square$

#### 4.4 Direct Consistency Maintaining and Increasing Transformations

Let a constraint  $c$  in UANF and graphs  $G$  with  $G \not\models c$  and  $H$  with  $H \models c$  be given. The transformation  $t : G \Rightarrow_{\rho, \text{id}_G} H$  via the rule  $\rho = G \xleftarrow{l} \emptyset \xrightarrow{r} H$  is a consistency increasing transformation. Therefore, the notions of consistency increase- and maintenance, a similar example for a consistency maintaining transformation can easily be constructed, does allow insertions or deletions that are unnecessary in order to increase or maintain consistency. That is, the deletion of occurrences of existentially bound graphs, the deletions of occurrences  $p : C_k \hookrightarrow G$  of universally bound graphs  $C_k$  such that  $p \models \exists(a_k : C_k \hookrightarrow C_{k+1}, \text{true})$  or the insertion of occurrences of universally bound graphs and the insertion of occurrences  $p$  of intermediate graphs  $C' \in \text{IG}(C_{k-1}, C_k)$  with  $C_k$  being existentially bound, such that each occurrence  $q$  of  $C_{k-1}$  with  $q = p \circ a_{k-1}^r$  already satisfied  $\exists(a_{k-1}^r : C_{k-1} \hookrightarrow C_k, \text{true})$ .

*Direct consistency increasing* and *maintaining* transformations are more restricted, in the sense that these unnecessary deletions and insertions are leading to a transformation not being direct consistency increasing or maintaining, respectively. 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 a 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. We start by introducing *direct consistency maintaining* transformations. Let  $t : G \Rightarrow H$  be a transformation.

Intuitively, formulas (4.1) and (4.2) check that no new violations of the first two unsatisfied layers are introduced, that means, every occurrence  $p$  of  $C_{\text{k}_{\max}(c, G)+2}$  that is not destroyed by  $t$  and satisfies  $\exists C'$  still satisfies  $\exists C'$  in  $H$  with the intermediate graph  $C' \in \text{IG}(C_{\text{k}_{\max}(c, G)+2}, C_{\text{k}_{\max}(c, G)+3})$ . To check that  $p$  has not been destroyed by  $t$ , we use the notion of total morphisms, since it has been shown that  $\text{tr}_t \circ p$  is total if and only if  $p$  has not been destroyed by  $t$  [3]. Additionally, every newly inserted occurrence of  $C_{\text{k}_{\max}(c, G)+2}$  satisfies  $d$  with  $d = \text{false}$  if  $\text{sub}_{\text{k}_{\max}(c, G)+2}(C) = \text{false}$  and  $d = \exists C_{\text{k}_{\max}(c, G)+3}$  otherwise. This case discrimination arises as a consequence of the fact that conditions in UANF are also allows to end with a condition of the form  $\forall(a : C_0 \hookrightarrow C_1, \text{false})$ . Formulas (4.3) and (4.4) ensure that the satisfaction at layer has not been decreased by checking that no occurrences of universally bound graphs  $C_j$  with  $j < \text{k}_{\max}$  have been inserted and that no occurrences of existentially bound graphs  $C_j$  with  $j \leq \text{k}_{\max}$

have been destroyed. Only in these cases, the satisfaction at layer can be decreased. Of course, this does not always lead to a decrease of satisfaction at layer but it can always be considered as an unnecessary insertion or deletion.

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 4.15 (direct consistency maintaining transformations).** Let  $G$  be a graph  $\rho$  a rule and  $c$  a constraint in UANF. If  $G \models c$ , a transformation  $t : G \Rightarrow_{\rho, m} H$  is called direct consistency maintaining w.r.t.  $c$  if  $H \models c$ . Otherwise, if  $G \not\models c$ , let  $k = k_{\max}(c, G) + 2$ ,  $e = \text{sub}_k(c)$ . A transformation  $t : G \Rightarrow_{\rho, m} H$  is called direct consistency maintaining w.r.t.  $c$  if the following equations hold.

1. If  $e \neq \text{false}$ , every occurrence of  $C_k$  in  $G$  that satisfies  $\text{IC}_0(e, C')$  for any  $C' \in \text{IG}(C_k, C_{k+1})$  still satisfies  $\text{IC}_0(e, C')$  in  $H$ .

$$\forall p : C_k \hookrightarrow G \left( \bigwedge_{C' \in \text{IG}(C_k, C_{k+1})} (p \models \text{IC}_0(e, C') \wedge \text{tr}_t \circ p \text{ is total}) \implies \text{tr}_t \circ p \models \text{IC}_0(e, C') \right) \quad (4.1)$$

Otherwise, if  $e = \text{false}$ , (4.1) is set equal to true.

2. Let  $d = \text{IC}_0(e, C_{k+1})$  if  $e \neq \text{false}$  and  $d = \text{false}$  otherwise. Every newly inserted occurrence of  $C_k$  satisfies  $d$ .

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

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

$$\bigwedge_{\substack{i < k_{\max} \\ C_i \text{ universally}}} \forall p : C_i \hookrightarrow H (\exists p' : C_i \hookrightarrow G (p = \text{tr}_t \circ p')) \quad (4.3)$$

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

$$\bigwedge_{\substack{i \leq k_{\max} \\ C_i \text{ existentially}}} \forall p : C_i \hookrightarrow G (\text{tr}_t \circ p \text{ is total}) \quad (4.4)$$

Before we continue with the definition of direct consistency increasing, let us first show that each direct consistency maintaining transformation is indeed consistency maintaining. For this, we start by showing that the satisfaction of formulas (4.3) and (4.4) guarantee that the satisfaction at layer has not been decreased.

**Lemma 4.16.** Let a transformation  $t : G \Rightarrow H$  and a constraint  $c$  in UANF be given, such that (4.3) and (4.4) 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, i.e.  $C_i$  is universally and  $C_j$  is existentially bound. It follows immediately that either (4.3) or (4.4) is not satisfied. This is a contradiction.  $\square$

With this, we will now show that direct consistency maintaining transformation is also consistency maintaining.

**Theorem 4.3.** *Let a graph  $G$ , a constraint  $c$  in UANF, a rule  $\rho$  and a direct consistency maintaining transformation  $t : G \implies_{\rho, m} H$  w.r.t.  $c$  be given. Then,  $t$  is also a consistency maintaining transformation.*

*Proof.* With Lemma 4.16 follows that  $k_{\max}(c, G) \leq k_{\max}(c, H)$  and it also holds that  $nvc_{k_{\max}(c,G)+1}(H) \neq \infty$ . It remains to show that  $nvc_k(H) \leq nvc_k(G)$  for all  $0 \leq k < \text{nl}(c)$ . In particular, we only need to show that  $nvc_{k_{\max}(c,G)+1}(H) \leq nvc_{k_{\max}(c,G)+1}(G)$  since for all  $0 \leq j < k_{\max}(c, G) + 1$  it holds that  $nvc_j(H) = nvc_j(G) = 0$ . Also, since  $nvc_j(G) = \infty$  for all  $k_{\max}(c, G) + 1 < j < \text{nl}(c)$  it follows that  $nvc_j(H) \leq nvc_j(G)$ .

Let  $k = k_{\max}(c, G) + 1$  and  $d = \text{sub}_{k+1}(c)$ . We show that (4.1) and (4.2) imply that  $nvc_k(H) \leq nvc_k(G)$ . Assume that  $nvc_k(H) > nvc_k(G)$ .

Therefore, a morphism  $p : C_{k+1} \hookrightarrow H$  with  $p \not\models \text{IC}_0(d, C')$  for any  $C' \in \text{IG}(C_{k+1}, C_{k+2})$  exists, such that either 1 or 2 is satisfied. Note that this is only the case if  $d \neq \text{false}$ . Otherwise, a morphism  $p$  satisfying 2 must exist.

1. There does exist a morphism  $q' : C_k \hookrightarrow G$  with  $q' \models \text{IC}_0(d, C')$  and  $p = \text{tr}_t \circ q'$ .
2. There does not exist a morphism  $q : C_k \hookrightarrow G$  with  $p = \text{tr}_t \circ q$ .

This is a contradiction, if 1 is satisfied,  $q'$  does not satisfy equation (4.1) and if (2) is satisfied  $q$  does not satisfy equation (4.2) since  $q$  only satisfies  $\text{IC}_0(d, C_{k+2})$  if  $q$  satisfies  $\text{IC}_0(d, C')$  for all  $C' \in \text{IG}(C_{k+1}, C_{k+2})$ . It follows that

$$nvc_k(H) \leq nvc_k(G).$$

Therefore,  $t$  is a consistency maintaining transformation.  $\square$

Let us now introduce the notion of *direct consistency increasing* transformations. Similar to the definition of consistency maintaining and increasing transformation, again this notion is based on the notion of direct consistency maintaining transformations, in the sense that a direct consistency increasing transformation is also a direct consistency maintaining one. Since a direct consistency maintaining transformation  $t$  does not insert

any new violations it is sufficient that  $t$  deletes at least one violation to state that  $t$  is direct consistency increasing. For this, (4.5) checks that either an occurrence  $p$  of  $C_{k_{\max}+2}$  has been deleted or  $\text{tr}_t op \models \exists C'$  for an intermediate graph  $C' \in \text{IG}(C_{k_{\max}+2}, C_{k_{\max}+3})$  such that  $p \not\models \exists C'$ .

**Definition 4.17 (direct consistency increasing).** Let a graph  $G$ , a rule  $\rho$  and a constraint  $c$  in UANF with  $G \not\models c$  be given. Let  $e = \text{sub}_{k_{\max}+2}(c)$  and

A transformation  $t : G \Rightarrow_{\rho, m} H$  is called direct consistency increasing w.r.t.  $c$  if it is direct consistency maintaining w.r.t.  $c$  and the following holds:

$$\exists p : C_{k_{\max}+2} \hookrightarrow G(d) \quad (4.5)$$

with

$$d := \begin{cases} \text{tr}_t op \text{ is not total} \\ \bigvee_{C' \in \text{IG}(k_{\max}+2, k_{\max}+3)} (p \not\models \text{IC}_0(e, C') \wedge (\text{tr}_t op \text{ is not total} \vee \text{tr}_t op \models \text{IC}_0(e, C'))) \end{cases} \quad \begin{array}{l} \text{if } e = \text{false} \\ \text{otherwise.} \end{array}$$

Note that (4.3) and (4.4) not only secure that the satisfaction up to layer does not decrease, as shown in Lemma 4.16, 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.

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.

**Theorem 4.4.** Let a graph  $G$ , a constraint  $c$  in UANF with  $G \not\models c$ , 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.* With Theorem 4.3 follows that  $t$  is a consistency maintaining transformation. Therefore, it is sufficient to show that  $\text{nvc}_{k_{\max}(c, G)+1}(H) < \text{nvc}_{k_{\max}(c, G)+1}(G)$ . Let  $k = k_{\max}(c, G) + 2$  and  $d = \text{sub}_k(c)$  with  $d \neq \text{false}$ .

Since (4.5) is satisfied, a morphism  $p : C_k \hookrightarrow G$  with  $p \not\models \text{IC}_0(d, C')$ , such that either  $\text{tr}_t op$  is total and  $\text{tr}_t op \models \text{IC}_0(d, C')$  or  $\text{tr}_t op$  is not total exists, for a graph  $C' \in \text{IG}(C_k, C_{k+1})$ . In both cases the following holds:

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

Since  $t$  is direct consistency maintaining 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  $C \in \text{IG}(C_k, C_{k+1})$ . Additionally, this inequality is strictly satisfied if  $C = C'$ . Therefore, it follows that  $\text{nvc}_k(H) < \text{nvc}_k(G)$  and  $t$  is a consistency increasing transformation.

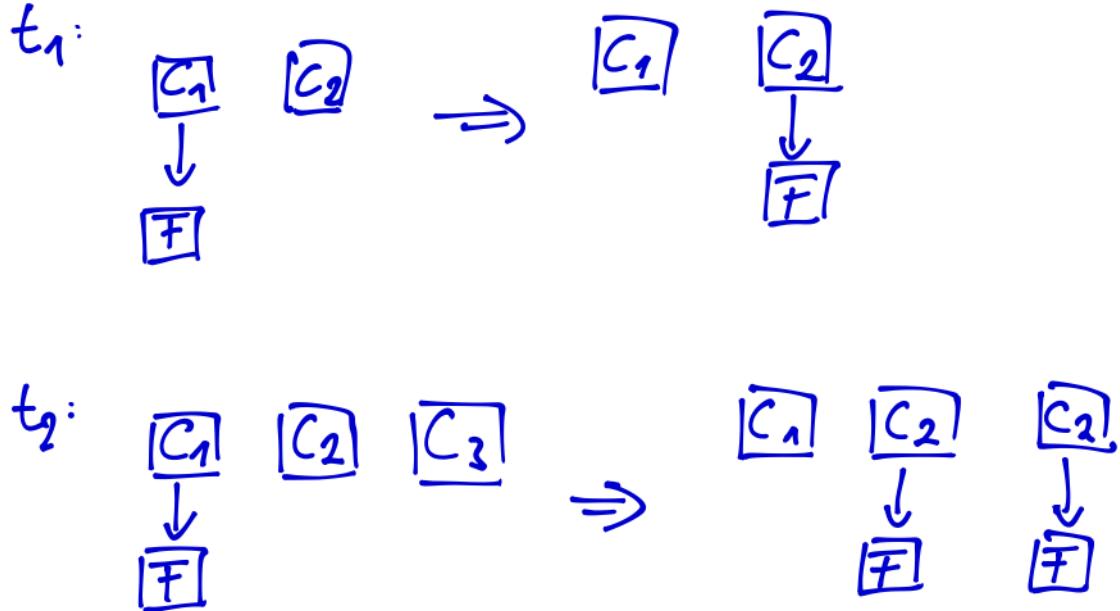


Figure 4: example

If  $d = \text{false}$  it holds that

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

since  $t$  is a direct consistency maintaining transformation and it can be shown in a similar manner as above that (4.5) implies

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

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

**Example 4.6.** Consider the transformations  $t_1$  and  $t_2$  given in Figure 4 and constraint  $c_1$  given in Figure 2. Then,  $t_1$  is a consistency maintaining transformation since the number of violations in both graphs is equal to 2. But,  $t_1$  is not a direct consistency maintaining transformation since one occurrence of a node of type `Class` satisfying  $\exists C_2^2$  in the first but not in the second graph of the transformation exists. Therefore (4.1) is not satisfied.

The transformation  $t_2$  is consistency increasing w.r.t.  $c_1$  since the number of violations is equal to 4 in the first and equal to 2 in the second graph. But,  $t_2$  is not a direct consistency increasing transformation since (4.1) is not satisfied.

#### 4.5 Comparison with other concepts of Consistency

In this chapter, the notions of (direct) consistency increase- and maintenance are compared to the already known concepts of consistency guaranteeing, consistency preserving

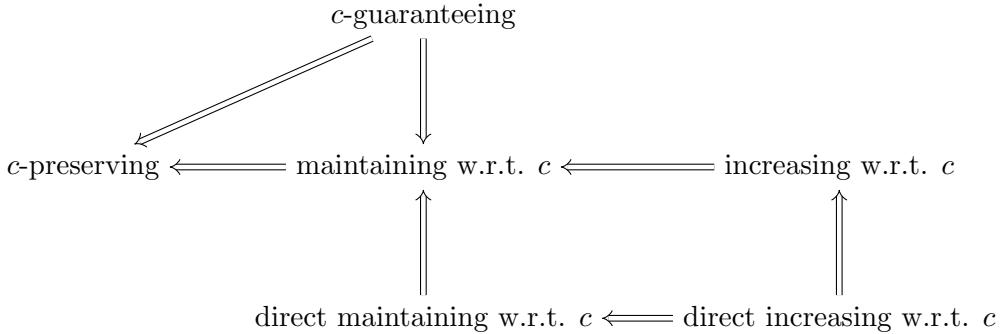


Figure 5: Relations of consistency notions.

[2], (direct) consistency increasing and sustaining [3] in order to reveal relationships between them and ensure that (direct) consistency increase- and maintenance are indeed a new notions of consistency. These relations are displayed in Figure 5.

First, we compare (direct) consistency increase- and maintenance with the notions of consistency-guaranteeing and -preserving. We start by showing that consistency maintaining implies preserving but the backwards implication does not hold.

**Lemma 4.18.** *Let a constraint  $c$  in UANF, graphs  $G$  and  $H$  and a transformation  $t : G \Rightarrow H$  be given. Then,*

$$\begin{aligned} t \text{ is maintaining w.r.t. } c &\implies t \text{ is } c\text{-preserving} && \text{and} \\ t \text{ } c\text{-preserving} &\not\implies t \text{ is maintaining w.r.t. } c \end{aligned}$$

*Proof.* 1. Let  $t$  be a consistency maintaining transformation w.r.t.  $c$ . If  $G \not\models c$ ,  $t$  is a  $c$ -preserving transformation. If  $G \models c$ , it holds that  $nvc_j(G) = 0$  for all  $0 \leq j < nl(c)$ . Since  $t$  is consistency maintaining it follows that  $nvc_j(H) = 0$  for all  $0 \leq j < nl(c)$  and therefore  $H \models c$ . It follows that  $t$  is a  $c$ -preserving transformation.

2. Consider graphs  $C_1^1, C_2^2$  and constraint  $c_1$  given in Figure 2. Then, the transformation  $t : C_2^2 \Rightarrow C_1^1$  is  $c$ -preserving, since  $C_2^2 \not\models c_1$ , but not consistency maintaining w.r.t.  $c$  since  $nvc_0(C_2^2) = 2$  and  $nvc_0(C_1^1) = 5$ .

□

Obviously, guaranteeing implies consistency maintaining, since this property is embedded in the definition of consistency maintenance. The inversion of this implication does not hold, since maintaining is a way stricter notion, in the sense, that the number of removed violations has to be greater or equal than the number of introduced violations. For guaranteeing transformations this is not the case, an arbitrary number of violations can be inserted, as 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.



Figure 6: example

**Lemma 4.19.** Let a constraint  $c$  in UANF, graphs  $G$  and  $H$  and a transformation  $t : G \Rightarrow H$  be given. Then,

$$\begin{aligned} t \text{ is } c\text{-guaranteeing} &\implies t \text{ is maintaining w.r.t } c \quad \text{and} \\ t \text{ is } c\text{-guaranteeing} &\nRightarrow t \text{ is direct maintaining w.r.t } c \text{ and} \\ t \text{ is maintaining w.r.t. } c &\nRightarrow t \text{ is } c\text{-guaranteeing} \end{aligned}$$

- Proof.*
1. Let  $t$  be a  $c$ -guaranteeing transformation, then  $H \models c$ . It holds that  $nvc_j(H) = 0$  for all  $0 \leq j < nl(c)$  and since the number of violations cannot be negative,  $t$  is a maintaining transformation w.r.t.  $c$ .
  2. Let  $t$  be a  $c$ -guaranteeing transformation. Then,  $t$  is also a  $c$ -preserving transformation and since each direct maintaining transformation is also a maintaining one the statement follows directly with Lemma 4.18.
  3. Consider the transformation  $t : G \Rightarrow H$  shown in Figure 6 and constraint  $c_1$  shown in Figure 2. Then,  $t$  is a maintaining transformation w.r.t  $c$ , in particular,  $c$  is a consistency increasing transformation w.r.t.  $c$  since  $nvc_0(G) = 10$  and  $nvc_0(H) = 7$ . But,  $t$  is not  $c$ -guaranteeing since both occurrences of nodes of type Class do not satisfy  $\exists C_2^1$ . □

Let  $t : G \Rightarrow H$  be a  $c$ -guaranteeing transformation. If  $G \models c$ , the transformation is, by definition, not consistency increasing. If  $G \not\models c$ ,  $t$  is always also a consistency increasing transformation w.r.t.  $c$ .

**Lemma 4.20.** Let graphs  $G, H$ , a constraint  $c$  with  $G \not\models c$  and a transformation  $t : G \Rightarrow H$  be given. Then

$$t \text{ is } c\text{-guaranteeing} \implies t \text{ is increasing w.r.t } c.$$

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

The definition of consistency improvement only differs from guaranteeing if the corresponding constraint is universally bound and these notions are identical for existentially bound constraint. Therefore, with Lemmas 4.19 and 4.20, we can state the following.

**Corollary 4.21.** *Let an existentially bound constraint  $c$  in ANF and a transformation  $t : G \Rightarrow H$  be given. Then,*

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

If  $G \not\models c$ , we can also state that

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

The notions of increase- and improvement are equivalent for universally bound constraints with nesting level 1. Note that, with corollary 4.21, this equivalence does not hold for existentially bound constraints with nesting level 1.

**Lemma 4.22.** *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  $\text{sub}_1(c) = \text{false}$ ,  $\text{nvc}_0(G)$  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.  $\square$

For universally bound constraints  $c$  with  $\text{nl}(c) \geq 2$ , the notions of (direct) consistency increase- and maintainment are not related to (direct) consistency improve- and sustainment. By definition, (direct) consistency improvement implies (direct) consistency sustainment [3]. Therefore, it is sufficient to show that direct improvement does not imply maintainment and that direct increasement does not imply consistency sustainment.

**Lemma 4.23.** *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,*

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

*Proof.* 1. Consider transformation  $t_1$  given in Figure 7 and constraint  $c_1$  given in Figure 2. The transformation  $t_1$  is direct consistency improving but not maintaining since  $\text{nvc}_0(G) = 4$  and  $\text{nvc}_0(H) = 5$ .

2. Consider the constraint  $c = \forall(C_1^1, \exists(C_2^1, \forall(C_4^2, d)))$  with  $d$  being an existentially bound constraint in ANF with  $d \neq \text{false}$  composed of the graphs given in Figure

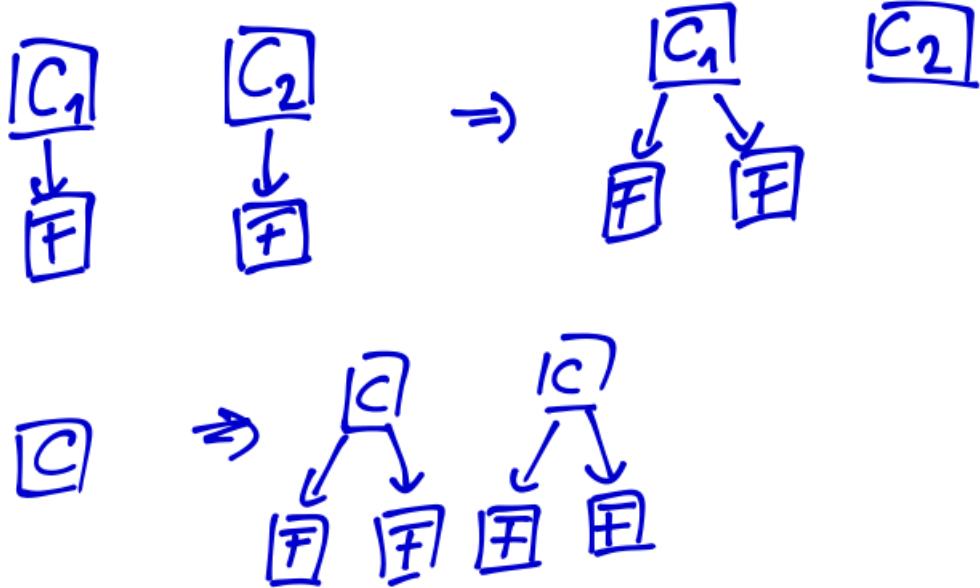


Figure 7: example

2 and transformation  $t_2$  given in Figure 7. Then,  $t$  is direct consistency increasing, (4.1), (4.2), (4.3) and (4.4) are trivially satisfied and (4.5) is satisfied since one occurrence of  $C_1^1$  that did not satisfy  $\exists C_2^1$  in  $G$  satisfies  $\exists C_2^1$  in  $H$ , but not consistency sustaining since the number of occurrences of  $C_1^1$  that not satisfying  $\exists(C_2^1, \forall(C_4^2, d))$  in  $H$  is greater than the number of occurrences of  $C_1^1$  in  $G$  not satisfying  $\exists(C_2^1, \forall(C_4^2, d))$ .

□

## 5 Application Conditions

To guarantee that each transformation  $t$  is (direct) consistency increasing or maintaining w.r.t to a constraint  $c$ , we present applications conditions ensuring this property. Given a constraint  $c$ , this application conditions are designed to only consider graph of  $c$  up to a certain layer. In particular, this is useful to reduce the restrictiveness of these application conditions, since all graphs  $C_j$  of  $c$  with  $j > k_{\max} + 2$  do not affect whether an transformation is considered as consistency maintaining or increasing. Additionally, in the case of consistency increasing application conditions, this design is necessary since it has to be ensured that violations at layer  $k_{\max} + 1$  will be removed. Therefore, let us introduce a weaker notion of consistency increasing and maintaining rules, namely *consistency increasing and maintaining rules at layer*. As the name suggests, given a constraint  $c$ , a rule is consistency increasing or maintaining at layer  $0 \leq k < nl(c)$  if all of its applications at graphs  $G$ , with  $k_{\max} = k$ , are consistency increasing or maintaining w.r.t.  $c$ , respectively.

- delete
- insert

move Feature :



assign Feature :



Figure 8: rules

$$\begin{aligned}
 & (\exists (C \ F_1) \hookrightarrow C \ F_1, \ \forall \exists (C \ F_1) \hookrightarrow C \overset{c}{\hookleftarrow} F_1, \text{true}) \\
 & \wedge \ \exists (C \ F_1) \hookrightarrow C \ F_1, \ \forall \exists (C \ F_1) \hookrightarrow C \underset{F_1}{\overset{c}{\downarrow}} \ F_1, \text{true}) \\
 & \wedge \ \exists (C \ F_1) \hookrightarrow C \underset{F_1}{\overset{c}{\downarrow}} \ F_1, \text{true}
 \end{aligned}$$

Figure 9: application condition

**Definition 5.1 ((direct) consistency increasing and maintaining rule at layer).** Let a constraint  $c$  be given. A rule  $\rho$  is called (direct) consistency maintaining at layer  $-1 \leq k < \text{nl}(c)$  w.r.t.  $c$  if all transformations  $t : G \Rightarrow_{\rho} H$ , with  $\text{k}_{\max}(G, c) = k$ , are (direct) consistency maintaining w.r.t.  $c$ . Additionally,  $\rho$  is called (direct) consistency increasing at layer  $0 \leq k < \text{nl}(c)$  w.r.t.  $c$  if all transformations  $t : G \Rightarrow_{\rho} H$ , with  $\text{k}_{\max}(G, c) = k$ , are (direct) consistency increasing w.r.t.  $c$ .

Note, that a consistency maintaining rule at layer  $\text{nl}(c)-1$  w.r.t.  $c$  is also a consistency maintaining or increasing rule w.r.t.  $c$ . But, a consistency increasing rule at layer  $\text{nl}(c)-1$  must not necessarily be a consistency increasing rule w.r.t.  $c$ .

## 5.1 General Application Conditions

We start by introducing consistency maintaining application conditions, i.e. a rule equipped with this application condition will only be applicable if the corresponding transformation is consistency maintaining. In particular, we will show that this transformation is even direct consistency maintaining.

The maintaining application condition consists of the three parts  $\text{ned}_k(\rho')$ ,  $\text{nui}_k(\rho')$  and  $\text{nds}_k(\rho')$  that are connected via the boolean  $\wedge$ -operator with  $\rho'$  being a plain rule and  $0 \leq k < \text{nl}(c)$  for a given constraint  $c$ . As already discussed, the satisfaction at layer is decreased if and only if occurrences of existentially bound graphs  $C_j$  have been deleted or occurrences of universally bound graphs  $C_j$  have been inserted, with  $0 \leq j \leq k$ . To check that no existentially bound graph  $C_j$ ,  $0 \leq j \leq k$ , will be deleted, we use the condition constructed by  $\text{ned}_k(\rho')$  that checks that no overlap  $P$  of the left-hand-side of  $\rho'$  with  $C_j$ , such that the application of  $\rho'$  at  $P$  leads to a deletion of  $C_j$ , exists in the originating graph. To ensure that no universally bound graph  $C_j$ ,  $0 \leq j \leq k$ , will be inserted, the condition constructed by  $\text{nui}_k(\rho')$  checks that no overlap  $P$  of the right-hand-side of  $\rho'$  and  $C_j$ , such that the application of  $\rho'^{-1}$  at  $P$  would destroy the occurrence of  $C_j$ , exists in the derived graph. Note, that a condition constructed in this way is a right application condition. Therefore, we use the shift over rule operator to construct an equivalent left application condition. To verify that the number of violations is not decreased, we use the condition constructed by  $\text{nds}_k(\rho')$ , that checks, in the same manner as  $\text{nui}_k(\rho')$ , that no occurrences of graphs  $C' \in \text{IG}(C_{k+2}, C_{k+3})$  will be deleted if  $k < \text{nl}(c) - 2$ . Otherwise, this condition is set to true.

**Definition 5.2 (consistency maintaining application condition).** Let a rule  $\rho = (\text{ac}, \rho')$  with  $\rho' = L \longleftrightarrow K \longleftrightarrow R$  and a constraint  $c$  in UANF be given. The maintaining application condition of  $c$  for  $\rho$  at layer  $-1 \leq k < \text{nl}(c)$  is defined as  $\text{ac} \wedge \text{main}_k(\rho')$  with

$$\text{main}_k(\rho') := \text{ned}_k(\rho') \wedge \text{nui}_k(\rho') \wedge \text{nds}_k(\rho')$$

and

1. Let  $E$  be the set of all existentially bound graphs  $C_j$  with  $j \leq k+1$  and  $\mathbf{P}_{C_j}$

be the set all overlaps  $P'$  of  $L$  and  $C_j$  with  $i_L^{P'}(L \setminus K) \cap i_{C_j}^{P'}(C_j) \neq \emptyset$ :

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

2. Let  $U$  be the set of all universally bound graphs  $C_j$  with  $j \leq k+2$ , and  $\mathbf{P}_{C_j}$  be the set of all overlaps  $P'$  of  $R$  and  $C_j$  with  $i_R^{P'}(R \setminus K) \cap i_{C_j}^{P'}(C_j) \neq \emptyset$ :

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

3. If  $k > \text{nl}(c) - 3$  or  $\text{sub}_k(c)$  is universally bound,  $\text{nds}_k(\rho') = \text{true}$ . Otherwise, Let  $E$  be the set of all overlaps of  $L$  and  $C'$  with  $i_L^{P'}(L \setminus K) \cap i_{C'}^{P'}(C') \neq \emptyset$  for all  $C' \in \text{IG}(C_{k+2}, C_{k+3})$ :

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

### Example 5.1.

Let us now show that each rule equipped with the according application condition is a consistency maintaining rule at layer.

**Theorem 5.1.** Let a constraint  $c$  in UANF be given. Each rule  $\rho = (\text{ac}', \rho')$  with  $\text{ac}' = \text{ac} \wedge \text{main}_k(\rho')$  and  $-1 \leq k < \text{nl}(c)$  is a consistency maintaining rule at layer  $k$  w.r.t.  $c$ .

*Proof.* Let a graph  $G$ , such that  $k_{\max} = k$ , and a transformation  $t : G \Rightarrow_{\rho} H$  be given. It is sufficient to show that  $t$  is a direct consistency maintaining transformation. If  $k < \text{nl}(c) - 1$ , it follows that  $G \not\models c$  and that  $k$  is odd. We show that  $t$  satisfies (4.1), (4.2), (4.3) and (4.4).

1. Assume that (4.1) does not hold. Then,  $e = \text{sub}_{k+2}(c) \neq \text{false}$  and a morphism  $p : C_{k+2} \hookrightarrow G$  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  $C' \in \text{IG}(C_{k+2}, C_{k+3})$ . Therefore, an overlap  $P$  of  $L$  and  $C'$  such that  $i_{C_{k+2}}^P \models \exists(a_{k+2}^r : C_{k+2} \hookrightarrow C', \text{true})$  with  $i_L^P(L \setminus K) \cap i_{C'}^P(C' \setminus C_{k+2}) \neq \emptyset$  must exist and  $m \models \exists(i_L^P : L \hookrightarrow P, \text{true})$  holds. Thus,  $\text{nds}_k(\rho')$  and consequently also  $\text{main}_k(\rho')$  cannot be satisfied.
2. Assume that (4.2) does not hold and let

$$d := \begin{cases} \text{IC}_0(\text{sub}_{k+2}(c), C_{k+3}) & \text{if } \text{sub}_{k+2}(c) \neq \text{false} \\ \text{false} & \text{otherwise.} \end{cases}$$

Then, a morphism  $p' : C_{k+2} \hookrightarrow H$  with  $p' \not\models d$  exists, such that no morphism  $p : C_{k+2} \hookrightarrow G$  with  $\text{tr}_t \circ p = p'$  exists. Therefore, an overlap  $P$  of  $R$  and  $C_{k+2}$  with  $i_R^P(R \setminus K) \cap i_{C_{k+2}}^P(C_{k+2}) \neq \emptyset$  exists, such that  $m \models \text{Left}(\exists(i_R^P : R \hookrightarrow P, \text{true}), \rho')$ . Hence,  $m$  does not satisfy  $\text{nui}_k(\rho')$ .

3. Assume that (4.3) 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,  $m \not\models \text{nui}_k(\rho')$ .
4. Assume that (4.4) does not hold. Then, a morphism  $p : C_j \hookrightarrow G$  with  $C_j$  being existentially bound and  $j \leq k$  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,  $m \not\models \text{ned}_k(\rho')$ .

If  $k = \text{nl}(c) - 1$ , it follows immediately that  $G \models c$ . Therefore, we need to show that  $H \models c$ . It follows that  $\text{nds}_k(\rho') = \text{true}$ . Assume that  $H \not\models c$ . Therefore, either an occurrence  $p : C_j \hookrightarrow H$  of an universally bound graph  $C_j$  exists such that no  $q : C_j \hookrightarrow G$  with  $p = \text{tr}_t \circ q$  exists or an occurrence  $p' : C_{j'} \hookrightarrow G$  of an existentially bound graph  $C_{j'}$  exists, such that  $\text{tr}_t \circ p'$  is not total. If the first case applies, an overlap  $P$  of  $C_j$  and  $R$  with  $i_{C_j}^P(C_j) \cap i_R^P(R \setminus K) \neq \emptyset$  exists, such that  $m \models \text{Left}(\neg\exists(i_R^P : R \hookrightarrow P, \text{true}), \rho')$  and therefore  $m \not\models \text{nui}_k(\rho')$ . If the second case applies, an overlap  $P$  of  $C_{j'}$  and  $L$  with  $i_{C_{j'}}^P(C_{j'}) \cap i_L^P(L \setminus K) \neq \emptyset$  exists, such that  $m \models \neg\exists(i_L^P : L \hookrightarrow P, \text{true})$  and therefore  $m \not\models \text{ned}_k(\rho')$ . By contradiction follows that  $H \models c$ .

In total follows that  $\rho$  is a consistency maintaining rule at layer  $k$  w.r.t.  $c$ .  $\square$

For an application condition, such that a rule equipped with it is consistency increasing at layer, we have additionally to ensure that at least one violation will be removed. To check this via an application condition, it is necessary (a) to check that an occurrence  $p$  of the universally bound graph  $C_{k_{\max}+2}$  exists, such that  $p$  and the match  $m$  do overlap, i.e  $p(C_{k_{\max}+2}) \cap m(L) \neq \emptyset$ , and, if the sub-condition at layer  $k_{\max}+2$  is not equal to  $\text{false}$ , (b) that  $p$  does not satisfy  $c' := \exists C_{k_{\max}+3}$ . 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  $P$  of  $L$  and  $C_{k_{\max}+3}$  such that  $p \models c'$  has to be checked. For this, we introduce *extended overlaps*. Intuitively, given an overlap  $C$  of  $L$  and  $C_{k_{\max}+2}$  with  $p : C_{k_{\max}+2} \hookrightarrow C$ , the overlap is extended with elements of  $C_{k_{\max}+3}$  such that  $p \models C_{k_{\max}+3}$ .

**Definition 5.3 (extended overlaps).** Let graphs  $C_0$ ,  $C_1$ ,  $G$  and morphisms  $i_{C_0}^G : C_0 \hookrightarrow G$  and  $i_{C_0}^{C_1} : C_0 \hookrightarrow C_1$  be given. The set of extended overlaps of  $G$  with  $i_{C_0}^G$  and  $i_{C_0}^{C_1}$ , denoted by  $\text{eol}(G, i_{C_0}^G, i_{C_0}^{C_1})$  is defined as:

$$\text{eol}(G, i_{C_0}^G, i_{C_0}^{C_1}) := \{P \in \text{ol}(G, C_1) \mid i_G^P \circ i_{C_0}^G \models \exists(i_{C_0}^{C_1} : C_0 \hookrightarrow C_1, \text{true})\}$$

Via the notion of extended overlaps we are now able to check that a violation exists, to decide whether a transformation is able to remove a violation at all. It remains to check whether such a violation will be removed.

In the definition below,  $\text{exv}(\cdot, \cdot)$  and  $\text{remv}(\cdot, \cdot)$  ensure that a violation will be removed, with  $\text{exv}(\cdot, \cdot)$  ensuring that a violation is present and  $\text{remv}(\cdot, \cdot)$  ensuring that this violation will be removed. Note, that the construction of these is divided in two cases. Firstly, either  $k \leq \text{nl}(c) - 3$  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  $\text{exv}(\cdot, \cdot)$ , the first case will be checked via extended overlaps as already described above. In the second case, it is sufficient to check whether an occurrence  $p$  of  $C_{\text{nl}(c)}$  with  $m(L) \cap p(C_{\text{nl}(c)}) \neq \emptyset$  exists.

For  $\text{remv}(\cdot, \cdot)$ , in the first case, a violation can be removed by either deleting an occurrence  $p$  of  $C_{k+2}$  or inserting elements of  $C_{k+3}$ , such that  $p \not\models \exists C'$  and  $\text{tr}_t \circ p \models \exists C'$  for a graph  $C' \in \text{IG}(C_k k + 2, C_{k+3})$ . In the second case, a violation can only be removed by deleting an occurrence  $p$  of  $C_1$ . This will only occur if  $m(L \setminus K) \cap p(C_1) \neq \emptyset$ .

**Definition 5.4 (consistency increasing application condition).** Let a rule  $\rho = (\text{ac}, \rho')$  with  $\rho' = L \hookleftarrow K \hookrightarrow R$  and a constraint  $c$  in UANF be given. Let  $0 \leq k < \text{nl}(c)$  be even, i.e.  $\text{sub}_k(c)$  is universally bound, and  $C \in \text{IG}(C_{k+1}, C_{k+2})$  if  $\text{sub}_{k+1}(c) \neq \text{false}$  and  $C = C_{k+1}$  otherwise. The increasing application condition of  $c$  for  $\rho$  at layer  $k$  with  $C$  is defined as

$$\text{incr}_k(C, \rho) := \text{ac} \wedge \text{main}_{k-1}(\rho) \wedge \left( \bigvee_{P \in \text{ol}(L, C_{k+1})} \text{exv}(P, C) \wedge \text{remv}(P, C) \right) \quad (5.1)$$

with

- Let  $a^r : C_{k+1} \hookrightarrow C$  be the 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:

$$\text{exv}(P, C') := \begin{cases} \exists(i_L^P : L \hookrightarrow P, \text{true}) & \text{if } \text{sub}_{k+1}(c) = \text{false} \\ \exists(i_L^P : L \hookrightarrow P, \bigwedge_{Q \in \text{ol}(P, i_{C_{k+1}}^P, a^r)} \neg \exists(i_P^Q : P \hookrightarrow Q, \text{true})) & \text{otherwise} \end{cases}$$

- 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  $C_{k+1}$ . If this transformation does not exist, we set  $\text{remv}(P, C') := \text{false}$ . Let  $a^r : C_{k+1} \hookrightarrow C$  be the restricted morphism of  $a_{k+1}$ , then

$$\text{remv}(P, C') := \begin{cases} \text{false} & \text{if } \text{sub}_{k+2}(c) = \text{false} \\ \bigvee_{Q \in \text{ol}(P', i_{C_{k+1}}^{P'}, a^r)} \text{Left}(\forall(i_R^Q : R \hookrightarrow P', \exists(i_{P'}^Q : P' \hookrightarrow Q, \text{true})), \rho) & \text{otherwise.} \end{cases}$$

Note, that, in case that no occurrence of  $C_{k+1}$  not satisfying  $\exists C_{k+2}$  will be removed,  $\text{incr}_k(C', \rho)$  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  $\text{tr}_t \circ p \models \exists(a_{k+1}^r : C_{k+1} \hookrightarrow C', \text{true})$  exists. For any smaller improvements, i.e a similar improvement for a sub-graph  $C'' \in \text{IG}(C_{k+1}, C_{k+2})$  of  $C'$ ,  $\text{incr}_k(C', \rho)$  would be evaluated with `false`. For any bigger improvements, i.e the same improvement for a super-graph  $C'' \in \text{IG}(C_{k+1}, C_{k+2})$  of  $C'$ ,  $\text{incr}_k(C', \rho)$  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

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

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 `main()` can be removed, since they are identical for each  $\text{incr}_k(C', \rho)$  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`; the maintaining part of the condition 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 `main()` 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.

**Example 5.2.** Consider the rule `assignFeature` of figure 8 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 5.4 is shown in figure 9. 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 5.2.** *Let a constraint  $c$  in UANF be given. Each rule  $\rho = (\text{ac}', \rho')$  with  $\text{ac}' = \text{ac} \wedge \text{incr}_k(C, \rho)$  with  $0 \leq k < \text{nl}(c)$  being even and  $C = Ck + 1$  if  $\text{sub}_{k+1}(c) = \text{false}$  and  $C \in \text{IG}(C_{k+1}, C_{k+2})$  otherwise, Then,  $\rho$  is a consistency increasing rule at layer  $k - 1$  w.r.t.  $c$ .*

*Proof.* Let a transformation  $t : G \implies_\rho H$  with  $\text{k}_{\max}(c, G) = k - 1$  be given. Since  $\text{main}_{k-1}(\rho)$  is contained in  $\text{incr}_k(C, \rho)$ ,  $t$  is a consistency maintaining transformation at layer  $k - 1$  with Theorem 5.1. It remains to show that  $t$  satisfies (4.5).

Let  $\text{sub}_k(c) = \forall(a_k : C_k \hookrightarrow C_{k+1}, e)$  be the sub-condition of  $c$  at layer  $k$ .

1. If  $e = \text{false}$ , assume that (4.5) does not hold, then, no morphism  $p : C_{k+1} \hookrightarrow G$  exists, such that  $\text{tr}_t \circ p$  is not total. Therefore, 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. It follows that  $\text{remv}(P, C') = \text{false}$  and  $m \not\models \text{incr}_k(C, \rho)$
2. Otherwise, let  $P \in \text{ol}(L, C_{k+1})$ . We show that  $m \models \text{exv}(P, C) \wedge \text{remv}(P, C)$  implies that (4.5) holds. If  $m \models \text{exv}(P, C)$ , there does exist a morphism  $p : P \hookrightarrow G$  with  $m = p \circ i_L^P$  and  $p \models \neg \exists(i_P^Q : P \hookrightarrow Q, \text{true})$  for all  $Q \in \text{eol}(P, a^r)$ . Therefore,  $q \not\models \exists(a_{k+1} : C_{k+1} \hookrightarrow C_{k+2}, \text{true})$  with  $q = p \circ i_{C_{k+1}}^P$ .

If  $i_L^P(L \setminus K) \cap i_{C_{k+1}}^P(C_{k+1}) \neq \emptyset$ ,  $\text{tr}_t \circ q$  is not total, since all occurrence  $i$  of  $C_{k+1}$  with  $i = p' \circ i_{C_{k+1}}^P$ ,  $p' : P \hookrightarrow G$  and  $m = p' \circ i_L^P$  in  $G$  will be removed by  $t$ . Otherwise,  $\text{tr}_t \circ q$  is total and there does exist a morphism  $p : P' \hookrightarrow H$  such that  $\text{tr}_t \circ q = p \circ i_{C_{k+1}}^{P'}$ . Since  $m \models \text{remv}(P, C)$ , all morphisms  $p \circ i_{C_{k+1}}^{P'}$  with  $n = p \circ i_R^Q$  satisfy  $\text{IC}_0(e, C)$ . Therefore,  $\text{tr}_t \circ q \models \text{IC}_0(e, C)$ . It follows that (4.5) is satisfied.

Therefore,  $\rho$  is a consistency increasing rule at layer  $k - 1$ .  $\square$

## 5.2 Basic Increasing and Maintaining Rules

The construction of the application conditions introduced in the previous section, as well as the constructed application conditions itself, are very complex. For a certain set of rules, which we will call *basic consistency increasing rules*, for which, we are able to construct application conditions with the same property, namely that a rule equipped with this application condition is consistency increasing at layer, in a less complex manner. The main idea is, that these rules are (a) not able to delete occurrences of existentially or insert occurrences of universally bound graphs and (b) are able to increase consistency at a certain layer. That means, given a basic increasing rule  $\rho$ , there does exist a transformation  $t : G \implies_\rho H$  such that  $t$  is a consistency increasing transformation w.r.t to a constraint  $c$ .

To ensure that (a) is met, we firstly introduce *basic consistency maintaining rules up to layer*, which means that, given a constraint, a plain rule is not able to delete existentially bound and insert universally bound graphs up to a certain layer. For the definition, we use the notion of consistency maintaining rules up to layer. The set of basic consistency maintaining rules up to layer is indeed a subset of the set of consistency maintaining rules up to layer since these rules have to be plain rules, whereas consistency maintaining rules up to layer are allowed to be equipped with application conditions, i.e.  $\text{main}(\cdot, \cdot)$ .

**Definition 5.5 (basic consistency maintaining rule up to layer).** *Let a plain rule  $\rho$  and a constraint  $c$  in UANF be given. Then,  $\rho$  is called basic consistency maintaining rule up to layer  $-1 \leq k < \text{nl}(c)$  w.r.t.  $c$  if  $\rho$  is a direct consistency maintaining rule at layer  $k$ .*

**Example 5.3.** Consider the rules `moveFeature` and `assignFeature` given in Figure 8 and constraint  $c_1$  given in Figure 2. The rule `assignFeature` is a basic consistency maintaining rule up to layer 1 w.r.t.  $c$ , whereas `moveFeature` is not a basic consistency maintaining rule.

Since infinite many transformations via a plain rule  $\rho$  exist, it is impossible to check whether  $\rho$  is a basic consistency maintaining rule up to layer based on the definition above. Therefore, we present a characterisation of basic consistency maintaining 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)  $\rho$  deletes 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 via  $\rho$  such that an occurrence of  $C_j$  is deleted by deleting this node since the dangling edge condition would not be satisfied. A rule satisfying this properties does not decrease the satisfaction at layer. Additionally, we have to ensure that the number of violations will not be increased. For this, we have to check that  $\rho$  is not able to insert occurrences of the corresponding universally bound, as described above, and that  $\rho$  is not able to remove occurrence of any intermediate graph. This is only ensured, if  $\rho$  does not remove any elements of  $C'$  if the set of intermediate graphs is given by  $\text{IG}(C, C')$ .

To check that a plain rule satisfies these properties, we make use of the dangling edge condition, or in other words, we check that the rule is not applicable at certain overlaps of  $L$  and an existentially bound graph or that the inverse rule is not applicable at certain overlaps of  $R$  and an universally bound graph.

**Lemma 5.6.** Let a plain rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  and a constraint  $c$  in EANF be given. Then,  $\rho$  is a basic consistency maintaining rule up to layer  $-1 \leq k < \text{nl}(c)$  w.r.t  $c$  if 1 and 2 apply for all  $k$  and 3 applies if  $k \leq \text{nl}(c) - 3$  and  $\text{sub}_k(c)$  is existentially bound, i.e.  $k$  is odd.

1. Let  $E$  be the set of all existentially bound graphs  $C_j$  with  $0 \leq j \leq k+1$ . For each graph  $C \in E$  and each overlap  $P \in \text{ol}(L, C)$  with  $i_L^P(L \setminus K) \cap i_C^P(C) \neq \emptyset$ , the transformation

$$t : P \Longrightarrow_{\rho, i_L^P} H$$

does not exist.

2. Let  $U$  be the set of all universally bound graphs  $C_j$  with  $0 \leq j \leq k+2$ . For each graph  $C \in U$  and each overlap  $P \in \text{ol}(L, C)$  with  $i_R^P(R \setminus K) \cap i_C^P(C) \neq \emptyset$  the transformation

$$t : P \Longrightarrow_{\rho^{-1}, i_R^P} H$$

does not exist.

3. For all graphs  $P \in \text{ol}(L, C_{k+3})$  it holds that

$$i_L^P(L \setminus K) \cap i_{C_{k+3}}^P(C_{k+3} \setminus C_{k+1}) = \emptyset$$

*Proof.* Let  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  be a rule that satisfies the characterisations listed in Lemma 5.6 with  $-1 \leq k < \text{nl}(c)$ . Let us assume that  $\rho$  is not a direct consistency maintaining rule up to layer  $k$  w.r.t.  $c$ . Then, there exists a transformation  $t : G \Longrightarrow_{\rho, m} H$  with  $\text{k}_{\max}(c, G) = k$  such that  $t$  is not direct consistency maintaining w.r.t.  $c$ . Therefore, either (4.1), (4.2), (4.3) or (4.4) is not satisfied.

1. If (4.1) is not satisfied, then  $k < \text{nl}(c) - 2$ , let  $j = \text{k}_{\max}(c, G) + 2$ , there does exist an occurrence  $p : C_j \hookrightarrow G$  such that  $p \models \text{IC}_0(\text{sub}_j(c), C)$  and  $\text{tr}_t \circ p \models \text{IC}_0(\text{sub}_j(c), C)$  with  $C \in \text{IG}(C_j, C_{j+1})$ . Since  $i_L^P(L \setminus K) \cap i_{C_{k+3}}^P(C_{k+3} \setminus C_{k+1}) = \emptyset$  for all  $P \in \text{ol}(L, C_{k+3})$  this case can to apply since  $\rho$  does not delete any elements of  $C_{j+1} \setminus C_j$ .
2. If (4.2) is not satisfied, there does exist an occurrence  $p : C_j \hookrightarrow H$  such that no morphism  $q : C_j \hookrightarrow G$  with  $p = \text{tr}_t \circ q$  exists and  $p \not\models \text{false}$  if  $\text{sub}_j(c) = \text{false}$  and  $p \not\models \text{IC}_0(\text{sub}_j(c), C_j + 1)$  otherwise. Since  $\rho$  satisfies 2  $G$  must have dangling edges and therefore,  $G$  is not a graph.
3. If (4.3) is not satisfied, there does exist an occurrence  $p : C_j \hookrightarrow H$  of an universally bound graph with  $j \leq \text{k}_{\max}(c, G)$  such that no morphism  $q : C_j \hookrightarrow G$  with  $\text{tr}_t \circ q = p$  exists. Again, since  $\rho$  satisfies 2  $G$  must have dangling edges and therefore,  $G$  is not a graph.
4. If (4.4) is not satisfied, there does exist an occurrence  $p : C_j \hookrightarrow H$  of an existentially bound graph with  $j \leq \text{k}_{\max}(c, G)$  such that  $\text{tr}_t \circ p$  is not total. Since  $\rho$  satisfied 1 the dangling edge condition would not bee satisfied and  $H$  would contain dangling edges.

In total follows that  $\rho$  is a direct basic consistency maintaining rule up to layer  $k$ .  $\square$

Now, we are ready to introduce *basic increasing rules at layer  $k$*  with  $k$  being odd. The set of basic increasing rules is a subset of the set of maintaining rules at layer  $k$  which ensures that the satisfaction at layer as well as the number of violation will not be decreased. Additionally, the left-hand side of these rule do contain an occurrence  $p$  of the universally bound graph  $C_{k+2}$ , such that this occurrence either will be removed, i.e. elements of  $C_{k+2} \setminus C_{k+1}$  will be deleted, or an intermediate graph  $C \in \text{IG}(C_{k+2}, C_{k+3})$  will be inserted. Of course, this second case only occurs if  $k < \text{nl}(c) - 3$  with  $c$  being the respective constraint. This property yields the advantage that application conditions for basic increasing rules are less complex and smaller, since it can be exactly determined how this rule removes a violation and therefore, no overlaps have to be considered.

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  $C_{k+2}$  is a sub-graph 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  $p$  will be removed and *inserting basic increasing rules* if  $\text{sub}_k(c)$  an intermediate graph will be inserted. For our repairing process, we will introduce the restriction that deleting basic increasing rules are only allowed to delete edges but no nodes of  $C_{k+2}$  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  $C_{k+2}$ . Then, it is unknown whether this node is connected to nodes not belonging to  $C_{k+2}$  and it is unclear whether all occurrence of  $C_{k+2}$  could be destroyed by  $\rho$  since the dangling edge condition might be unsatisfied.

**Definition 5.7 (basic increasing rule).** Let a constraint  $c$  in UANF and a direct consistency maintaining rule  $\rho = (\text{ac}, L \xleftarrow{l} K \xrightarrow{r} R)$  up to layer  $-1 \leq k < \text{nl}(v) - 1$ , with  $k$  odd, be given. Then,  $\rho$  is called basic increasing w.r.t  $c$  at layer  $k$  if a morphism  $p : C_{k+2} \hookrightarrow L$ , called the increasing morphism, exists such that either 1 or 2 applies.

1.  $r \circ l^{-1} \circ p$  is not total. Then,  $\rho$  is called a deleting basic increasing rule.
2. If  $k < \text{nl}(c) - 2$ , there does exist an intermediate graph  $C \in \text{IG}(C_{k+2}, C_{k+3})$  such that  $p \not\models \exists(a_{k+2}^r : C_{k+2} \hookrightarrow C, \text{true})$ ,  $r \circ l^{-1} \circ p$  is total and  $r \circ l^{-1} \circ p \models \exists(a_{k+2}^r : C_{k+2} \hookrightarrow C, \text{true})$ . Then,  $\rho$  is called a inserting basic increasing rule with  $C$ .

**Example 5.4.** Consider the rule `assignFeature` given in Figure 8 and constraint  $c_1$  given in Figure 2. Then, `assign Feature` is a inserting basic rule with  $C_2^2 \in \text{IG}(C_1^1, C_2^1)$  w.r.t.  $c_1$  but not a inserting basic rule with respect to the constraint  $\forall(C_2^2, \exists C_2^1)$  since the left-hand side of `assignFeature` does not contain an occurrence of  $C_2^2$ .

Again, 5.7 relies on every transformation of a rule  $\rho$ . Therefore, we present an alternative method to determine whether  $\rho$  satisfies 5.7 or not by checking that  $\rho$  does not delete any edges or nodes of  $C_{k+1} \setminus C_k$ .

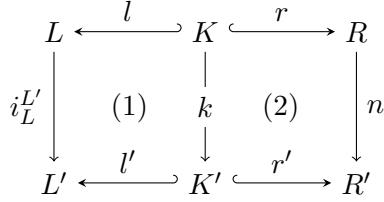


Figure 10: Pushout diagram for the construction of basic increasing rules.

As mentioned above, given a consistency maintaining rule  $\rho$  we can derive basic increasing rules that are only applicable if at a match and graph if  $\rho$  is applicable.

**Definition 5.8 (derived rules).** Let a constraint  $c$  in UANF and an rule  $\rho = (\text{ac}, L \xleftarrow{l} K \xrightarrow{r} R)$  be given. The set of derived rules of  $\rho$  at layer  $0 \leq k \leq \text{nl}(c) - 1$  contains rules characterized in the following way: Let

$$\mathbf{G} := \begin{cases} \text{IG}(C_k, C_{k+1}) & \text{if } \text{sub}_k(c) \text{ is existentially bound} \\ \{C_k\} & \text{otherwise} \end{cases}$$

For each  $P \in \mathbf{G}$  and  $L' \in \text{ol}(L, P)$ : If the diagram shown in Figure 10 is a transformation, i.e. (1) and (2) are pushouts, and for the characterisations of Definition 5.7 holds that

$$\begin{aligned}
\rho \text{ satisfies 1} &\implies L' \xleftarrow{l'} K' \xrightarrow{r'} R' \text{ satisfies 1} \wedge \\
\rho \text{ satisfies 2} &\implies L' \xleftarrow{l'} K' \xrightarrow{r'} R' \text{ satisfies 2}
\end{aligned}$$

the rule

$$\rho' = (\text{Shift}(\text{ac}, i_L^{L'}), L' \xleftarrow{l'} K' \xrightarrow{r'} R')$$

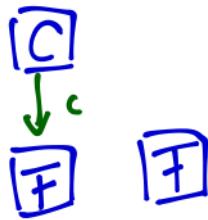
is a derived rule of  $\rho$  at layer  $k$ .

**Example 5.5.** Consider the rule `assignFeature` given in Figure 8 and constraint  $c_1$  given in Figure 2. The set of derived rules of  $\rho$  at layer 1 is given in Figure 11.

Obviously, a rule  $\rho'$  contained in the set of derived rules of a rule  $\rho$  is only applicable at a match  $m'$  if  $\rho$  is applicable at match  $m$  with  $m = m' \circ i$  with  $i$  being the inclusion of the left-hand side of  $\rho$  into the left-hand side of  $\rho'$ . Therefore, given a rules set  $\mathcal{R}$ , the extension of  $\mathcal{R}$  by the set of all derived rules for each rule of  $\mathcal{R}$  does not extend the expressiveness of  $\mathcal{R}$ . The main idea of the concept of derived rules is the extension of a given rule set by as many basic increasing rules as possible without extending the expressiveness of this set.

**Lemma 5.9.** Let a constraint  $c$  in UANF, and a rule  $\rho = (\text{ac}, L \xleftarrow{l} K \xrightarrow{r} R)$  be given, such that  $\rho$  is a maintaining rule up to layer  $-1 \leq k < \text{nl}(c) - 1$ , with  $k$  odd, and satisfies 1 and 2 of Definition 5.7. Then, each rule contained in the set of derived rules of  $\rho$  at layer  $k$  is a basic increasing rule.

assignFeature 2:



assignFeature 3:

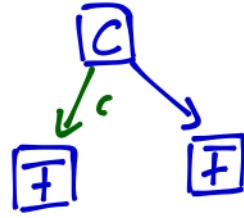


Figure 11: Derived rules of `assignFeature` and  $c_1$ .

*Proof.* Let  $\rho' = (\text{ac}', L' \xleftarrow{l'} K' \xrightarrow{r'} R')$  be one of those derived rules. Since  $\rho'$  deletes an inserts exactly the same elements as  $\rho$  and  $m' \models \text{ac}' \iff m' \circ i_L^{L'} \models$ ,  $\rho'$  is a consistency maintaining rule up to layer  $k$  and satisfies 1 and 2 of Definition 5.7. Since a morphism  $i : C_k \hookrightarrow L$  exists,  $\rho'$  is a basic increasing rule at layer  $k$ .  $\square$

In transformations via a rule  $\rho$ , such that the match intersects with an occurrence of an universally bound graph  $C_k$ ,  $\rho$  can be replaced by a derived rule of  $\rho$  at layer  $k$ .

**Lemma 5.10.** *Let a constraint  $c$  in UANF and a rule  $\rho = (\text{ac}, L \xleftarrow{l} K \xrightarrow{r} R)$  be given. Then, for each transformation*

$$t : G \Longrightarrow_{\rho, m} H$$

such that an occurrence  $p : C_k \hookrightarrow G$  of a universally bound graph  $C_k$  with  $p(C_k) \cap m(L) \neq \emptyset$  exists, there does exist a transformation

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

with  $\rho'$  being a derived rule of  $\rho$  at layer  $k$ .

*Proof.* Since  $p(C_k) \cap m(L) \neq \emptyset$  there does exist an overlap  $P \in \text{ol}(C_k, L)$  such that a morphism  $q : P \hookrightarrow G$  with  $m = q \circ i_L^P$  and  $p = q \circ i_{C_k}^P$  exists. Since  $t$  exists, there does exist a derived rule  $\rho' = (\text{ac}', L' \xleftarrow{l'} K' \xrightarrow{r'} R)$  with  $L' = P$ . We set  $m' = p$ , since  $m = m' \circ i_L L' \models \text{ac}$  it follows that  $m' = \text{ac}'$ . Since  $\rho$  does removes and inserts the same elements as  $\rho'$ , there does exist a transformation  $t' : G \Longrightarrow_{\rho', m'} H$ .  $\square$

Via this, we are able to replace consistency increasing transformations via a rule  $\rho$  that is consistency maintaining up to layer  $k$  and satisfies satisfies 1 and 2 of Definition 5.7 by a derived rule of  $\rho$  at layer  $k$ , that is, a basic increasing rule at layer  $k$ .

### 5.3 Application Conditions for Basic Rules

Let us now introduce the application conditions for basic increasing rules. Since these rules are maintaining rules up to a certain layer  $k$ , given a basic increasing rule  $\rho$ , it is sufficient to check whether  $m \circ i \not\models \exists C_{k+1}$  with  $i$  is a deleting rule and do check whether  $m \circ i \models \exists C$  if  $\rho$  is a inserting rule with  $C$  with  $i$  being the increasing morphism of  $\rho$ .

**Definition 5.11 (application conditions for basic increasing rules).** Let a constraint  $c$  in UANF and a basic increasing rule  $\rho = (\text{ac}, L \xleftarrow{l} K \xrightarrow{r} R)$  w.r.t  $c$  at layer  $-1 \leq k < \text{nl}(c) - 1$ , with  $k$  odd, be given. The basic application condition of  $\rho$  w.r.t.  $c$  at layer  $-1 \leq j < \text{nl}(c)$  is given by

$$\text{ac}' = \text{ac} \wedge \text{basic}_j(\rho)$$

with

$$\text{basic}_j(\rho) := \begin{cases} \bigwedge_{P \in \text{eol}(L, a_{k+2}^r, i)} \neg \exists(i_L^P : L \hookrightarrow P, \text{true}) & \text{if } j = k \text{ and } k < \text{nl}(c) - 2 \\ \text{true} & \text{if } k = \text{nl}(c) - 2 \\ \text{false} & \text{otherwise} \end{cases}$$

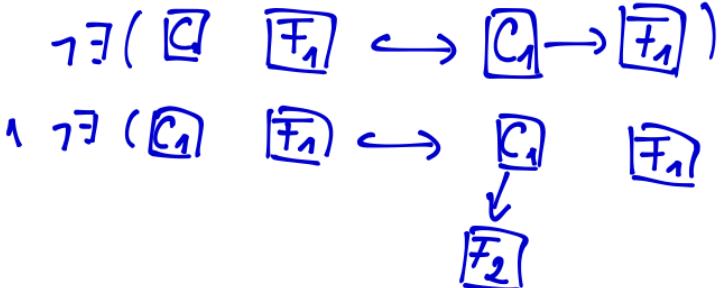
and  $a_{k+2}^r = a_{k+2}$ , if  $\rho$  is a deleting rule,  $a_{k+2}^r : C_{k+2} \hookrightarrow C$  if  $\rho$  is an inserting rule with  $C$  and  $i$  being the increasing morphism of  $\rho$ .

This application conditions are way easier to construct and smaller than the ones constructed by Definition 5.4. Note, that in case of a inserting basic rule  $\rho$  that inserts an intermediate graph  $C$ , the application condition only checks whether the increasing morphism does not satisfy  $\exists C$ . But, an application of this rule could also lead to a consistency increasing transformation w.r.t.  $c$  if the increasing morphism satisfied  $\exists C$ , if another intermediate graph  $C'$  will be inserted. To check this, conditions similar to the ones constructed via Definition 5.4 need to be constructed. On first sight, this seams like a restriction, but via the notion of derived rules we are able to dissolve this restriction, since the set of derived rules of  $\rho$  will contain a inserting basic increasing rule with  $C'$ , such that this rule, equipped with the according basic application condition can be used to perform this consistency increasing transformation. For example, consider the rule `assignFeature`, there does exist a consistency increasing transformation  $t : C_2^2 \Rightarrow_{\text{assignFeature}, m} C_2^1$  such that  $m \not\models \text{basic}_1(\text{assignFeature})$ , but there does also exist a transformation  $t : C_2^2 \Rightarrow_{\text{assignFeature3}, m'} C_2^1$  with  $m' \models \text{basic}_1(\text{assignFeature3})$ .

Let us now show that basic increasing rules equipped with the application condition constructed by Definition 5.11 are direct consistency increasing rules at layer.

**Theorem 5.3.** Let a constraint  $c$  in UANF and a basic increasing rule  $\rho = (\text{ac}, L \xleftarrow{l} K \xrightarrow{r} R)$  w.r.t  $c$  at layer  $k$  be given. Then,  $\rho' = (\text{ac} \wedge \text{basic}_k(\rho), L \xleftarrow{l} K \xrightarrow{r} R)$  is a direct consistency increasing rule at layer  $k$ .

*assign Feature:*



*assignFeature3:*

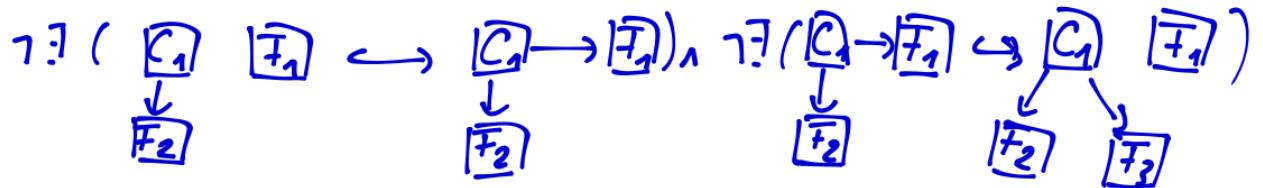


Figure 12: Application condition for `assignFeature` and `assignFeature3` with  $c_1$  at layer 1.

*Proof.* Let  $G$  be a graph with  $k_{\max} = k$ . We show that each transformation  $t : G \Rightarrow_{\rho',m} H$  is a direct consistency increasing transformation. Since,  $\rho$  is a direct basic increasing rule at layer  $k$ ,  $\rho'$  is also a consistency maintaining transformation at layer  $k$  and  $t$  satisfies (4.1), (4.2), (4.3) and (4.4). Therefore, we only need to show that  $t$  satisfies (4.5).

1. If  $\rho$  is a deleting rule,  $r \circ l^{-1} \circ i$  is not total, with  $i$  being the increasing morphism of  $\rho$ . If  $k = nl(c) - 2$ , this transformation is satisfied (4.5) since one occurrence of  $C_{k+2}$  has been destroyed. If  $k < nl(c) - 2$ , since  $m \models \text{basic}_k(\rho)$  the morphism  $m \circ i$  does not satisfy  $\exists C_{k+3}$ . Since this occurrence will be destroyed,  $t$  satisfies (4.5).
2. If  $\rho$  is a inserting rule with  $C \in \text{IG}(C_{k+2}, C_{k+3})$ , the morphism  $m \circ i$  does not satisfy  $\exists C$  since  $m \models \text{basic}_k(\rho)$ . But it holds that  $\text{tr}_t \circ m \circ i \models \exists C$  and therefore  $t$  satisfies (4.5).

In total follows that  $\rho'$  is a direct consistency increasing rule at layer  $k$ .  $\square$

**Example 5.6.** Again, consider the rule `assignFeature`, its derived rule `assignFeature3` and  $c_1$ . The application condition for these rules at layer  $k$  is given in Figure 12.

## 6 Rule-based Graph Repairs

### 6.1 Repairing rule Sets

**Definition 6.1 (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

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

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 6.2.** 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

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

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

**Theorem 6.1.** 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 6.1 is satisfied.

1. Assume that 1. of theorem 6.1 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 6.1 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  $\text{tr}_{t_0} \circ q \models \exists(a'_k : C_k \hookrightarrow P, \text{true})$ . 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 6.3.** *If a set of rule  $\mathcal{R}$  is a repairing set of  $c$  at layer  $k \leq \text{nl}(c)$  and 1. of theorem 6.1 applies, then  $\mathcal{R}$  is a repairing set of  $c$  at layer  $j$  for all  $k \leq j \leq \text{nl}(c)$ .*

**Lemma 6.4.** *Let a graph  $G_0$ , a constraint  $c$  in EANF and a repairing set  $\mathcal{R}$  at layer  $c_{\max}^{G_0} + 1$  be given, such that each rule in  $\mathcal{R}$  is a basic increasing. Then, for every sequence*

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

*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*

### 6.1.1 Construction of repairing sets

**Definition 6.5 (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*

$$\mathbf{D} \subseteq \{D_v \mid v \in V_{C_1 \setminus C_0}\}$$

*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 6.6.** *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 6.7.** *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 6.8.** 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 6.5. 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.

### Example 6.1.

## 6.2 Conflicts within and between Conditions

**Definition 6.9 (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  $t : G \Rightarrow_\rho H$  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  $C_k$  is has a conflict with an existentially bound graph  $C_j$  if a transformation  $t : G \Rightarrow_\rho H$  with  $\rho = C_k \xleftarrow{i_{C'}^k} C' \xrightarrow{\text{id}} C'$  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 6.10.** 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 6.11 (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 6.12 (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 (4.3) and (4.4) hold.

**Lemma 6.13.** Let a graph  $G$  and a constraint  $c$  in UANF be given. Let  $P$  be the set of all  $P$ . Then,  $G$  is conflict free w.r.t  $c$  if and only if either  $c$  is conflict free or

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

*Proof.*

□

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

**Theorem 6.2.** 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 6.14.** The set rule set constructed as described in construction 6.8 is a repairing set.

*Proof.*

□

**Theorem 6.3.** Algorithm 1 is correct and does terminate.

*Proof.*

□

---

**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  $G \implies \dots \implies H$ ;
5   end
6 end
```

---

---

**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  $G \implies \dots \implies H$ ;
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
```

---

---

**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
```

---

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

**Theorem 6.4.** *Algorithm 2 terminates and returns an consistent graph.*

*Proof.* □

## 6.5 Rule based repair for non-repairing sets

## 7 Conclusion

## 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.