

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

Alexander Lauer

February 18, 2023

**Abstract**

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Preliminaries</b>	<b>3</b>
2.1	Graphs and Graph morphisms . . . . .	4
2.2	Nested Graph Conditions and Constraints . . . . .	5
2.3	Rules and Graph Transformations . . . . .	7
2.4	Concepts of Consistency . . . . .	10
<b>3</b>	<b>Consistency Increase and Maintainment</b>	<b>12</b>
3.1	Universally quantified ANF . . . . .	12
3.2	Conditions up to Layer . . . . .	13
3.3	Consistency Increasing and Maintaining Transformations and Rules . . . . .	19
3.4	Direct Consistency Maintaining and Increasing Transformations . . . . .	21
3.5	Comparison with other concepts of Consistency . . . . .	27
<b>4</b>	<b>Application Conditions</b>	<b>30</b>
4.1	General Application Conditions . . . . .	32
4.2	Basic Consistency-increasing and Consistency-maintaining Rules . . . . .	40
4.3	Application Conditions for Basic Rules . . . . .	45
<b>5</b>	<b>Rule-based Graph Repair</b>	<b>47</b>
5.1	Conflicts within Conditions . . . . .	47
5.2	Repairing rule Sets . . . . .	52
5.3	Rule-based Graph Repair for one Constraint . . . . .	54
5.4	Rule-based Graph Repair for multiple Constraints . . . . .	59
<b>6</b>	<b>Related Work</b>	<b>61</b>
<b>7</b>	<b>Conclusion</b>	<b>63</b>

# 1 Introduction

Graphs and graph transformations are an appropriate framework for model-driven software engineering. The models (graphs) must be consistent with respect to a set of constraints. The concept of *nested graph conditions* is suitable to describe these constraints based on second order logic [?]. During development, inconsistencies may arise and a method for resolving them, called *model repair*, is necessary. Due to the versatility of graphs and graph transformations, the concept of model repair can be used to resolve inconsistencies for all kinds of graph-like structures. For example, in traffic light control or road networks.

For model repair, a notion is needed to assess the degree of inconsistency of a graph with respect to a constraint. There are already concepts for this, such as consistency preserving and guaranteeing transformations, and rules [?] that assess the change in consistency due to a graph transformation. These notions only look at the first graph of a constraint or whether the derived graph is consistent. They are not able to detect small increases or decreases in consistency caused by the insertion or deletion of single elements. In this paper we introduce the finer grained notions of *consistency-maintaining* and *consistency-increasing* transformations and rules that are able to detect these small increases or decreases in consistency. In addition, we introduce the more restricted notions of *direct consistency-maintaining* and *direct consistency-increasing* transformations and rules, which completely prohibit the insertion of new inconsistencies and have the advantage that second-order formulas can be used to specify these notions. We will also introduce direct consistency-maintaining and direct consistency-increasing application conditions for general rules, and direct consistency-increasing application conditions for a specific type or rules, called *basic rules*. It is known that the application conditions for these concepts are either very restrictive or large. For basic rules, we are able to construct smaller and less restrictive application conditions. We also present a rule-based graph repair approach capable of repairing specific constraints, namely *circular conflict free constraints*, and we extend this repair approach to repair specific sets of constraints, namely *circular conflict free sets of constraints*.

This paper is structured as follows: Formal prerequisites are introduced in section ???. The notions of (direct) consistency-maintaining and (direct) consistency-increasing transformations are given in section ??? and the construction of application conditions and characterisation of basic rules is given in section ???. The repair process is presented in section ??, and we summarise related graph repair approaches in section ??, before concluding the paper in section ??.

## 2 Preliminaries

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

## 2.1 Graphs and Graph morphisms

We start by introducing graphs and graph morphisms according to [?].

**Definition 2.1 (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. The edge  $e \in E$  connects the vertices  $\text{tar}(e)$  and  $\text{src}(e)$ .

If no tuple as above is given,  $V_G$ ,  $E_G$ ,  $\text{tar}_G$  and  $\text{src}_G$  denote the sets of vertices, edges and target and source mappings, respectively.

For the rest of this paper we will assume that all graphs are finite, i.e. given a graph  $G$ , the sets  $V_G$  and  $E_G$  are finite.

**Definition 2.2 (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. This means

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

holds. A graph morphism  $f$  is called injective (surjective) if  $f_E$  and  $f_V$  are injective (surjective) mappings. An injective morphism  $f : G \rightarrow H$  is called inclusion if  $f_E(e) = e$  and  $f_V(v) = v$  for all edges  $e \in E_G$  and all nodes  $v \in V_G$ . If  $f$  is injective, it is denoted with  $f : G \hookrightarrow H$ . 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 maintainment, we also need to consider *subgraphs*, *overlaps* of graphs, and so-called *intermediate graphs*. Intuitively, intermediate graphs are graphs  $G'$  which lie between two given graphs  $G$  and  $H$ . That is,  $G$  is a subgraph of  $G'$  and  $G'$  is a subgraph of  $H$ .

**Definition 2.3 (subgraph).** Let the graphs  $G$  and  $H$  be given. Then  $G$  is called a subgraph of  $H$  if an inclusion  $p : G \hookrightarrow H$  exists.  $G$  is called a proper subgraph of  $H$  if the  $p$  is not bijective.

Note that since the inclusion can also be surjective, by this definition every graph  $G$  is a subgraph of itself.

**Definition 2.4 (intermediategraph).** Let  $G$  and  $H$  be graphs such that  $G$  is a subgraph of  $H$ . A graph  $C$  is called an intermediategraph of  $G$  and  $H$ , if  $G$  is a proper subgraph of  $C$  and  $C$  is a subgraph of  $H$ . The set of intermediate-graphs of  $G$  and  $H$  is denoted by  $\text{IG}(G, H)$ .

**Definition 2.5 (overlap).** Let the graphs  $G_1$  and  $G_2$  be given. An overlap  $P = (H, i_{G_1}, i_{G_2})$  consists of a graph  $H$  and a jointly surjective pair of injective morphisms  $i_{G_1} : G_1 \hookrightarrow H$  and  $i_{G_2} : G_2 \hookrightarrow H$  with  $i_{G_1}(G_1) \cap i_{G_2}(G_2) \neq \emptyset$ . The set of all overlaps of  $G_1$  and  $G_2$  is denoted by  $\text{ol}(G_1, G_2)$ . If a tuple as above is not given, then  $G_P$ ,  $i_{G_1}^P$  and  $i_{G_2}^P$  denote the graph and morphisms of a given overlap  $P \in \text{ol}(G_1, G_2)$ .

Note that  $(H, i_{G_1}, i_{G_2})$  with  $i_{G_1}$  and  $i_{G_2}$  being jointly surjective and  $i_{G_1}(G_1) \cap i_{G_2}(G_2) = \emptyset$  could also be considered as an overlap of  $G_1$  and  $G_2$ . In this paper we only need to consider overlaps with  $i_{G_1}(G_1) \cap i_{G_2}(G_2) \neq \emptyset$ . So we have embedded this property directly into the definition.

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

**Definition 2.6 (restriction of a morphism).** *Let the graphs  $G, H$  and a morphism  $f : G \rightarrow H$  be given. Then, a morphism  $f' : G' \rightarrow H'$  is called a restriction of  $p$  if inclusions  $i : G' \hookrightarrow G$  and  $i' : H' \hookrightarrow H$  exist such that*

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

A restriction of  $p$  is denoted by  $p^r$ .

Note that given a morphism  $p : G \rightarrow H$  a restriction  $p^r : G' \rightarrow H'$  of  $p$  is uniquely determined by  $G'$  and  $H'$ . Assume two restrictions  $p^r : G' \rightarrow H'$  and  $q^r : G' \rightarrow H'$  of  $p$  are given. It holds that  $i' \circ p^r = p \circ i = i' \circ q^r$  and  $p^r = q^r$  follows because  $i'$  is injective.

## 2.2 Nested Graph Conditions and Constraints

*Nested graph constraints* are useful for specifying graph properties. The more general notion of *nested graph conditions* allows the specification of properties for graph morphisms and the definition of graph conditions and constraints in a recursive manner. Within these conditions, only quantifiers and Boolean operators are used [?].

**Definition 2.7 (nested graph condition).** *A nested graph condition over a graph  $C_0$  is defined recursively 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 inclusion 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 **false** =  $\neg$  **true**.*

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

Since these are the only types of conditions that will be used in this paper, we will refer to them only as *conditions* and *constraints*. We will use the more compact notations  $\exists(C_1, d)$  for  $\exists(a_0 : C_0 \hookrightarrow C_1, d)$  and  $\forall(C_1, d)$  for  $\forall(a_0 : C_0 \hookrightarrow C_1, d)$  if  $C_0$  and  $a_0$  are clear from the context.

**Example 2.1.** *Given a condition  $c = \forall(a_0 : \emptyset \hookrightarrow C_1, \exists(a_2 : C_1 \hookrightarrow C_2, \text{true}))$ . The compact notation of this condition is given by  $\forall(C_1, \exists(C_2, \text{true}))$ .*

**Definition 2.8 (semantic of graph conditions).** *Given a graph  $G$ , a condition  $c$  over  $C_0$  and a graph morphism  $p : C_0 \hookrightarrow G$ . Then  $p$  satisfies  $c$ , denoted by  $p \models c$ , if*

1.  $c = \text{true}$ .
2.  $c = \exists(a_0 : C_0 \hookrightarrow C_1, d)$  and there exists an injective morphism  $q : C_1 \hookrightarrow G$  with  $p = q \circ a_0$  and  $q \models d$ .
3.  $c = \neg d$  and  $p \not\models d$ .
4.  $c = d_1 \wedge d_2$  and  $p \models d_1$  and  $p \models d_2$ .
5.  $c = d_1 \vee d_2$  and  $p \models d_1$  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 constraint, constraints without any boolean operators. Each of these conditions can be transformed into an equivalent condition in so-called *alternating quantifier normal form* [?]. As the name suggests, these are conditions with alternating quantifiers and without any Boolean operators.

**Definition 2.9 (alternating quantifier normal form (ANF)).** *Conditions in alternating quantifier normal form (ANF) are defined recursively 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 an universally bound condition over  $C_1$  in ANF or  $d = \text{true}$ .
3.  $\forall(a_0 : C_0 \hookrightarrow C_1, d)$  is a condition in ANF if either  $d$  is an existentially bound condition over  $C_1$  in ANF or  $d = \text{false}$ .

*Every condition is a subcondition of itself. In cases ?? and ??,  $d$  is called a subcondition of  $\exists(a : C_0 \hookrightarrow C_1, d)$  or  $\forall(a : C_0 \hookrightarrow C_1, d)$  respectively. All subcondition of  $d$  are also subconditions 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  $\text{nl}(\text{true}) = \text{nl}(\text{false}) = 0$  and  $\text{nl}(\exists(a : P \hookrightarrow Q, d)) = \text{nl}(\forall(a : P \hookrightarrow Q, d)) := \text{nl}(d) + 1$ .*

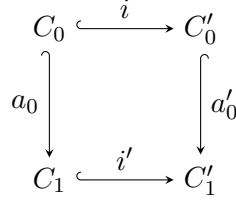


Figure 1: Diagram for the Shift operator.

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

In the following, we assume that all graphs and the nesting level of a condition are finite.  $c$

Using the *shift over morphism* construction, we are able to transform a nested condition over  $C$  into a nested condition over  $C'$  via an injective morphism  $i : C \hookrightarrow C'$  [?].

**Definition 2.10 (shift over morphism).** *Let a condition  $c$  over  $C_0$  and a morphism  $i : C_0 \hookrightarrow C'_0$  be given. The shift of  $c$  over  $i$ , denoted by  $\text{Shift}(c, i)$ , is given by*

1. If  $c = \text{true}$ ,  $\text{Shift}(c, i) = \text{true}$ .
2. If  $c = \exists(a_1 : C_0 \hookrightarrow C_1, d)$ ,  $\text{Shift}(c, i) = \bigvee_{(a', i') \in \mathcal{F}} \exists(a', \text{Shift}(d, i'))$  with  $\mathcal{F}$  being the set of all pairs  $(a', i')$  of injective morphisms that are jointly surjective and  $i' \circ a = a' \circ i$ , i.e., the diagram shown in Figure ?? commutes.
3. If  $c = \neg d$ ,  $\text{Shift}(c, i) = \neg \text{Shift}(d, i)$
4. If  $c = d_1 \wedge d_2$ ,  $\text{Shift}(c, i) = \text{Shift}(d_1, i) \wedge \text{Shift}(d_2, i)$
5. If  $c = d_1 \vee d_2$ ,  $\text{Shift}(c, i) = \text{Shift}(d_1, i) \vee \text{Shift}(d_2, i)$

**Lemma 2.11** ([?]). *Let a condition  $c$  over  $C_0$  and a morphism  $i : C_0 \hookrightarrow C'_0$  be given. Then, for each morphism  $m : C'_0 \hookrightarrow G$ ,*

$$m \models \text{Shift}(c, i) \iff m \circ i \models c$$

### 2.3 Rules and Graph Transformations

Via *rules* and *graph transformation* graphs can be modified by inserting or deleting nodes and edges. We will use the concept of the double-pushout approach for rules and transformations, which is based on category theory [?]. A rule consists of the

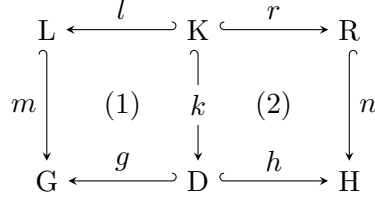


Figure 2: Diagram of a transformation in the double-pushout approach.

three graphs  $L$ , called the *left-hand side*,  $K$ , called *context*, and  $R$ , called *right-hand side*, where  $K$  is a subgraph of  $L$  and  $R$ . During a transformation, denoted by  $G \Rightarrow H$ , elements of  $L \setminus K$  are removed and elements of  $R \setminus K$  are inserted so that a new morphism  $p : R \hookrightarrow H$  is created. In addition, the so-called *dangling edge condition* must be satisfied, which means that for every edge  $e \in E_H$  there are vertices  $u, v \in V_H$  such that  $\text{tar}(e) = u$  and  $\text{src}(e) = v$  or vice versa. We also define application conditions. These are nested conditions over  $L$  and  $R$  that prevent the transformation if they are not satisfied. Later, we will use application conditions to ensure that transformations cannot reduce consistency. For example, application conditions that prevent a transformation if  $G \models c$  and  $H \not\models c$ .

**Definition 2.12 (rules and application conditions).** A plain rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  consists of graphs  $L, K, R$  and inclusions  $l : K \hookrightarrow L$  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  $(\text{ap}_L, \rho, \text{ap}_R)$  consists 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 2.13 (graph transformation).** Let a rule  $\rho = (\text{ap}_L, \rho', \text{ap}_R)$ , a graph  $G$  and a morphism  $m : L \hookrightarrow G$ , called the match, be given. Then, a graph transformation  $t : G \Rightarrow_{\rho, m} H$  is given in Figure ?? if the squares (1) and (2) are pushouts in the sense of category theory,  $m \models \text{ap}_L$  and the morphism  $n : L \hookrightarrow H$ , called the co-match of  $t$ , satisfies  $\text{ap}_R$ . The morphisms  $g : D \hookrightarrow G$  and  $h : D \hookrightarrow H$  are called the transformations morphisms of  $t$ .

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. The transformation must therefore be reversed if the co-match does not satisfy this condition. To avoid this, we introduce the *shift over rule* operation, which is capable of transforming a right into an equivalent left application condition [?].

**Definition 2.14 (shift over rule).** Let a plain rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  and a right application condition  $\text{ap}$  of  $\rho$  be given. The shift of  $\text{ap}$  over  $\rho$ , denoted with  $\text{Left}(\text{ap}, \rho)$ , is defined as

1. If  $\text{ap} = \text{true}$ ,  $\text{Left}(\text{ap}, \rho) := \text{true}$ .



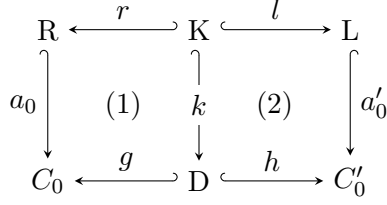


Figure 3: Transformation for the shift over rule operator.

2. If  $\text{ap} = \neg d$ ,  $\text{Left}(\text{ap}, \rho) := \neg \text{Left}(d, \rho)$ .
3. If  $\text{ap} = d_1 \wedge d_2$ ,  $\text{Left}(\text{ap}, \rho) := \text{Left}(d_1, \rho) \wedge \text{Left}(d_2, \rho)$ .
4. If  $\text{ap} = d_1 \vee d_2$ ,  $\text{Left}(\text{ap}, \rho) := \text{Left}(d_1, \rho) \vee \text{Left}(d_2, \rho)$ .
5. If  $\text{ap} = \exists(a_0 : R \hookrightarrow C_1, d)$ ,  $\text{Left}(\text{ap}, \rho) := \exists(a'_0 : L \hookrightarrow C'_0, \text{Left}(d, \rho'))$  where  $\rho' = C_0 \xleftarrow{g} D \xrightarrow{h} C'_0$  is the rule derived in Figure ?? by applying  $\rho^{-1}$  at match  $a_0$ . If this transformation does not exist, we set  $\text{Left}(\text{ap}, \rho) := \text{false}$ .

Shift over rule produces an equivalent left application condition, meaning that, given a right application condition  $\text{ap}$  and a plain rule  $\rho$ , a match of a transformation satisfies  $\text{Left}(\text{ap}, \rho)$  if and only if the co-match satisfies  $\text{ap}$  [?].

**Lemma 2.15.** *Let a plain rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$ , a right application condition  $\text{ap}$  for  $\rho$  and a transformation  $t : G \Longrightarrow_{\rho, m} H$  be given. Then,*

$$m \models \text{Left}(\text{ap}, \rho) \iff n \models \text{ap}.$$

Since every right application condition can be transformed into an equivalent left application condition, we will assume from now on that each rule contains only left application conditions. These rules are denoted by  $(\text{ap}, \rho)$ .

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

**Definition 2.16 (track morphism).** *Consider the transformation  $t$  shown in figure ?? . The track morphism,  $\text{tr}_t : G \dashrightarrow H$ , of  $t$  is a partial morphism, 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  $t : G \Longrightarrow H$ , the track morphism can be used to check whether a morphism  $p : C \hookrightarrow G$  extends to the derived graph  $H$  by checking whether  $\text{tr}_t \circ p$  is total, or whether 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 [?]. We will use these results later on.

**Lemma 2.17 ([?]).** *Let a transformation  $t : G \Longrightarrow G$  with transformation morphisms  $g : D \hookrightarrow G$ ,  $h : D \hookrightarrow H$  and an occurrence  $p : C \hookrightarrow G$  of a graph  $C$  be given. Then,*

1. The track morphism  $\text{tr}_t$  of  $t$  is total when restricted to  $p(C)$ , i.e.  $\text{tr}_t \circ p$  is total, if and only if there is an injective morphism  $p' : C \hookrightarrow D$  such that  $p = g \circ p'$ .
2. Given an injective morphism  $p : C \hookrightarrow H$ ,  $p(C)$  is contained in  $\text{tr}_t(G)$  if and only if there is an injective morphism  $p' : C \hookrightarrow D$  such that  $p = h \circ p'$ .

**Lemma 2.18** ([?]). *Given a transformation  $t : G \Rightarrow H$  with the transformation morphisms  $g : D \hookrightarrow G$ ,  $h : D \hookrightarrow H$  and a constraint  $c$  in ANF that contains the morphism  $a_i : C_{i-1} \hookrightarrow C_i$ . Given injective morphisms  $p_{i-1} : C_{i-1} \hookrightarrow G$  and  $p_i : C_i \hookrightarrow G$  such that  $p_{i-1} \circ a_i = p_i$  and the track morphism  $\text{tr}_t : G \dashrightarrow H$  is total when restricted to  $p_i(C_i)$ . Then,  $p'_{i-1} = p'_i \circ a_i$  where  $p'_{i-1} = \text{tr}_t \circ p_{i-1}$  and  $p'_i = \text{tr}_t \circ p_i$ . Also, given injective morphisms  $p'_{i-1} : C_{i-1} \hookrightarrow H$  and  $p'_i : C_i \hookrightarrow H$  such that  $p'_{i-1} = p'_i \circ a_i$  and  $p_i(C_i)$  is contained in  $\text{tr}_t(G)$ , then  $p_{i-1} = p_i \circ a_i$  where  $p_{i-1} = \text{tr}_t^{-1} \circ p'_{i-1}$  and  $p_i = \text{tr}_t^{-1} \circ p'_i$ .*

**Definition 2.19 (new and removed occurrences).** *Given a transformation  $t : G \Rightarrow H$  and a graph  $C$ . An occurrence  $p : C \hookrightarrow G$  is called a removed occurrence of  $C$  by  $t$  if  $\text{tr}_t \circ p$  is not total. An occurrence  $p' : C \hookrightarrow H$  is called a inserted occurrence of  $C$  by  $t$  if there is no morphism  $p : C \hookrightarrow G$  such that  $p' = \text{tr}_t \circ p$ .*

Given a sequence of transformations, the notion of *concurrent rules* can be used to describe this sequence by a rule. In other words, any sequence of transformations can be replaced by a transformation via its concurrent rule [?].

**Definition 2.20 (concurrent rule).** *Let the rules  $\rho_1 = L_1 \xleftarrow{l_1} K_1 \xrightarrow{r_1} R_1$ ,  $\rho_2 = L_2 \xleftarrow{l_2} K_2 \xrightarrow{r_2} R_2$  and a sequence of transformations*

$$G_1 \Rightarrow_{\rho_1, m_1} G_2 \Rightarrow_{\rho_2, m_2} G_3$$

*be given. Then,  $\rho' = G_1 \xleftarrow{l'} K \xrightarrow{r'} G_3$  is called the concurrent rule of the transformation sequence if the square (5) in Figure ?? is a pullback.*

A transformation sequence  $G_1 \Rightarrow_{\rho_1, m_1} G_2 \Rightarrow_{\rho_2, m_2} G_3$  can be replaced by a transformation  $G_1 \Rightarrow_{\rho', \text{id}} G_3$  via its concurrent rule. By inductive application, a concurrent rule for a transformation sequence  $G_1 \Rightarrow_{\rho_0} \dots \Rightarrow_{\rho_n} G_n$  of arbitrary finite length can be derived.

## 2.4 Concepts of Consistency

Now, we will introduce familiar consistency concepts. Namely, the notions of consistency preserving and guaranteeing transformations [?] and the notions of (direct) consistency sustaining and improving transformations [?]. Later, we will examine how these concepts differ from and correlate to our newly introduced concept of consistency.

**Definition 2.21 (consistency preserving and guaranteeing transformations).** *Let a constraint  $c$  and a transformation  $t : G \Rightarrow H$  be given. Then,  $t$  is called  $c$ -preserving if*

$$G \models c \Rightarrow H \models c.$$

*The transformation  $t$  is called  $c$ -guaranteeing if  $H \models c$ .*

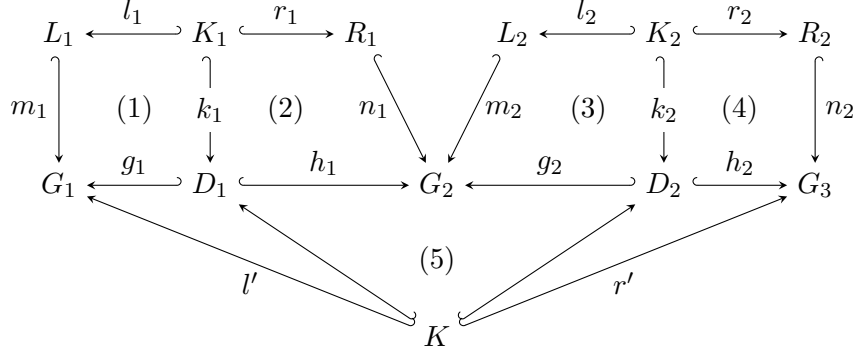


Figure 4: Pushout diagram of the transformation sequence  $G_1 \Rightarrow_{\rho_1, m_1} G_2 \Rightarrow_{\rho_2, m_2} G_3$  using the rules  $\rho_1 = L_1 \xleftarrow{l_1} K_1 \xrightarrow{r_1} R_1$  and  $\rho_2 = L_2 \xleftarrow{l_2} K_2 \xrightarrow{r_2} R_2$ .

While consistency preserving and guaranteeing transformations are defined for nested conditions, the finer-grained notions of (direct) consistency sustaining and improving transformations are defined only for conditions in ANF.

**Definition 2.22 (consistency sustaining and improving transformations).** *Let a constraint  $c$  in ANF and a transformation  $t : G \Rightarrow_{\rho} H$  be given. If  $c$  is existentially bound,  $t$  is called consistency sustaining w.r.t.  $c$  if it is  $c$ -preserving and  $t$  is called consistency improving w.r.t.  $c$  if it is  $c$ -guaranteeing. If  $c = \forall(a_0 : \emptyset \hookrightarrow C_1, d)$  is universally bound,  $t$  is called consistency sustaining w.r.t.  $c$  if*

$$|\{p : C_1 \hookrightarrow G \mid p \not\models d\}| \geq |\{p : C_1 \hookrightarrow H \mid p \not\models d\}|$$

*and  $t$  is called consistency improving w.r.t.  $c$  if*

$$|\{p : C_1 \hookrightarrow G \mid p \not\models d\}| > |\{p : C_1 \hookrightarrow H \mid p \not\models d\}|.$$

*The number of elements of these sets are called the number of violations in  $G$  and number of violations in  $H$  respectively.*

The even stricter notion of direct sustaining and improving transformations prohibits the insertion of new violations altogether.

**Definition 2.23 (direct sustaining and improving transformations).** *Let a constraint  $c$  in ANF and a transformation  $t : G \Rightarrow_{\rho} H$  be given. If  $c$  is existentially bound,  $t$  is called direct consistency sustaining w.r.t.  $c$  if  $t$  is  $c$ -preserving and  $t$  is called direct consistency improving w.r.t.  $c$  if  $t$  is  $c$ -guaranteeing.*

*If  $c = \forall(a_0 : \emptyset \hookrightarrow C_1, d)$ ,  $t$  is called consistency sustaining w.r.t.  $c$  if*

$$\begin{aligned} \forall p : C_0 \hookrightarrow G ((p \models d \wedge \text{tr}_t \circ p \text{ is total}) \implies \text{tr}_t \circ p \models d) \text{ and} \\ \forall p' : C_0 \hookrightarrow H (\neg \exists q : C_0 \hookrightarrow G (p' = \text{tr}_t \circ q) \implies p' \models d) \end{aligned}$$

*and  $t$  is called consistency improving w.r.t.  $c$  if additionally*

$$\begin{aligned} \exists p : C_0 \in G (p \not\models d \text{ and } \text{tr}_t \circ p \text{ is total} \wedge \text{tr}_t \circ p \models d) \vee \\ \exists p : C \hookrightarrow G (p \not\models d \wedge \text{tr}_t \circ p \text{ is not total}). \end{aligned}$$

### 3 Consistency Increase and Maintainment

In the following we will introduce the notions of *Satisfaction until after layer*, *(direct) consistency-maintaining* and *(direct) consistency-increasing* transformations and rules and compare them with the notions of consistency introduced in the previous section.

**Definition 3.1 (layer of a subcondition).** *Let a condition  $c$  in ANF and a subcondition  $d$  of  $c$  be given. 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 increases layer by layer, and that even small improvements, such as inserting 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, where consistency increasing indicates that the consistency has actually increased and consistency maintaining indicates that the consistency has not decreased.

#### 3.1 Universally quantified ANF

The definition of consistency increase and maintainment requires that each condition begins with a universal quantifier. Otherwise, case discrimination is required. 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). Furthermore, we will show that these sets are expressively equivalent by showing that every condition in ANF can be transformed into an equivalent condition in UANF.

**Definition 3.2 (universally quantified alternating quantifier normal form).** *A conditions  $c$  in ANF is in universally quantified ANF (UANF) if it is universally bound.*

Note that in our notation, given a condition  $c$  in UANF, any subcondition of  $c$  at layer  $0 \leq k \leq \text{nl}(c)$  is universal if  $k$  is an even number and existential if  $k$  is an odd number. Furthermore, a graph  $C_k$  of  $c$  is universally bound if  $k$  is an odd number and existentially bound if  $k$  is an even number. It is already known that an existentially bound condition  $c$  can be extended to the equivalent condition  $\exists(\text{id}_{C_0} : C_0 \hookrightarrow C_0, d)$  [?]. Analogously, we show that every condition in ANF has an equivalent condition in UANF.

**Lemma 3.3.** *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, then  $c$  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  $p \models c$ . Then  $p \models c'$ , because  $p$  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  $p \models c'$ , then all morphisms  $q : C_0 \hookrightarrow G$  with  $p = q \circ \text{id}_{C_0}$  satisfy  $c$ . Since  $p = p \circ \text{id}_{C_0}$ , it immediately follows that  $p \models c$ .  $\square$

$$c_1 = \forall(C_1^1, \exists(C_2^1, \text{true}))$$

$$c_2 = \forall(C_1^1, \exists(C_2^2, \forall(C_3^2, \exists(C_4^2, \text{true}))))$$

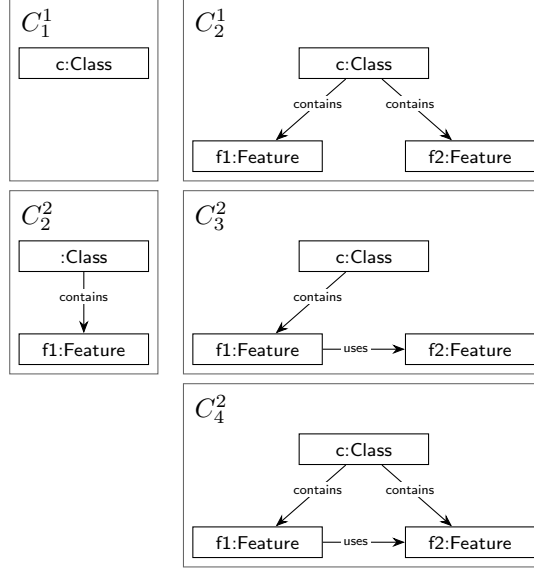


Figure 5: constraints

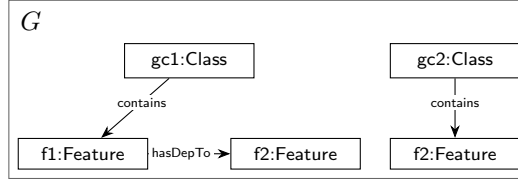


Figure 6: graph

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 every condition in ANF can be transformed into an equivalent condition in ANF that satisfies this property, by showing that  $\exists(\text{id}_{C_0} : C_0 \hookrightarrow C_0, \forall(a_1 : C_0 \hookrightarrow C_2, d))$  is equivalent to  $\forall(a_1 : C_0 \hookrightarrow C_2, d)$  and that  $\forall(\text{id}_{C_0} : C_0 \hookrightarrow C_0, \exists(a_1 : C_0 \hookrightarrow C_2, d))$  is equivalent to  $\exists(a_1 \circ a_0 : C_0 \hookrightarrow C_2, d)$  [?]. In addition, given a condition  $c$  in UANF, we will denote the first graph of  $c$  with  $C_0$ , the first morphism with  $a_0$ , the second graph with  $C_1$ , the second morphism with  $a_1$ , and so on. This means that we always write constraints as  $\forall(a_0 : C_0 \hookrightarrow C_1, \exists(a_1 : C_1 \hookrightarrow C_2, \dots))$ .

### 3.2 Conditions up to Layer

The goal of our approach is to increase the consistency of a constraint layer by layer, as we have already mentioned. To do this, we introduce a notion of partial consistency,

called *Satisfaction until after layer*, which allows us to check whether a constraint is satisfied at a particular layer by checking whether the so-called *truncated condition after layer* is satisfied at that layer.

We first define the *subcondition at layer*  $-1 \leq k \leq \text{nl}(c)$  of a condition  $c$ . As the name suggests, the subcondition at layer  $0 \leq k \leq \text{nl}(c)$  denotes the subcondition of  $c$  with layer  $k$ . We also define the subcondition at layer  $-1$ , which is *true*. This will be useful for defining the satisfaction at layer when a graph does not satisfy any layer of a constraint.

**Definition 3.4 (subcondition at layer).** *Let a condition  $c$  in ANF be given. The subcondition at layer  $-1 \leq k \leq \text{nl}(c)$ , denoted by  $\text{sub}_k(c)$ , is the subcondition  $d$  of  $c$  with  $\text{lay}(d) = k$  if  $0 \leq k \leq \text{nl}(c)$  and *true* if  $k = -1$ .*

Note that by definition the subcondition at layer  $k$  is always a condition over the graph  $C_k$  and the morphism is denoted by  $a_k$ .

**Example 3.1.** *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,  $\text{sub}_1(c) = \exists(a_1 : C_1 \hookrightarrow C_2, \forall(a_2 : C_2 \hookrightarrow C_3, \text{false}))$ .*

1.  $\text{sub}_{-1}(c) = \text{true}$ .
2.  $\text{sub}_0(c) = \forall(a_0 : C_0 \hookrightarrow C_1, \exists(a_1 : C_1 \hookrightarrow C_2, \forall(a_2 : C_2 \hookrightarrow C_3, \text{false}))) = c$ .
3.  $\text{sub}_1(c) = \exists(a_1 : C_1 \hookrightarrow C_2, \forall(a_2 : C_2 \hookrightarrow C_3, \text{false}))$ .
4.  $\text{sub}_2(c) = \forall(a_2 : C_2 \hookrightarrow C_3, \text{false})$ .
5.  $\text{sub}_3(c) = \text{false}$ .

Let us first introduce an operator which allows to replace a subcondition  $\text{sub}_k(c)$  by an arbitrary condition over  $C_k$ , called *replacement starting from layer*.

**Definition 3.5 (replacement starting from layer).** *Given 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. The replacement starting from layer  $k$  in  $c$  by  $e$ , denoted by  $\text{rep}_k(c, e)$ , is defined recursively 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 3.2.** *Consider the conditions  $c := \forall(a_0 : C_0 \hookrightarrow C_1, \exists(a_1 : C_1 \hookrightarrow C_2, \text{true}))$  and  $e = \exists(a'_1 : C_1 \hookrightarrow C_3, d)$ . The replacement of layer 1 in  $c$  by  $e$  is given by*

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

We now define *truncated conditions after layer* using the concept of replacement starting from layer. Intuitively, a condition is truncated after a particular layer by replacing the subcondition at the next layer with *true* or *false*, depending on which quantifier the replaced subcondition is bound by.

**Definition 3.6 (truncated condition after layer).** Let a condition  $c$  in UANF be given. The truncated condition of  $c$  after layer  $-1 \leq k < \text{nl}(c)$ , denoted by  $\text{cut}_k(c)$ , is defined as

$$\text{cut}_k(c) := \begin{cases} \text{true} & \text{if } k = -1 \\ \text{rep}_{k+1}(c, \text{true}) & \text{if } \text{sub}_k(c) \text{ is a existential condition, i.e. } k \text{ is odd} \\ \text{rep}_{k+1}(c, \text{false}) & \text{if } \text{sub}_k(c) \text{ is a universal condition, i.e. } k \text{ is even.} \end{cases}$$

**Example 3.3.** Consider constraint  $c_2$  given in Figure ?? . The truncated conditions of  $c_2$  after layer  $-1 \leq k < 3$  are given by

1.  $\text{cut}_{-1}(c_2) = \text{true}$ .
2.  $\text{cut}_0(c_2) = \forall(C_1^1, \text{false})$ .
3.  $\text{cut}_1(c_2) = \forall(C_1^1, \exists(C_2^2, \text{true}))$ .
4.  $\text{cut}_2(c_2) = \forall(C_1^1, \exists(C_2^2, (\forall C_3^2, \text{false})))$ .
5.  $\text{cut}_3(c_2) = \forall(C_1^1, \exists(C_2^2, (\forall C_3^2, \exists(C_4^2, \text{true})))) = c_2$ .

Note that the truncated condition of a condition  $c$  at layer  $\text{nl}(c) - 1$  is  $c$  itself. With these prerequisites we can now introduce *satisfaction until after layer*, which allows us to check whether a condition is satisfied up to a given layer. A morphism or graph satisfies a condition or constraint until after a layer if it satisfies the truncated condition after that layer.

**Definition 3.7 (satisfaction until after layer).** Let a graph  $G$  and a condition  $c$  in UANF be given. A morphism  $p : C_0 \hookrightarrow G$  satisfies  $c$  until after layer  $-1 \leq k < \text{nl}(c)$ , denoted by  $p \models_k c$ , if

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

A graph  $G$  satisfies a constraint  $c$  until after layer  $-1 \leq k < \text{nl}(c)$ , denoted by  $G \models_k c$ , if  $q : \emptyset \hookrightarrow G$  satisfies  $\text{cut}_k(c)$ . The largest  $-1 \leq k < \text{nl}(c)$  such that  $G \models_k c$  and there is no  $k < j < \text{nl}(c)$  with  $G \models_j c$ , called the largest satisfied layer, is denoted by  $k_{\max}(c, G)$ . When  $c$  and  $G$  are clear from the context, we use the abbreviation  $k_{\max}$ .

Note that given a graph  $G$  and a constraint  $c$ ,  $k_{\max}(c, G)$  always exists, since  $\text{cut}_{-1}(c) = \text{true}$  and every graph satisfies  $\text{true}$ . Moreover, if  $p \models_{\text{nl}(c)-1} c$ , i.e.  $k_{\max} = \text{nl}(c) - 1$ , it immediately follows that  $p \models c$ .

**Example 3.4.** Consider the graph  $G$  given in Figure ?? and the constraint  $c_2$  given in Figure ?? . This graph does not satisfy  $c_2$  because the occurrence of **Class** denoted with **gc2** does not satisfy  $\text{sub}_1(c_2) = \exists(C_2^2, (\forall C_3^2, (\exists C_4^2, \text{true})))$ , but it satisfies  $\text{cut}_1(c_2) = \forall(C_1^1, \exists(C_2^2, \text{true}))$  and therefore

$$G \models_1 c_2 \text{ and } k_{\max} = 1.$$

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

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

Given a graph  $G$ , a condition  $c$  and a morphism  $p : C_0 \hookrightarrow G$ . Suppose that  $p \models_k c$  with  $0 \leq k < \text{nl}(c)$ . Then we can infer results for the satisfaction until after other layers. If  $k$  is even, i.e.  $\text{sub}_k(c)$  is a universal condition, we can conclude that  $p \models_j c$  for all  $k < j < \text{nl}(c)$  and especially  $p \models c$ . It also follows that  $p \models_j c$  for all odd  $0 \leq j < k$ , i.e.  $\text{sub}_j(c)$  is an existential condition. We present these results in the following lemmas, and an overview is given in Table ??.

We start by examining the consequences for the satisfaction until after layer  $\text{nl}(c) > j > k$  if  $p \models_k c$ . Our first lemma shows that replacing the subcondition  $\text{sub}_{k+1}(c)$  by any condition over  $C_{k+1}$  leads to a condition that is satisfied by  $p$  if  $k$  is even.

**Lemma 3.8.** *Given a graph  $G$ , a condition  $c$  in UANF and a morphism  $p : C_0 \hookrightarrow G$  with  $p \models_k c$  and  $-1 \leq k < \text{nl}(c)$  even. Then, for any condition  $f$  over  $C_{k+1}$  it holds that*

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

*Proof.* We start by showing the statement for the smallest  $-1 \leq j < \text{nl}(c)$  such that  $\text{sub}_j(c)$  is universally bound and  $p \models_j c$ . After this, we can conclude that this statement holds for all  $-1 \leq i < \text{nl}(c)$  such that  $\text{sub}_i(c)$  is universally bound and  $p \models_i c$ .

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 morphism must exist, since  $j$  is the smallest even number with  $p \models_j c$ . 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  $p \models \text{rep}_{k+1}(c, d)$  because  $\text{rep}_{k+1}(c, d) = \text{rep}_{j+1}(c, \text{sub}_{j+1}(\text{rep}_{k+1}(c, d)))$ .  $\square$

As a direct consequence of the previous lemma, a morphism which satisfies the condition at layer  $k$ , where  $k$  is even, also satisfies the condition at layer  $j$  for all  $j > k$ .

**Lemma 3.9.** *Given a graph  $G$ , a morphism  $p : C_0 \hookrightarrow G$  and a condition  $c$  in UANF. If  $0 \leq k < \text{nl}(c)$  is even, i.e.  $\text{sub}_k(c)$  is a universal condition, 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 ?? and setting  $f$  equal to  $\text{sub}_{k+1}(\text{cut}_j(c))$ .  $\square$



$p \models_k c$	$k_{\max} < \text{nl}(c) - 1$ $k \leq k_{\max}$	$k_{\max} = \text{nl}(c) - 1$ $k < k_{\max}$
$k$ even	$\times$	?
$k$ odd	$\checkmark$	$\checkmark$

Table 2: Overview of the satisfaction until after layer  $k$  with respect to  $k_{\max}$ , where “ $\checkmark$ ” indicates that  $p \models_k c$ , “ $\times$ ” indicates that  $p \not\models_k c$  and “?” indicates that it cannot be concluded from  $p \models_k c$  whether  $p \models_j c$  or  $p \not\models_j c$ .

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

**Corollary 3.10.** *Given a graph  $G$ , a morphism  $p : C_0 \hookrightarrow G$  and a condition  $c$  in UANF. 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. Given a graph  $G$ , a morphism  $p : C_0 \hookrightarrow G$  and a condition  $c$  such that  $p \models_k c$  for an even  $-1 \leq k < \text{nl}(c)$ . It follows that  $p \models c$  and in particular  $p \models c'$  for each condition  $c'$  with  $\text{cut}_k(c) = \text{cut}_k(c')$ .

Let us now examine the satisfaction until after layer  $j$  with  $-1 < j < k$ . If  $j$  is odd, i.e.  $\text{sub}_j(c)$  is an existential condition, we can conclude that  $p \models_j c$  as shown in the next lemma. If  $j$  is even, i.e.  $\text{sub}_j(c)$  is universally bound, we can only make statements that depend on  $k_{\max}$ . If  $k_{\max} < \text{nl}(c) - 1$ , then  $p \not\models_j c$ . Otherwise  $p \models c$  and therefore  $k_{\max} = \text{nl}(c) - 1$  would immediately follow Corollary from ???. If  $k_{\max} = \text{nl}(c) - 1$ , we can say that there is at least one even  $j \leq k_{\max}$  with  $p \models_j c$  if  $c$  ends with  $\forall(C_{\text{nl}(c)}, \text{false})$ . An overview of these relations is given in Table ???.

**Lemma 3.11.** *Given a graph  $G$ , a morphism  $p : C_0 \hookrightarrow G$  and a constraint  $c$  in UANF. Then for all odd  $-1 \leq k \leq k_{\max}$ , i.e.  $\text{sub}_k(c)$  is an existential condition, we have*

$$p \models_k c.$$

*Proof.* If there is an even  $0 \leq j < k_{\max}$ , i.e.  $\text{sub}_j(c)$  is universal, with  $p \models_j c$ , let  $j'$  be the smallest of these. Lemma ??? implies that  $p \models_\ell c$  for all  $j' \leq \ell < \text{nl}(c)$ . Otherwise we set  $j' = k_{\max}$ .

Let  $\ell < j'$ , such that  $\text{sub}_\ell(c)$  is an existential condition and let  $d = \text{sub}_\ell(\text{cut}_{j'}(c)) = \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, e)$  be the subcondition at layer  $\ell$  of the truncated condition after layer  $j'$  of  $c$ . Since  $\ell < j'$ , there must be a morphism  $q : C_\ell \hookrightarrow G$  with  $q \models d$  and therefore there must be a morphism  $q' : C_{\ell+1} \hookrightarrow G$  with  $q = q' \circ a_\ell$  and  $q' \models e$ . It follows that  $q \models \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, \text{true})$  and thus  $p \models_\ell c$ .  $\square$

**Example 3.5.** *We will show counterexamples for all “?” in Table ??? and Table ???. Consider constraint  $c_2 = \forall(C_1^1, \exists(C_2^2, \forall(C_3^2, \exists(C_4^2, \text{true}))))$  given in Figure ???. We begin with Table ???.*

1.  $j \leq k$  and  $j$  is even. We set  $k = 2$  and  $j = 0$ . It follows that  $\text{cut}_2(c) = \forall(C_1^1, \exists(C_2^2, \forall(C_3^2, \text{false})))$  and  $\text{cut}_0(c) = \forall(C_1^1, \text{false})$ . Then,  $C_3^2 \models_2 c_2$ ,  $C_3^2 \not\models_0 c_2$  and  $\emptyset \models_2 c_2$  and  $\emptyset \models_0 c_2$ .
2. If  $k = 3$ , then  $C_4^2 \models_3 c_2$ ,  $C_2^2 \not\models_0 c_2$  and  $\emptyset \models_3 c_2$  and  $\text{emptyset} \models_0 c_2$ .
3. If  $k = 1$ , then  $C_2^2 \models_2 c_2$ ,  $C_2^2 \models_3 c_2$  and  $C_3^2 \not\models_2 c_2$  and  $C_3^2 \not\models_3 c_2$ .

For the “?” in Table ??, consider the graph  $C_2^2$ . It follows that  $k_{\max} = 3 = \text{nl}(c_2) - 1$ ,  $C_2^2 \models_2 c_2$  and  $C_2^2 \not\models_0 c_2$ .

By satisfaction at layer, an increase of consistency can be detected in the following way: Let  $t : G \implies H$  be a transformation. If the largest satisfied layer in  $H$  is greater than the largest satisfied layer in  $G$ , i.e.  $k_{\max}(c, G) < k_{\max}(c, H)$ , we consider the transformation as consistency increasing. However, the notion of consistency increasing should also be able to detect the smallest changes made by a transformation that lead to an increase of consistency, namely the insertion of a single edges or nodes of an existentially bound graph. To remedy this problem, we introduce *intermediate conditions*, which are used to detect this type of increase by checking whether an intermediate condition not satisfied by  $G$  is satisfied by  $H$ . Obviously, a decrease of consistency can be detected in a similar way, by checking whether an intermediate condition satisfied by  $G$  is not satisfied by  $H$ . Intuitively, the last graph of a truncated condition  $c$  will be replaced by an intermediate graph of the penultimate graph and the last graph.

If  $c$  ends with an existentially bound condition, the constructed intermediate condition is weaker than  $c$ , in the sense that the satisfaction of  $c$  implies the satisfaction of the intermediate condition. Conversely, if  $c$  ends with a universally bound condition, the opposite holds: satisfying an intermediate condition implies satisfying  $c$ . This is why we have designed intermediate conditions so that they only replace graphs on existentially bounded layers.

**Definition 3.12 (intermediate condition).** Let a condition  $c$  in UANF be given and let  $0 \leq k < \text{nl}(c)$  be 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 3.6.** Consider the constraint  $c_1$  given in Figure ??. Since  $C_2^2 \in \text{IG}(C_1^1, C_2^1)$ , we can construct an intermediate condition of  $c_1$  at layer 1 with  $C_2^2$  as  $\text{IC}_1(c_1, C_2^2) = \forall C_1^1 \exists C_2^2$ . While  $c_1$  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.

Notice that  $\text{IC}_k(c, C_k)$  is equivalent to  $\text{cut}_{k-1}(c)$ , because the condition  $\forall(a_{k-1} : C_{k-1} \hookrightarrow C_k, \exists(a_k^r : C_k \hookrightarrow C_k, \text{true}))$  is satisfied by every morphism  $p : C_{k-1} \hookrightarrow G$ .

### 3.3 Consistency Increasing and Maintaining Transformations and Rules

With the results above, we are now ready to define the notions of *consistency increase-ment* and *maintainment*, where increasement is a special case of maintainment. A transformation  $t$  is considered as consistency maintaining if it does not decrease consistency, while  $t$  is considered as consistency increasing if it increases the consistency.

These notions are designed to detect only 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  $c$ , a transformation  $t : G \Rightarrow H$  is considered as consistency maintaining if the largest satisfied layer has not decreased, i.e. if  $k_{\max}(c, G) \leq k_{\max}(c, H)$ , and at least as many increasing insertions or deletions have been made as decreasing ones. An increasing deletion is the deletion of an occurrence of  $C_{k_{\max}(c, G)+2}$  that does not satisfy  $\exists(C_{k_{\max}(c, G)+3}, \text{true})$ , an increasing insertion is the insertion of elements, such that for at least one occurrence  $p$  of  $C_{k_{\max}(c, G)+2}$  it holds that  $p \not\models \exists(C', \text{true})$  and  $\text{tr}_t \circ p \models \exists(C', \text{true})$  for an intermediate graph  $C' \in \text{IG}(C_{k_{\max}(c, G)+2}, C_{k_{\max}(c, G)+3})$ . Decreasing insertions and deletions are the opposite of increasing ones. That is, a decreasing insertion is the insertion of an occurrence of  $C_{k_{\max}(c, G)+2}$  that does not satisfy  $\exists(C_{k_{\max}(c, G)+3}, \text{true})$  and a decreasing deletion is the deletion of elements such that for one occurrence  $p$  of  $C_{k_{\max}(c, G)+2}$  with  $p \models \exists(C', \text{true})$  it holds that  $\text{tr}_t \circ p \not\models \exists(C', \text{true})$  for an intermediate graph  $C' \in \text{IG}(C_{k_{\max}(c, G)+2}, C_{k_{\max}(c, G)+3})$ . 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  $C_{k_{\max}+2}$  the number of graphs  $C' \in \text{IG}(C_{k_{\max}+2}, C_{k_{\max}+3})$  with  $p \not\models \exists C'$  is added up, and by comparing these numbers for  $G$  and  $H$  it can be determined whether there have been more increasing insertions and deletions than decreasing ones.

The number of violations is defined for each layer of the constraint, but only for the first unsatisfied layer is the sum 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  $\infty$ . In this way, a transformation  $t : G \Rightarrow H$  that increases the largest satisfied layer can be easily detected, since the number of violations in  $H$  at the layer  $k_{\max} + 1$  will be set to 0.

**Definition 3.13 (number of violations).** *Given a graph  $G$  and a constraint  $c$  in UANF, and let  $e = \text{sub}_{k_{\max}+2}(c)$ . The number of violations  $\text{nv}_j(c, G)$  at layer  $-1 \leq j < \text{nl}(c)$  in  $G$  is defined as:*

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

Note that the second and third cases of Definition ?? only occur if  $G \not\models c$  and  $\text{sub}_{k_{\max}}(c)$  is existentially bound. So  $e$  is also existentially bound or equal to false if  $c$

ends with  $\forall(C_{\text{nl}(c)}, \text{false})$  and  $k_{\max} = \text{nl}(c) - 2$ . Using the number of violations, we now define *consistency maintaining* and *increasing* transformations and rules by checking that the number of violations has not increased or, in the case of consistency increasing, has decreased.

**Definition 3.14 (consistency maintaining and increasing transformations and rules).** *Given a graph  $G$ , a constraint  $c$  in UANF and a rule  $\rho$ . A transformation  $t : G \Rightarrow_{\rho, m} H$  is called consistency-maintaining w.r.t.  $c$ , if*

$$\text{nv}_k(c, H) \leq \text{nv}_k(c, G)$$

*for all  $-1 \leq k < \text{nl}(c)$ . The transformation is called consistency-increasing w.r.t.  $c$  if*

$$\text{nv}_k(c, H) < \text{nv}_k(c, G)$$

*for all  $-1 \leq k < \text{nl}(c)$ . A rule  $\rho$  is called consistency maintaining or increasing w.r.t  $c$ , if all of its transformations are.*

Note that if  $G \models c$ , there is no consistency-increasing transformation  $G \Rightarrow H$  w.r.t  $c$ , since  $\text{nv}_j(c, G) = 0$  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  $t : G \Rightarrow_{\rho, m} H$  can always be constructed. Therefore, every consistency-increasing rule must have at least one application condition.

As mentioned above, a transformation is considered to be consistency-increasing if the largest satisfied layer is increasing. This property is already indirectly embedded in the definition of consistency-increasing transformations.

**Theorem 3.1.** *Given a rule  $\rho$  a constraint  $c$  in UANF and a graph  $G$  with  $G \not\models c$ . 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,  $\text{nv}_{k_{\max}(c, G)+1}(c, G) > 0$  and  $\text{nv}_{k_{\max}(c, G)+1}(c, G) \neq \infty$ . Since  $k_{\max}(c, H) > k_{\max}(c, G)$ ,  $\text{nv}_{k_{\max}(c, G)+1}(c, H) = 0$ , which immediately implies that  $t$  is consistency increasing w.r.t.  $c$ .  $\square$

Since there are no consistency-increasing transformations starting from consistent graphs, there are no infinitely long sequences of consistency-increasing transformations.

**Theorem 3.2.** *Let  $c$  be a constraint in UANF. Every sequence of consistency-increasing transformations w.r.t.  $c$  is finite.*

*Proof.* Let

$$G_0 \Rightarrow_{\rho_0, m_0} G_1 \Rightarrow_{\rho_1, m_1} G_2 \Rightarrow_{\rho_2, m_2} \dots$$

be a sequence of consistency-increasing transformations w.r.t.  $c$ . We assume that  $k_{\max}(c, G_0) < \text{nl}(c) - 1$ , otherwise  $\text{nv}_j(c, G_0) = 0$  for all  $0 \leq j < \text{nl}(c)$  and there is no consistency-increasing transformation  $G_0 \Rightarrow H$  with respect to  $c$ .

We show that  $G_j \models_{k_{\max}(c, G_0)+2} c$  holds after a maximum of  $j := \text{nv}_{k_{\max}(c, G_0)+1}(c, G_0)$  transformations. Note that  $j$  must be finite, since  $G_0$  contains only a finite number of occurrences of  $C_{j+1}$ . Since every transformation is consistency-increasing w.r.t.  $C$ , it follows that  $\text{nv}_{k_{\max}(c, G_i)+1}(c, G_{i+1}) \leq \text{nv}_{k_{\max}(c, G_i)+1}(c, G_i) - 1$  after each transformation. Therefore, after at most  $j$  transformations,  $\text{nv}_{k_{\max}(c, G_0)+1}(c, G_j) \leq \text{nv}_{k_{\max}(c, G_0)+1}(c, G_0) - j = 0$  and thus  $G_j \models_{k_{\max}(c, G_0)+2} c$ . If this is applied iteratively, it follows that after a finite number of transformations there must exist a graph  $G_k$  with  $G_k \models c$ . Since there is no consistency increasing transformation  $G_k \Rightarrow_{\rho_k, m_k} G_{k+1}$ , the sequence must be finite.  $\square$

### 3.4 Direct Consistency Maintaining and Increasing Transformations

We will now introduce stricter versions of consistency-increasing and consistency-maintaining transformations, called *direct consistency-maintaining* and *direct consistency-increasing* transformations. These can also be considered as consistency-maintaining and consistency-increasing transformations, which do not perform any unnecessary insertions and deletions. For example, given a constraint  $c$  in UANF and graphs  $G$  with  $G \not\models c$  and  $H$  with  $H \models c$ , 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-increasing and consistency-maintaining - a similar example of a consistency-maintaining transformation can be easily constructed - allow insertions or deletions that are unnecessary to increase or maintain consistency. That is, deleting occurrences of existentially bound graphs, deleting occurrences  $p : C_k \hookrightarrow G$  of universally bound graphs  $C_k$  satisfying  $\exists(C_{k+1}, \text{true})$  or inserting occurrences of universally bound graphs and inserting 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 satisfies  $\exists(C', \text{true})$ .

*Direct consistency-increasing* and *maintaining* transformations are more restricted, in the sense that these unnecessary deletions and insertions cause a transformation not to be direct consistency-increasing or direct consistency maintaining, respectively. In addition, we can use second-order logic formulas to check whether a transformation is direct consistency-maintaining. Furthermore, it is ensured that no new violations are introduced, since these can always be considered as unnecessary insertions or deletions. Thus, the removal of one violation is sufficient to state that the transformation is (direct) consistency-increasing, which can also be checked using a second-order logic formula. We start by introducing *direct consistency-maintaining*. Its definition consists of the following conditions

1. *Deleting condition*: This condition ensures that no new violations are introduced by deleting intermediate graphs  $C' \in \text{IG}(C_{k_{\max}+2}, C_{k_{\max}+3})$ . This leads to the insertion of new violations only if an occurrence of  $C_{k_{\max}+2}$  which satisfies  $\exists(C', \text{true})$

in the originating graph does not satisfy  $\exists(C', \text{true})$  in the derived graph of the transformation. Therefore, this condition checks that this case does not occur.

2. *Inserting condition:* This condition ensures that no new violations are introduced by inserting an occurrence of  $C_{k_{\max}+2}$ . Again, this only causes a new violation to be inserted if that occurrence does not satisfy  $\exists(C_{k_{\max}+3}, \text{true})$ . The condition checks that this is not the case.
3. *Universally condition:* This condition ensures that the largest satisfied layer is not reduced by inserting a universally bound graph  $C_j$ . This can only happen if  $j \leq k_{\max}$ , and the condition checks that no occurrences of such universally bound graphs are inserted.
4. *Existentially condition:* This condition ensures that the largest satisfied layer is not reduced by deleting an existentially bound graph  $C_j$ . Again, this can only happen if  $j \leq k_{\max}$ . The condition checks that no occurrences of such existentially bound graphs are deleted.

The deleting and inserting conditions ensure that the number of violations is not increased, and the universally and existentially conditions ensure that the largest satisfied layer is not reduced. Of course, the violation of the universally and existentially conditions does not necessarily lead to a decrease of the largest satisfied layer, but it can also be considered as an unnecessary insertion or deletion.

Since a condition  $c$  in UANF is also allowed to end with  $\forall(C_{\text{nl}(c)}, \text{false})$  the deleting and inserting conditions contain case discrimination. If the constraint  $c$  ends with  $\forall(C_{\text{nl}(c)}, \text{false})$  and  $k_{\max} = \text{nl}(c) - 2$ , there is no graph  $C_{k_{\max}+3}$  and thus no intermediate graphs. Therefore, the deleting condition is always satisfied and we will set it to **true**. The inserting condition will check that no new occurrences of  $C_{k_{\max}+3}$  are introduced at all.

**Definition 3.15 (direct consistency maintaining transformations).** *Given a graph  $G$ , a rule  $\rho$  and a constraint  $c$  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  $e = \text{sub}_k(c)$  and  $k_{\max} = k_{\max}(c, G)$ . A transformation  $t : G \Rightarrow_{\rho, m} H$  is called direct consistency maintaining w.r.t.  $c$  if the following conditions are satisfied.*

1. *Deleting condition:* If  $e \neq \text{false}$ , then each occurrence of  $C_{k_{\max}+2}$  in  $G$  which satisfies  $\text{IC}_0(e, C')$  for any  $C' \in \text{IG}(C_{k_{\max}+2}, C_{k_{\max}+3})$  still satisfies  $\text{IC}_0(e, C')$  in  $H$ :

$$\begin{aligned} \forall p : C_k \hookrightarrow G \left( \bigwedge_{C' \in \text{IG}(C_{k_{\max}+2}, C_{k_{\max}+3})} (p \models \text{IC}_0(e, C') \wedge \text{tr}_t \circ p \text{ is total}) \right. \\ \left. \implies \text{tr}_t \circ p \models \text{IC}_0(e, C') \right) \end{aligned} \quad (3.1)$$

*Otherwise, if  $e = \text{false}$ , the deleting condition is satisfied.*

2. Inserting condition: Let  $d = \text{IC}_0(e, C_{k_{\max}+3})$  if  $e \neq \text{false}$  and  $d = \text{false}$  otherwise. Each newly inserted occurrence of  $C_{k_{\max}+2}$  satisfies  $d$ .

$$\forall p' : C_{k_{\max}+2} \hookrightarrow H(\neg \exists p : C_{k_{\max}+2} \hookrightarrow G(p' = \text{tr}_t \circ p) \implies p' \models d) \quad (3.2)$$

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

$$\bigwedge_{\substack{i < k_{\max} \\ C_i \text{ universally}}} \forall p : C_i \hookrightarrow H(\exists p' : C_i \hookrightarrow G(p = \text{tr}_t \circ p')) \quad (3.3)$$

4. Existentially condition: No occurrence of an existentially bound graph  $C_j$  with  $j \leq k_{\max}+1$  is deleted.

$$\bigwedge_{\substack{i \leq k_{\max} \\ C_i \text{ existentially}}} \forall p : C_i \hookrightarrow G(\text{tr}_t \circ p \text{ is total}) \quad (3.4)$$

Before continuing with the definition of direct consistency increasing, let us first show that every direct consistency maintaining transformation is indeed consistency maintaining. To do this, we first show that satisfying the universally and existentially conditions guarantees that the satisfaction at layer has not been decreased.

**Lemma 3.16.** *Given a transformation  $t : G \implies H$  and a constraint  $c$  in UANF such that the universally and existentially condition is satisfied. Then*

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

*Proof.* Let us assume that  $H \not\models_{k_{\max}(c, G)} c$ . Then either a new occurrence of a universally bound graph  $C_i$  with  $i < k_{\max}(c, G)$  has been inserted, or an occurrence of an existentially bound graph  $C_j$  with  $j \leq k_{\max}(c, G)$  has been destroyed. Therefore, the following applies:

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

where  $i, j \leq k_{\max}(c, G)$ ,  $i$  is even and  $j$  is odd, i.e.  $C_i$  is universally and  $C_j$  is existentially bound. It follows immediately that either the universally or the existentially condition is not satisfied. This is a contradiction.  $\square$

With this we are now going to show that a direct consistency-maintaining transformation is also a consistency-maintaining transformation.

**Theorem 3.3.** *Given 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.  $c$ . Then,  $t$  is also a consistency-maintaining transformation.*

*Proof.* With Lemma ?? follows that  $k_{\max}(c, G) \leq k_{\max}(c, H)$  and thus  $nv_{k_{\max}(c, G)+1}(c, H) \neq \infty$ . It remains to show that  $nv_k(c, H) \leq nv_k(c, G)$  for all  $0 \leq k < nl(c)$ . In particular, we only need to show that  $nv_{k_{\max}(c, G)+1}(c, H) \leq nv_{k_{\max}(c, G)+1}(c, G)$  since for all  $-1 \leq j < k_{\max}(c, G) + 1$  it holds that  $nv_j(c, H) = nv_j(c, G) = 0$ . Also, since  $nv_j(c, G) = \infty$  for all  $k_{\max}(c, G) + 1 < j < nl(c)$ , it follows that  $nv_j(c, H) \leq nv_j(c, G)$ .

Let  $k_{\max} = k_{\max}(c, G)$  and  $d = \text{sub}_{k_{\max}+2}(c)$ . We show that the deleting and insertion conditions imply that  $nv_{k_{\max}+1}(c, H) \leq nv_{k_{\max}+1}(c, G)$ .

Let us assume that  $nv_{k_{\max}+1}(c, H) > nv_{k_{\max}+1}(c, G)$ . Therefore, there is a morphism  $p : C_{k_{\max}+2} \hookrightarrow H$  with  $p \not\models IC_0(d, C')$  for any  $C' \in IG(C_{k_{\max}+2}, C_{k_{\max}+3})$  such that either ?? or ?? is satisfied. Note that this is only the case if  $d \neq \text{false}$ . Otherwise there must be a morphism  $p$  which satisfies ??.

1. There is a morphism  $q' : C_{k_{\max}+2} \hookrightarrow G$  with  $q' \models IC_0(d, C')$  and  $p = \text{tr}_t \circ q'$ .
2. There is a morphism  $q : C_{k_{\max}+2} \hookrightarrow G$  with  $p = \text{tr}_t \circ q$ .

This is a contradiction, if ?? is satisfied,  $q'$  does not satisfy equation (??) and the deletion condition is not satisfied. If (??) is satisfied  $p$  does not satisfy equation (??) since  $p$  only satisfies  $IC_0(d, C_{k+2})$  if  $p$  satisfies  $IC_0(d, C')$  for all  $C' \in IG(C_{k+1}, C_{k+2})$ . In this case, the inserting condition is not satisfied. It follows that

$$nv_k(c, H) \leq nv_k(c, G)$$

holds and  $t$  is a consistency-maintaining transformation. □

Let us now introduce the notion of *direct consistency-increasing* transformations. Similar to the definition of consistency-maintaining and consistency-increasing transformations, 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 transformation. Since a direct consistency-maintaining transformation  $t$  does not introduce any new violations, it is sufficient that  $t$  deletes at least one violation to say that  $t$  is direct consistency-increasing.

We need to check that at least one violation is removed. Again, we need case discrimination if the constraint ends with  $\forall(C_{nl(c)}, \text{false})$  and  $k_{\max} = nl(c) - 2$ . So we introduce two new conditions, one for the general case and one for this special case.

1. *General increasing condition:* This condition is satisfied if either an occurrence of  $C_{k_{\max}+2}$  that does not satisfy  $\exists(k_{\max}+3, \text{true})$  is deleted, or an occurrence of  $C_{k_{\max}+2}$  which does not satisfy  $\exists(C', \text{true})$  in the first graph satisfies  $\exists(C', \text{true})$  in the second graph of the transformation with  $C' \in IG(C_{k_{\max}+2}, C_{k_{\max}+3})$ . Both cases result in the removal of a violation.
2. *Special increasing condition:* This condition is satisfied if an occurrence of  $C_{k_{\max}+2}$  is removed. In the special case that this is the only way to remove a violation.



**Definition 3.17 (direct consistency increasing).** Given a constraint  $c$  in UANF, a rule  $\rho$  and a graph  $G$  with  $G \not\models c$  and let  $e = \text{sub}_{k_{\max}+2}(c)$ .

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 either the special increasing condition is satisfied if  $\text{sub}_{\text{nl}(c)-1}(c) = \forall(C_{\text{nl}(c)}, \text{false})$  and  $k_{\max} = \text{nl}(c) - 2$  or the general increasing condition is satisfied otherwise.

1. General increasing condition:

$$\exists p : C_{k_{\max}+2} \hookrightarrow G \left( \bigvee_{C' \in \text{IG}(k_{\max}+2, k_{\max}+3)} (p \not\models \text{IC}_0(e, C') \wedge (\text{tr}_t \circ p \text{ is not total} \vee \text{tr}_t \circ p \models \text{IC}_0(e, C'))) \right) \quad (3.5)$$

2. Special increasing condition:

$$\exists p : C_{k_{\max}+2} \hookrightarrow G (\text{tr}_t \circ p \text{ is not total}) \quad (3.6)$$

Note that the universally and existentially bound conditions not only ensure that the largest satisfied layer does not decrease, as shown in Lemma ??, but also prevent further unnecessary insertions and deletions, since inserting a universally bound graph and deleting an existentially bound graph will never lead to an increase in consistency.

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

**Theorem 3.4.** Given a constraint  $c$  in UANF, a rule  $\rho$ , a graph  $G$  with  $G \not\models c$  and a direct consistency-increasing transformation  $t : G \Rightarrow_{\rho, m} H$  w.r.t.  $c$ . Then,  $t$  is also a consistency-increasing transformation.

*Proof.* By Theorem ?? it follows that  $t$  is a consistency maintaining transformation. Therefore, it is sufficient to show that  $\text{nv}_{k_{\max}(c, G)+1}(c, H) < \text{nv}_{k_{\max}(c, G)+1}(c, G)$ . Let  $k_{\max} = k_{\max}(c, G)$  and  $d = \text{sub}_{k_{\max}+2}(c)$  with  $d \neq \text{false}$ .

Then, the general increasing condition is satisfied, so there exists a morphism  $p : C_{k_{\max}+2} \hookrightarrow G$  with  $p \not\models \text{IC}_0(d, C')$ , such that either  $\text{tr}_t \circ p$  is total and  $\text{tr}_t \circ p \models \text{IC}_0(d, C')$  or  $\text{tr}_t \circ p$  is not total, for a graph  $C' \in \text{IG}(C_{k_{\max}+2}, C_{k_{\max}+3})$ . In both cases the following applies:

$$\begin{aligned} p &\in \{q \mid q : C_{k_{\max}+2} \hookrightarrow G \wedge q \not\models \text{IC}_0(d, C')\} \wedge \\ \text{tr}_t \circ p &\notin \{q \mid q : C_{k_{\max}+2} \hookrightarrow H \wedge q \not\models \text{IC}_0(d, C')\} \end{aligned}$$

Since  $t$  is direct consistency maintaining, it follows that

$$|\{q \mid q : C_{k_{\max}+2} \hookrightarrow G \wedge q \not\models \text{IC}_0(d, C')\}| \leq |\{q \mid q : C_{k_{\max}+2} \hookrightarrow H \wedge q \not\models \text{IC}_0(d, C')\}|.$$

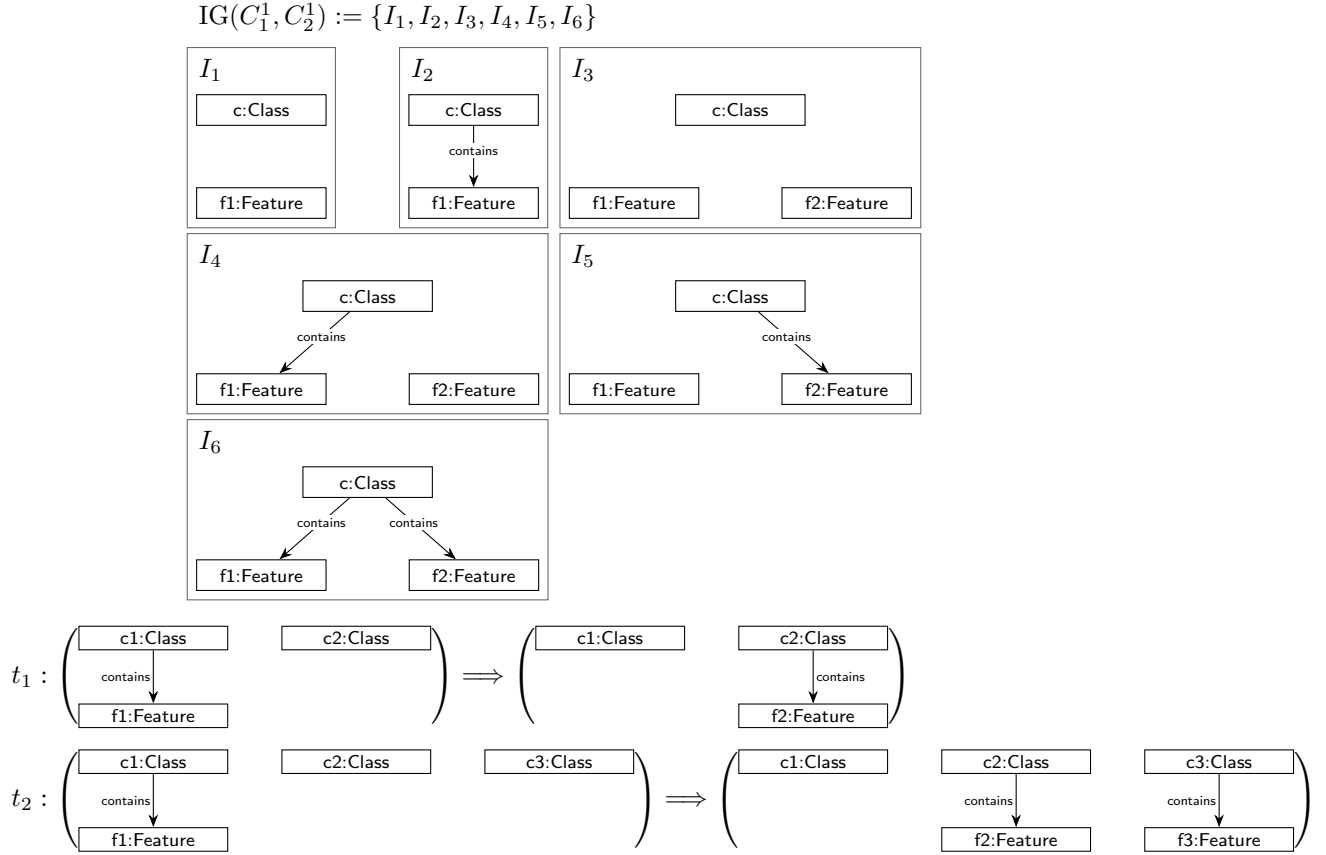


Figure 7: example

for all  $C \in \text{IG}(C_{k_{\max}+2}, C_{k_{\max}+3})$ . Furthermore, this inequality is strictly satisfied if  $C = C'$ . It immediately follows that  $\text{nv}_k(c, G) < \text{nv}_k(c, G')$  and  $t$  is a consistency-increasing transformation.

If  $d = \text{false}$ , i.e.  $\text{sub}_{\text{nl}(c)-1}(c) = \forall(C_{\text{nl}(c)}, \text{true})$  and  $k_{\max} = \text{nl}(c) - 2$ , the special increasing condition is satisfied. It holds that

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

, and since  $t$  is a direct consistency-maintaining transformation, it can be shown in a similar way as above that satisfying the special increasing condition implies that

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

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

**Example 3.7.** Consider constraint  $c_1$  given in Figure ??, the transformations  $t_1, t_2$  and the set  $\text{IG}(C_1^1, C_2^1)$  given in Figure ??. Then,  $t_1$  is a consistency maintaining transformation. The number of violations in both graphs is 7. In the first graph, the occurrence

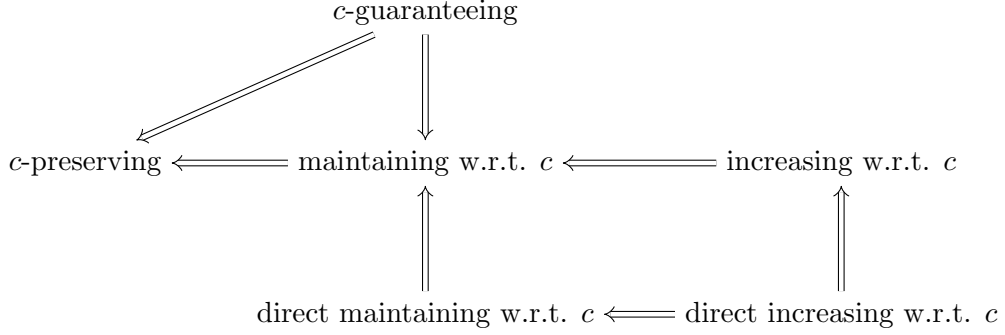


Figure 8: Relations of consistency notions.

$C_1$  does not satisfy  $\exists(\mathcal{I}_4, \text{true})$ ,  $\exists(\mathcal{I}_5, \text{true})$  and  $\exists(\mathcal{I}_6, \text{true})$  and the occurrence  $C_2$  does not satisfy  $\exists(\mathcal{I}_3, \text{true})$ ,  $\exists(\mathcal{I}_4, \text{true})$ ,  $\exists(\mathcal{I}_5, \text{true})$  and  $\exists(\mathcal{I}_6, \text{true})$ . In the second graph, these roles are swapped, i.e.  $C_2$  satisfies exactly the intermediate conditions that  $C_1$  satisfied in the first graph, and vice versa. But,  $t_1$  is not a direct consistency-maintaining transformation, since the occurrence  $C_1$  satisfies  $\exists(\mathcal{I}_3, \text{true})$  in the first but not in the second graph. Therefore, the deleting condition is not satisfied.

The transformation  $t_2$  is consistency increasing w.r.t.  $c_1$ . The number of violations in the first graph is equal to 11. The occurrence  $C_1$  does not satisfy  $\exists(\mathcal{I}_4, \text{true})$ ,  $\exists(\mathcal{I}_5, \text{true})$  and  $\exists(\mathcal{I}_6, \text{true})$ . Both occurrences  $C_2$  and  $C_3$  do not satisfy  $\exists(\mathcal{I}_3, \text{true})$ ,  $\exists(\mathcal{I}_4, \text{true})$ ,  $\exists(\mathcal{I}_5, \text{true})$  and  $\exists(\mathcal{I}_6, \text{true})$ .

In the second graph,  $C_1$  does not satisfy  $\exists(\mathcal{I}_3, \text{true})$ ,  $\exists(\mathcal{I}_5, \text{true})$  and  $\exists(\mathcal{I}_6, \text{true})$  and both  $C_2$  and  $C_3$  do not satisfy  $\exists(\mathcal{I}_6, \text{true})$ . Therefore, the number of violations in the second graph is 5.

But  $t_2$  is not a direct consistency increasing transformation, since  $C_1$  satisfies  $\exists(\mathcal{I}_3, \text{true})$  in the first graph but not in the second, and the deleting condition is not satisfied.

### 3.5 Comparison with other concepts of Consistency

In this chapter, the notions of (direct) consistency increase- and maintainment are compared to the already known notions of consistency guaranteeing, consistency preserving [?], (direct) consistency increasing and sustaining [?], in order to reveal relations between them to and ensure that (direct) consistency increase- and maintainment are indeed a new notions of consistency. These relationships are shown in Figure ??.

First, we compare (direct) consistency increase- and maintainment with the notions of consistency-guaranteeing, preserving, sustaining and improving in the general case and later on, for some special cases. We begin by examining the conclusions that can be drawn about a consistency-maintaining or consistency-increasing transformation.

**Theorem 3.5.** *Given a condition  $c$  in UANF and a transformation  $t : G \Rightarrow H$ . Then,*

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

*Proof.* 1.  $t$  is consistency-maintaining w.r.t.  $c \implies t$  is  $c$ -preserving: Let  $t$  be a consistency-maintaining transformation w.r.t.  $c$ . If  $G \not\models c$ , then  $t$  is a  $c$ -preserving transformation. If  $G \models c$ , then  $\text{nv}_j(c, G) = 0$  for all  $0 \leq j < \text{nl}(c)$ . Since  $t$  is consistency maintaining it follows that  $\text{nv}_j(c, H) = 0$  for all  $0 \leq j < \text{nl}(c)$  and hence  $H \models c$ . It follows that  $t$  is a  $c$ -preserving transformation.

2.  $t$  is consistency maintaining w.r.t.  $c \not\Rightarrow t$  is  $c$ -guaranteeing: Consider the transformation  $t_2 : G \Rightarrow H$  shown in Figure ?? and constraint  $c_1$  shown in Figure ?. As discussed in Example ??,  $t_2$  is consistency- increasing and this consistency-maintaining w.r.t.  $c$ . But,  $t$  is not a  $c$ -guaranteeing transformation, since all occurrences of nodes of type **Class** do not satisfy  $\exists(C_2^1, \text{true})$ .

3.  $t$  is consistency maintaining w.r.t.  $c \not\Rightarrow t$  is consistency sustaining w.r.t.  $c$ : Consider the constraint  $c = \forall(C_1^1, \exists(C_2^1, \forall(C_4^2, d)))$ , where  $d$  is an existentially bound constraint in ANF with  $d \neq \text{false}$  composed of the graphs given in Figure ?. And consider transformation  $t_2$  given in Figure ?. Then,  $t$  is direct consistency increasing; the deletion, increasing, universally and existentially conditions are satisfied and the general increasing condition is satisfied because an occurrence of  $C_1^1$  that did not satisfy  $\exists(C_2^1, \text{true})$  in  $G$  satisfies  $\exists(C_2^1, \text{true})$  in  $H$ . This transformation is not consistency-sustaining since the number of occurrences of  $C_1^1$  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  $\exists(C_2^1, \forall(C_4^2, d))$ .

□

These results are not surprising, since consistency-maintaining and consistency-increasing are much stricter notions than guaranteeing and sustaining, in the sense that the notion of violation is finer-grained. For example, for guaranteeing transformations, an arbitrary number of violations can be introduced as long as the derived graph satisfies the constraint, and thus guaranteeing does not imply direct increasing, since a direct increasing transformation is not allowed to introduce new violations. Let us now examine whether a concept of consistency implies the notions of consistency-maintaining and increasing.

**Theorem 3.6.** *Given a condition  $c$  in UANF and a transformamtion  $t : G \Rightarrow H$ . Then,*

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

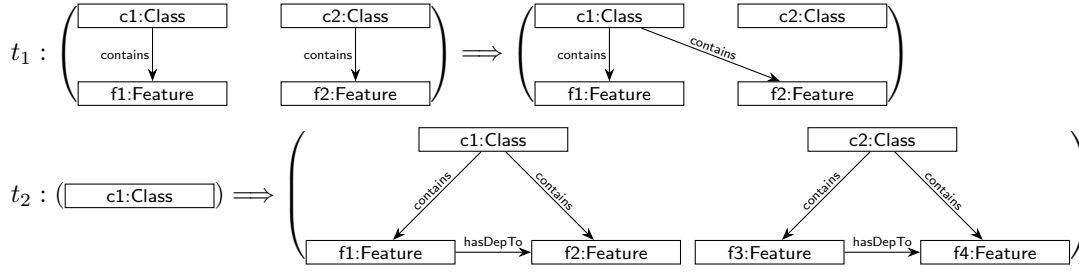


Figure 9: example

- Proof.*
1.  $t$  is  $c$ -guaranteeing  $\implies t$  is consistency-maintaining w.r.t.  $c$ : Let  $t$  be a  $c$ -guaranteeing transformation. Then,  $t$  is also a consistency-maintaining transformation w.r.t.  $c$  since  $H \models c$  and therefore  $\text{nv}_j(c, H) = 0$  for all  $-1 \leq j < \text{nl}(c)$ .
  2.  $t$  is  $c$ -guaranteeing  $\not\Rightarrow t$  is consistency-increasing w.r.t.  $c$ : Assume that  $G \models c$  and  $H \models c$ . Then,  $t$  is a  $c$ -guaranteeing transformation, but not a consistency-increasing one.
  3.  $t$  is  $c$ -preserving  $\not\Rightarrow t$  is consistency-maintaining w.r.t.  $c$ : Consider graphs  $C_1^1$ ,  $C_2^2$  and constraint  $c_1$  given in Figure ?? . The transformation  $t : C_2^2 \implies C_1^1$  is  $c$ -preserving, since  $C_2^2 \not\models c_1$ , but not consistency maintaining w.r.t.  $c$  since  $\text{nv}_0(c, C_2^2) = 2$  and  $\text{nv}_0(c, C_1^1) = 5$ .
  4.  $t$  is direct consistency-improving w.r.t.  $c \not\Rightarrow t$  is consistency-maintaining w.r.t.  $c$ : Consider transformation  $t_1$  given in Figure ?? and constraint  $c_1$  given in Figure ?? . The transformation  $t_1$  is direct consistency improving since no occurrence of  $C_1^1$  is inserted, no occurrence of  $C_1^1$  satisfying  $\exists(C_2^1, \text{true})$  is deleted, and one occurrence  $C_1$  satisfies  $\exists(C_2^1, \text{true})$ . But, this transformation is not consistency-maintaining w.r.t.  $c$  since the number of violations in the first graph is 2 and the number of violations in the second graph is 3.

□

This shows, that in general the notions of (direct) consistency increase and maintainment are not related to (direct) consistency improve- and sustainment. We have only shown some of these relations. Since by definition, (direct) consistency improvement implies (direct) consistency sustainment are related we can conclude results for all pairs of consistency types [?]. An overview of these is given in Table ??.

For some special cases, we can infer other types of relationships.

**Theorem 3.7.** *Given a constraint  $c$  in UANF and a transformation  $t : G \implies H$ .*

1. *If  $G \not\models c$ , then*

$t$  is  $c$ -guaranteeing  $\implies t$  is consistency-increasing w.r.t.  $c$ .

$\implies$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
maintaining (1)	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓
increasing (2)	✓	✓	✗	✗	✗	✗	✗	✗	✗	✓
direct maintaining(3)	✓	✗	✓	✗	✗	✗	✗	✗	✗	✓
direct increasing (4)	✓	✓	✓	✓	✗	✗	✗	✗	✗	✓
improving (5)	✗	✗	✗	✗	✓*	✓*	✗*	✗*	✗*	✓*
sustaining(6)	✗	✗	✗	✗	✗*	✓*	✗*	✗*	✗*	✓*
direct improving (7)	✗	✗	✗	✗	✓*	✓*	✓*	✓*	✗*	✓*
direct sustaining (8)	✗	✗	✗	✗	✗*	✓*	✗*	✓*	✗*	✓*
guaranteeing(9)	✓	✗	✗	✗	✓*	✓*	✓*	✓*	✓**	✓**
preserving (10)	✗	✗	✗	✗	✗*	✗*	✗*	✗*	✗**	✓**

Table 3: Overview of the relationships between consistency concepts, “✓” indicates that the notion in this row implies the notion in the column, and “✗” indicates that this implication does not hold. All results marked with “\*” are from [?] and those marked with “\*\*” are from [?].

2. If  $\text{nl}(c) = 1$ , then

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

*Proof.* 1. Let  $t$  be a  $c$ -guaranteeing transformation with  $G \not\models c$ . Then,  $t$  is a consistency-increasing transformation w.r.t.  $c$  since  $0 < \text{nv}_{k_{\max}(c,G)+1}(c, G) < \infty$  and  $\text{nv}_j(c, H) = 0$  for all  $-1 \leq j \leq \text{nl}(c)$ .

2. Let  $\text{nl}(c) = 1$ . Since  $c$  is in UANF,  $\text{sub}_1(c) = \text{false}$  and  $\text{nv}_0(c, G)$  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 immediately follows.  $\square$

## 4 Application Conditions

In the following, we present application conditions which ensure that any rule equipped with this application condition is ( direct ) consistency-increasing or ( direct ) consistency-maintaining. In particular, we present applications in the general case and later for specific rules, called *basic maintaining/increasing* rules. For basic rules, less complex application conditions can be constructed. Similar to the notions of consistency-maintaining and consistency-increasing, these application conditions will only consider graphs of the constraint up to a certain layer. This is useful to reduce the restrictiveness of these application conditions, since all graphs  $C_j$  with  $j < k_{\max} + 2$  do not affect whether a transformation is consistency-maintaining or consistency-increasing. So it is clear that a rule with these application conditions is not a consistency-maintaining or consistency-increasing rule. Therefore, let us introduce a weaker notion of consistency-maintaining

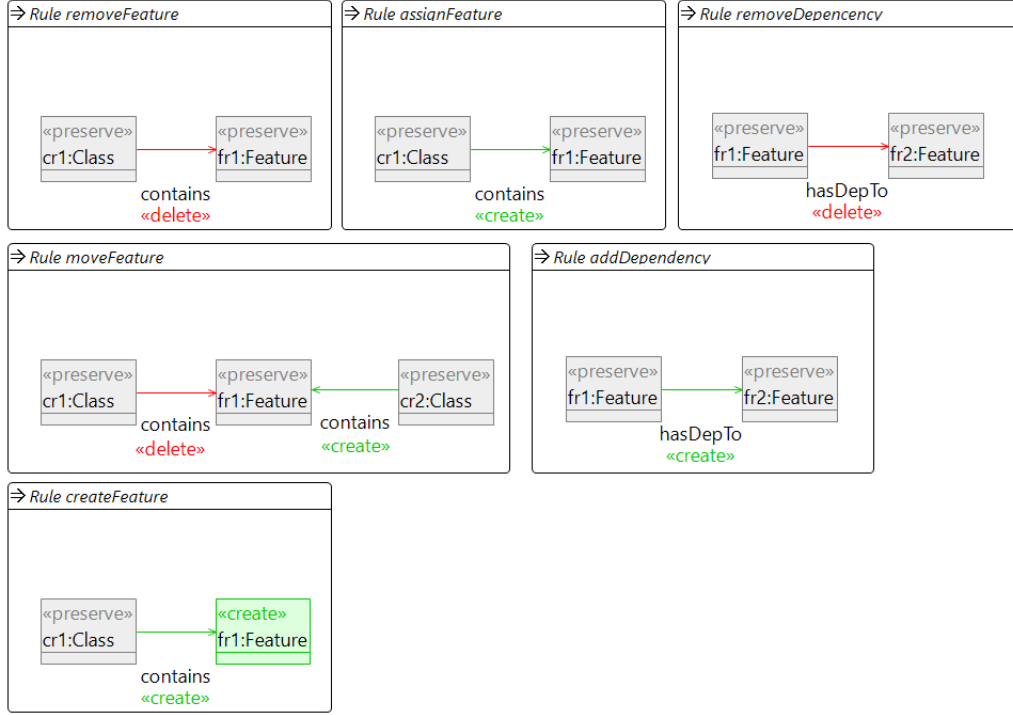


Figure 10: rules

$$\begin{aligned}
c = & \forall \left( \boxed{\text{cc1:Class}} ; \exists \left( \begin{array}{c} \boxed{\text{cc1:Class}} \\ \swarrow \text{contains} \searrow \text{contains} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array} , \forall \left( \begin{array}{c} \boxed{\text{cc1:Class}} \\ \swarrow \text{contains} \searrow \text{contains} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array} \xrightarrow{\text{hasDepTo}} \boxed{\text{cf3:Feature}} , \exists \left( \begin{array}{c} \boxed{\text{cc1:Class}} \quad \boxed{\text{cc2:Class}} \\ \swarrow \text{contains} \searrow \text{contains} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array} \xrightarrow{\text{hasDepTo}} \boxed{\text{cf3:Feature}} , \text{true} \right) \right) \right) \\
\text{ned}_{-1} = & \text{true} \\
\text{ned}_1 = & \neg \exists \left( \begin{array}{c} \boxed{\text{cr1} = \text{cc1:Class}} \\ \swarrow \text{contains} \searrow \text{contains} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{fr1} = \text{cf2:Feature}} \end{array} , \text{true} \right) \\
\text{ned}_3 = & \neg \exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \swarrow \text{contains} \searrow \text{contains} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{fr1} = \text{cf2:Feature}} \end{array} , \text{true} \right) \wedge \neg \exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \quad \boxed{\text{cc2:Class}} \\ \swarrow \text{contains} \searrow \text{contains} \\ \boxed{\text{fr1} = \text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array} \xrightarrow{\text{hasDepTo}} \boxed{\text{cf3:Feature}} , \text{true} \right) \wedge \\
\neg \exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \quad \boxed{\text{cc2:Class}} \\ \swarrow \text{contains} \searrow \text{contains} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{fr1} = \text{cf2:Feature}} \end{array} \xrightarrow{\text{hasDepTo}} \boxed{\text{cf3:Feature}} , \text{true} \right) \wedge \neg \exists \left( \begin{array}{c} \boxed{\text{cc1:Class}} \quad \boxed{\text{rc1} = \text{cc2:Class}} \\ \swarrow \text{contains} \searrow \text{contains} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array} \xrightarrow{\text{hasDepTo}} \boxed{\text{fr1} = \text{cf3:Feature}} , \text{true} \right)
\end{aligned}$$

Figure 11: no existentially deleted conditions for the constraint and rule **removeFeature**.

and consistency-increasing rules, called *(direct) consistency-increasing rule at layer* and *(direct) consistency-maintaining rules at layer* respectively.

As the name suggests, a rule is called (direct) consistency-increasing at layer  $k$  (or (direct) consistency-maintaining at layer  $k$ ) if all its transformations starting with graphs with  $k_{\max} = k$  are (direct) consistency-increasing or (direct) consistency-maintaining, respectively.

**Definition 4.1 ((direct) consistency increasing and maintaining rule at layer).**

Let a constraint  $c$  in UANF 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  $t : G \Longrightarrow_{\rho} H$ , with  $k_{\max}(G, c) = k$ , are (direct) consistency maintaining w.r.t.  $c$ . The rule  $\rho$  is called (direct) consistency increasing at layer  $0 \leq k < \text{nl}(c)$  w.r.t.  $c$  if all transformations  $t : G \Longrightarrow_{\rho} H$ , with  $k_{\max}(G, c) = k$ , 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 rule w.r.t.  $c$ , since all graphs to which this rule is applied satisfy  $c$ . For the same reason, there is no consistency-increasing transformation at layer  $\text{nl}(c) - 1$  w.r.t.  $c$ .

#### 4.1 General Application Conditions

Let us start by introducing consistency-maintaining application conditions, i.e., a rule equipped with this application condition is direct consistency-maintaining at a given layer. We will also introduce an application condition such that any rule equipped with it is a consistency-maintaining rule.

The maintaining application condition has a odd parameter  $k$  which specifies which graphs of a constraint are to be considered. In particular, only the graphs  $C_j$  with  $0 \leq j \leq k + 3$  are considered. Note that there is no graph with  $k_{\max}$  odd and  $k_{\max} \neq \text{nl}(c) - 1$ . So it is not a restriction that  $k$  must be odd. The maintaining application condition consists of the following three parts, which also use the  $k$  parameter:

1. *No existentially destroyed* ( $\text{ned}_k()$ ): This application condition checks that no occurrences of existentially bound graphs  $C_j$  with  $2 \leq j \leq k + 1$  are removed. It corresponds to the existentially condition, since a rule that satisfies this application condition also satisfies the existentially condition when applied to a graph with  $k_{\max} = k$ .
2. *No universally inserted* ( $\text{nui}_k()$ ): This application condition checks that no occurrences of universally bound graphs  $C_j$  with  $1 \leq j \leq k + 2$  are inserted. It corresponds to the universally and the inserting conditions, in the sense that, a rule that satisfies this application condition also satisfies the universally and inserting conditions when applied to a graph with  $k_{\max} = k$ .
3. *No violation inserted* ( $\text{vio}_k()$ ): This application condition checks that no new violations are introduced by removing occurrences of intermediate graphs from the set  $\text{IG}(C_{k+2}, C_{k+3})$ . It corresponds to the deleting condition in the sense that



a rule that satisfies this application condition also satisfies the deleting condition when applied to a graph with  $k_{\max} = k$ . There are several cases for this application condition. If  $k = \text{nl}(c) - 2$ , the constraint ends with  $\forall(C_{\text{nl}(c)}, \text{false})$ . Therefore no violations can be introduced by removing occurrences of intermediate graphs. In particular, there is no graph  $C_{k+3}$ . If  $k = \text{nl}(c) - 1$ , a rule equipped with this application condition always satisfies the deleting condition when applied to graphs with  $k_{\max} = k$ . Therefore, if  $k \geq \text{nl}(c) - 2$ , we set the application condition to **true**.

Recall that given a constraint  $c$  in UANF, each subcondition  $\text{sub}_k(c)$  is a condition over the graph  $C_k$  and the morphism is denoted by  $a_k$ .

**Definition 4.2 (consistency maintaining application condition).** *Given a rule  $\rho = (\text{ac}, \rho')$  with  $\rho' = L \longleftarrow K \longrightarrow R$ , a constraint  $c$  in UANF and an odd  $-1 \leq k < \text{nl}(c)$ . The maintaining application condition of  $c$  for  $\rho$  at layer  $k$  is defined as  $\text{ac} \wedge \text{main}_k(\rho')$  with*

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

and

1. *No existentially destroyed: Let  $E$  be the set of all existentially bound graphs  $C_j$  with  $2 \leq j \leq k+1$  and  $\mathbf{P}_{C_j}$  be the set all overlaps  $P'$  of  $L$  and  $C_j$  with  $i_L^{P'}(L \setminus K) \cap i_{C_j}^{P'}(C_j \setminus C_{j-1}) \neq \emptyset$ :*

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

2. *No universally inserted: Let  $U$  be the set of all universally bound graphs  $C_j$  with  $1 \leq j \leq k+2$ , and  $\mathbf{P}_{C_j}$  be the set of all overlaps  $P'$  of  $R$  and  $C_j$  with  $i_R^{P'}(R \setminus K) \cap i_{C_j}^{P'}(C_j \setminus C_{j-1}) \neq \emptyset$ :*

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

3. *No violation inserted: Let  $\mathbf{P}_{C'}$  be the set of all overlaps  $P$  of  $L$  and  $C'$  with  $i_L^P(L \setminus K) \cap i_{C'}^P(C' \setminus C_{k+2}) \neq \emptyset$  for  $C' \in \text{IG}(C_{k+2}, C_{k+3})$ :*

$$\text{vio}_k(\rho') := \begin{cases} \text{true} & \text{if } k \geq \text{nl}(c) - 2 \\ \bigwedge_{C' \in \text{IG}(C_{k+2}, C_{k+3})} \bigwedge_{P \in \mathbf{P}_{C'}} \neg \exists(i_L^P : L \hookrightarrow P', \text{true}) & \text{otherwise} \end{cases}$$

**Example 4.1.** 1. Consider the constraint  $c$  given in Figure ?? and the rule **removeFeature**. The conditions  $\text{ned}_{-1}(\text{removeFeature})$ ,  $\text{ned}_1(\text{removeFeature})$  and  $\text{ned}_3(\text{removeFeature})$  are also given in Figure ??;  $\text{ned}_{-1}(\text{removeFeature})$  is equal to **true** because there is no existentially bound graph  $C_0$ .  $\text{ned}_1(\text{removeFeature})$  checks that no occurrences

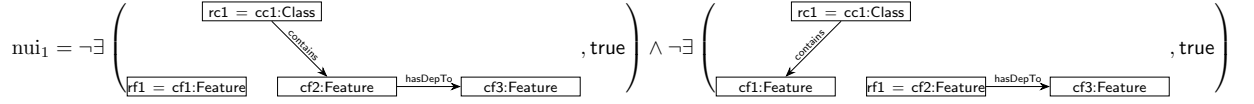


Figure 12: No universally inserted conditions for the constraint given in Figure ?? and rule `assignFeature`.

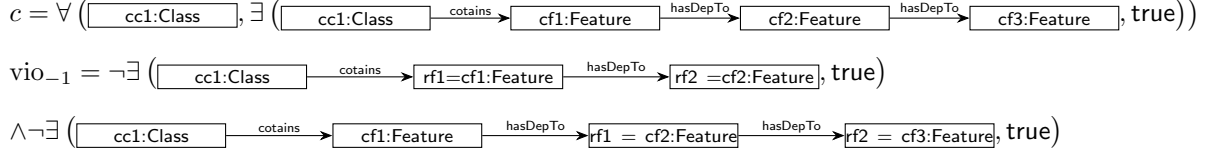


Figure 13: No violation inserted conditions for the constraint and rule `removeDependency`.

of  $C_2$  are inserted, while  $\text{ned}_3(\text{removeFeature})$  checks that no occurrences of  $C_2$  and  $C_4$  are inserted. Obviously,  $\text{ned}_1(\text{removeFeature})$  is contained in  $\text{ned}_3(\text{removeFeature})$ . Note that there are conditions in  $\text{ned}_3(\text{removeFeature})$  that imply each other. For example, the first condition implies the second and third. So they can be removed.

2. Again consider constraint  $c$  given in Figure ?? and the rule `assignFeature`. The condition  $\text{nui}_{-1}(\text{assignFeature})$  is equal to `true` since `assignFeature` cannot create a node of type `Class`. The condition  $\text{nui}_1(\text{assignFeature})$  is given in Figure ?. The condition  $\text{nui}_3(\text{assignFeature})$  is equal to  $\text{nui}_1(\text{assignFeature})$  since there is no universally bound graph  $C_j$  with  $j > 3$ .
3. Consider the constraint  $c$  given in Figure ?? and the rule `addDependency`. Then,  $\text{vio}_{-1}(\text{addDependency})$  is given in Figure ?. The condition  $\text{vio}_1(\text{addDependency})$  is equal to `true` since  $1 > \text{nl}(c) - 2 = 0$ .

Let us now show that every rule equipped with  $\text{main}_k()$  is a consistency-maintaining rule at layer  $k$ .

**Theorem 4.1.** *Given a constraint  $c$  in UANF. Every rule  $\rho = (\text{ac} \wedge \text{main}_k(\rho'), \rho')$  with  $-1 \leq k < \text{nl}(c)$  odd is a consistency maintaining rule at layer  $k$  w.r.t.  $c$ .*

*Proof.* Given a graph  $G$  with  $k_{\max} = k$  and a transformation  $t : G \Rightarrow_{\rho, m} H$ . We show that  $t$  is a direct consistency maintaining transformation w.r.t.  $c$ .

We show that  $t$  satisfies the deleting, inserting, universally and existentially conditions.

1. Assume that  $t$  does not satisfy the deleting condition. Then  $k_{\max} < \text{nl}(c) - 1$ ,  $e = \text{sub}_{k+2}(c) \neq \text{false}$  and there is a morphism  $p : C_{k+2} \hookrightarrow G$  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, there is an overlap  $P$  of  $L$  and  $C'$  such that  $i_{C'}^P \circ a_{k+2}^r \models \exists(a_{k+2}^r : C_{k+2} \hookrightarrow C', \text{true})$

with  $i_L^P(L \setminus K) \cap i_{C'}^P(C' \setminus C_{k+2}) \neq \emptyset$  and  $m \models \exists(i_L^P : L \hookrightarrow P, \text{true})$ . Thus,  $\text{vio}_k(\rho')$  and consequently also  $\text{main}_k(\rho')$  cannot be satisfied.

2. Assume that  $t$  does not satisfy the inserting condition. 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, there is a morphism  $p' : C_{k+2} \hookrightarrow H$  with  $p' \not\models d$  such that no morphism  $p : C_{k+2} \hookrightarrow G$  with  $\text{tr}_t \circ p = p'$  exists. Therefore, there is an overlap  $P$  of  $R$  and  $C_{k+2}$  with  $i_R^P(R \setminus K) \cap i_{C_{k+2}}^P(C_{k+2} \setminus C_{k+1}) \neq \emptyset$  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  $t$  does not satisfy the universally condition. Then, there is a morphism  $p : C_j \hookrightarrow H$  with  $0 \leq j < k$  and  $C_j$  universally bound such that no morphism  $p' : C_j \hookrightarrow G$  with  $\text{tr}_t \circ p' = p$  exists. Then, there is an overlap  $P$  of  $C_j$  and  $R$  with  $i_R^P(R \setminus K) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$  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  $t$  does not satisfy the existentially condition. There is a morphism  $p : C_j \hookrightarrow G$  with  $j \leq k+1$  and  $C_j$  existentially bound and such that  $\text{tr}_t \circ p$  is not total. Then, there is an overlap  $P$  of  $C_j$  and  $L$  with  $i_L^P(L \setminus K) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$ , such that  $m \models \exists(i_L^P : L \hookrightarrow P, \text{true})$ . Hence,  $m \not\models \text{ned}_k(\rho')$ .

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

**Theorem 4.2.** *Given a constraint  $c$  in UANF. Every rule  $\rho$  equipped with the application condition*

$$\left( \bigwedge_{\substack{-1 \leq i < \text{nl}(c) \\ i \text{ odd}}} \text{vio}_i(\rho) \right) \wedge \text{nui}_{\text{nl}(c)-1}(\rho)$$

*is a consistency-maintaining rule w.r.t.  $c$ .*

*Proof.* Let  $\rho = R \xleftarrow{r} K \xrightarrow{l} L$  be a rule equipped with this application condition. We show that  $\rho$  is a consistency-maintaining rule at layer  $k$  w.r.t.  $c$  for all  $-1 \leq k \leq \text{nl}(c) - 1$ . Obviously,  $\text{nui}_{\text{nl}(c)-1}(\rho)$  contains  $\text{nui}_j(\rho)$  for all  $-1 < j \leq \text{nl}(c) - 1$ . The set of intermediate graphs always contains both graphs on which this set was built, so  $\text{vio}_i(\rho)$  contains the condition

$$\bigwedge_{P' \in \mathbf{P}_{C_{i+3}}} \neg \exists(i_L^{P'} : L \hookrightarrow P', \text{true})$$

which checks that no occurrence of  $C_{i+3}$  is deleted. Therefore

$$\left( \bigwedge_{\substack{-1 \leq i < \text{nl}(c) \\ i \text{ odd}}} \text{vio}_i(\rho) \right)$$

must contain  $\text{ned}_{\text{nl}(c)-1}(\rho)$  and therefore it contains  $\text{ned}_j(\rho)$  for all  $-1 \leq j \leq \text{nl}(c) - 1$ . So we can rewrite this application condition into the equivalent condition

$$\left( \bigwedge_{\substack{-1 \leq i < \text{nl}(c) \\ i \text{ odd}}} \text{vio}_i(\rho) \wedge \text{nui}_i(\rho) \wedge \text{ned}_i(\rho) \right) = \left( \bigwedge_{\substack{-1 \leq i < \text{nl}(c) \\ i \text{ odd}}} \text{main}_i(\rho) \right)$$

It follows that  $\rho$  is a consistency-maintaining rule at layer  $k$  for all  $-1 \leq k < \text{nl}(c)$ . Hence,  $\rho$  is a consistency-maintaining rule w.r.t.  $c$ .  $\square$

The *extended overlap* will be a useful tool for our consistency-increasing application conditions. Intuitively, given an overlap and a morphism  $a$ , the overlap is extended such that one overlap morphism satisfies  $\exists(a, \text{true})$ .

**Definition 4.3 (extended overlaps).** *Given an overlap  $(G, i_{C_0}^G, i_{C_1}^G)$  of  $C_0$  and  $C_1$  and a morphism  $e : C_0 \hookrightarrow H$ . The set of extended overlaps of  $G$  at  $i_{C_0}^G$  with  $e$ , denoted by  $\text{eol}(G, i_{C_0}^G, e)$  is defined as:*

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

Using extended overlaps, we will be able to check whether an violation has been removed. Let us now consider consistency-increasing application conditions. For these we will use the maintaining application conditions described above. All that remains is to ensure that at least one violation is removed. To do this, we must first check that there is a violation in the match. This means that the match and a violation are overlapping. Finally, we need to check that this violation is removed.

Again, the increasing application condition has the odd parameter  $-1 \leq k < \text{nl}(c)-1$ , which specifies which constraint graphs are to be considered. Note that  $k$  must not be  $\text{nl}(c) - 1$ , since all graphs with  $k_{\max} = \text{nl}(c)$  satisfy the constraint. Therefore for these graphs there can be no consistency increasing transformations. It also has a second parameter  $C'$  which is an intermediate graph of  $C_{k+2}$  and  $C_{k+3}$  if  $C_{k+3}$  exists and is set to  $C_{k+2}$  otherwise. Again, a rule equipped with this application condition is a consistency-increasing rule at layer  $k$ . It consists of the following parts:

1. The maintaining application condition ( $\text{main}_k()$ ): As already discussed, this application condition ensures that a rule equipped with this application condition is a consistency-maintaining rule at layer  $k$ .
2. *Violation exists* ( $\text{exv}()$ ): This condition checks that there is a violation at the match, i.e. there is an occurrence  $p$  of  $C_{k+2}$  with  $m(L) \cap p(C_{k+2}) \neq \emptyset$  not satisfying  $\exists(C', \text{true})$ . There is a special case when  $k = \text{nl}(c) - 2$  and  $c$  ends with  $\forall(C_{\text{nl}(c)}, \text{false})$ , it is sufficient to check only that  $m(L) \cap p(C_{k+2}) \neq \emptyset$ . If there are no such occurrences, the transformation cannot be consistency-increasing.
3. *Violation removed* ( $\text{remv}()$ ): This condition checks that the violation is removed. This can be done in several ways, either by deleting the occurrence  $p$  or by inserting

elements such that  $p \models \exists(C', \text{true})$ . This leads to case discrimination. The first case is easy to check, if  $m(L \setminus K) \cap p(C_{k+2} \setminus C_{k+1}) \neq \emptyset$ ,  $p$  is removed and this condition can be set to **true**. Otherwise, we need to check that the violation has been removed by an additional condition that checks whether  $p$  satisfies  $\exists(C', \text{true})$  after the transformation. The last case is the special case where the constraint ends with  $\forall(C_{\text{nl}(c)}, \text{true})$  and  $k = \text{nl}(c) - 2$ . Then there is only one way to remove a violation, by removing the occurrence  $p$ . So the condition is set to **true** if  $m(L \setminus K) \cap p(C_{k+2} \setminus C_{k+1}) \neq \emptyset$  and to **false** otherwise.

**Definition 4.4 (consistency increasing application condition).** *Given a rule  $\rho = (\text{ac}, \rho')$  with  $\rho' = L \xleftarrow{l} K \xrightarrow{r} R$  and a constraint  $c$  in UANF. Let  $0 \leq k < \text{nl}(c) - 1$  be odd and  $C' = C_{k+2}$  if  $k = \text{nl}(c) - 2$  and  $C' \in \text{IG}(C_{k+2}, C_{k+3})$  otherwise. The increasing application condition of  $c$  for  $\rho$  at layer  $k$  with  $C'$  is defined as*

$$\text{incr}_k(C', \rho) := \text{ac} \wedge \text{main}_k(\rho') \wedge \left( \bigvee_{P \in \text{ol}(L, C_{k+2})} \text{exv}(P, C') \wedge \text{remv}(P, C') \right) \quad (4.1)$$

with

1. *Violation exists:* Let  $a_{k+2}^r : C_{k+2} \hookrightarrow C$  be the restricted morphism of  $a_{k+2}$  and  $i_L^P$  and  $i_P^Q$  be overlap morphisms of  $P$  and  $Q$ , respectively:

$$\text{exv}(P, C') := \begin{cases} \exists(i_L^P : L \hookrightarrow P, \text{true}) & \text{if } \text{sub}_{k+2}(c) = \text{false} \\ \exists(i_L^P : L \hookrightarrow P, \bigwedge_{Q \in \text{ol}(P, i_{C_{k+2}}^P, a_{k+2}^r)} \neg \exists(i_P^Q : P \hookrightarrow Q, \text{true})) & \text{otherwise} \end{cases}$$

2. *Violation removed:*

- (a) *If  $i_L^P(L \setminus K) \cap i_{C_{k+2}}^P(C_{k+2} \setminus C_{k+1}) \neq \emptyset$ ,  $\text{remv}(P, C') := \text{true}$ .*
- (b) *If  $\text{sub}_{k+2}(c) = \text{false}$ , i.e.  $k = \text{nl}(c) - 2$ ,*

$$\text{remv}(P, C') := \begin{cases} \text{true} & \text{if } i_L^P(L \setminus K) \cap i_{C_{k+2}}^P(C_{k+2} \setminus C_{k+1}) \neq \emptyset \\ \text{false} & \text{otherwise.} \end{cases}$$

- (c) *Otherwise, let  $P'$  be the graph derived by the transformation  $P \Rightarrow_{\rho, m} P'$ . Then,  $P'$  is an overlap of  $R$  and  $C_{k+2}$ . If this transformation does not exist, we set  $\text{remv}(P, C') := \text{false}$  and otherwise*

$$\text{remv}(P, C') := \bigvee_{Q \in \text{ol}(P', i_{C_{k+2}}^{P'}, a^r)} \text{Left}(\forall(i_R^Q : R \hookrightarrow P', \exists(i_{P'}^Q : P' \hookrightarrow Q, \text{true})), \rho).$$

Note that  $\text{incr}_k(C', \rho)$  is only evaluated to **true** if an occurrence of  $p : C_{k+2} \hookrightarrow G$  that does not satisfy  $\exists(C', \text{true})$  is either removed or satisfies  $\exists(C', \text{true})$  in the derived graph. For all smaller improvements, i.e. a similar improvement for a subgraph  $C'' \in$

$\text{IG}(C_{k+1}, C')$  of  $C'$ ,  $\text{incr}_k(C', \rho)$  would be evaluated as **false**. For any larger improvements, i.e. the same improvement for a supergraph  $C'' \in \text{IG}(C', C_{k+2})$  of  $C'$ ,  $\text{incr}_k(C', \rho)$  will also be evaluated as **false** if the repaired occurrence of  $C_{k+2}$  satisfies  $\exists(C', \text{true})$ . In both cases, the application condition would prohibit the transformation, even though it would be consistency-increasing. To solve this problem, several 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 as **true** if the cases described above occur, with the drawback that this results in a huge condition, even if duplicate conditions are removed. At least all duplicates of  $\text{main}()$  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 compromise between condition size and restrictiveness. They are very restrictive because they do not allow deletions of occurrences of existentially bound graphs and insertions of universally bound graphs. For example, any of these application conditions with the rule **moveFeature** and the constraint  $c_1$  will be equivalent to **false**; the maintaining part of the condition will always be evaluated as **false**, since **moveFeature** always removes occurrences of the existentially bound graph  $C_2^1$ . Changing the conditions constructed by  $\text{main}()$  to check whether two nodes of type **Feature** are connected to a node of type **Class** will give application conditions that can be satisfied with **moveFeature**. However, for a similar rule moving two nodes of type **Feature**, this newly constructed  $\text{main}()$  would still be evaluated as **false**. So this would only lead to a slight decrease in restrictiveness.

The conditions constructed by  $\text{ned}()$  and  $\text{nui}()$  could be modified in a similar way. For  $\text{ned}()$  and an occurrence  $p$  of the universally bound graph  $C_j$ , by checking whether there exist two occurrences  $p_1, p_2$  of  $C_j$  with  $p = p_1 \circ a_j = p_2 \circ a_j$ , and for  $\text{nui}()$ , by checking whether an introduced occurrence  $p$  of  $C_j$  satisfies  $\exists(C_{j+1}, \text{true})$ . The construction of these is similar to the construction of consistency guaranteeing application conditions as introduced by Habel and Pennemann [?], which is known to construct huge application conditions. Also, the application conditions constructed above become more and more restrictive as  $k$  increases, since the number of conditions also increases.

**Example 4.2.** 1. Consider constraint  $c_1$  given in Figure ?? and the rule **assignFeature**.

There is only one overlap  $P$  of  $L$  and  $C_1^1$  which is shown in Figure Consider Figure ?. The  $\text{exv}(P, C')$  and  $\text{exv}(P, C')$  parts of  $\text{incr}_{-1}(C', \text{assignFeature})$  with  $C' = C_2^2$  and  $C' = C_2^1$ , respectively, are also given in Figure ?.

2. Consider the constraint and rule given in Figure ?. The application increasing application condition of this rule and at layer  $-1$  with  $C_1$  of the constraint is also given in this Figure.

Note that we have already used the Left operator in all of these conditions.

Let us now show that a rule equipped with the consistency-increasing transformation condition at layer  $k$  is indeed a consistency-increasing rule at layer  $k$ .

$P$   

rc1 = cc1:Class

rf1 = cf1:Feature

$$\begin{aligned}
\text{exv}(P, C_2^2) &= \forall \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \end{array} , \neg \exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \downarrow \text{contains} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \end{array} , \text{true} \right) \right) \\
\text{remv}(P, C_2^2) &= \forall \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \end{array} , \neg \exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \end{array} , \text{true} \right) \right) \\
\text{exv}(P, C_2^1) &= \exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \end{array} , \neg \exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \downarrow \text{contains} \quad \downarrow \text{contains} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array} , \text{true} \right) \wedge \neg \exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \quad \boxed{\text{rf1} = \text{cf3:Feature}} \\ \downarrow \text{contains} \quad \downarrow \text{contains} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array} , \text{true} \right) \right) \\
\text{remv}(P, C_2^1) &= \forall \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \end{array} , \neg \exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \downarrow \text{contains} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array} , \text{true} \right) \right)
\end{aligned}$$

Figure 14: Examples for  $\text{exv}(P, C_2^2)$ ,  $\text{exv}(P, C_2^1)$ ,  $\text{remv}(P, C_2^2)$  and  $\text{remv}(P, C_2^1)$  using the rule **assignFeature** and constraint  $c_1$  given in Figure ??.

$$\begin{aligned}
c &= \forall \left( \begin{array}{c} \boxed{\text{cc1:Class}} \quad \boxed{\text{cf3:Feature}} \\ \downarrow \text{contains} \quad \uparrow \text{hasDepTo} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array} , \text{false} \right) \\
\text{incr}_{-1}(C_1) &= \text{true} \wedge \left( \left( \begin{array}{c} \boxed{\text{cc1:Class}} \quad \boxed{\text{cf3:Feature}} \\ \downarrow \text{contains} \quad \uparrow \text{hasDepTo} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \quad \boxed{\text{rf2} = \text{cf2:Feature}} \end{array} , \text{true} \right) \wedge \text{true} \right) \vee \left( \left( \begin{array}{c} \boxed{\text{cc1:Class}} \quad \boxed{\text{rf2} = \text{cf3:Feature}} \\ \downarrow \text{contains} \quad \uparrow \text{hasDepTo} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{rf1} = \text{cf2:Feature}} \end{array} , \text{true} \right) \wedge \text{true} \right)
\end{aligned}$$

Figure 15: Example of  $\text{incr}_{-1}(C_1, \rho)$  using this constraint and the rule **removeDependency**.

**Theorem 4.3.** *Given a constraint  $c$  in UANF. Every rule  $\rho = (\text{ac} \wedge \text{incr}_k(C', \rho'), \rho')$  with  $-1 \leq k < \text{nl}(c) - 1$  odd and  $C' \in \text{IG}(C_{k+2}, C_{k+3})$  if  $k < \text{nl}(c) - 2$  and  $C' = C_{k+2}$  otherwise is a consistency-increasing rule at layer  $k$  w.r.t.  $c$ .*

*Proof.* Let a transformation  $t : G \Rightarrow_\rho H$  with  $k_{\max}(c, G) = k$  be given. Since  $\text{incr}_k(C, \rho)$  contains  $\text{main}_k(\rho)$ ,  $t$  is a consistency maintaining transformation at layer  $k$  according to Theorem ???. It remains to show that  $t$  satisfies the general and special increasing conditions respectively.

1. If  $k = \text{nl}(c) - 2$ , i.e.  $c$  ends with a condition of the form  $\forall(C_{\text{nl}(c)}, \text{false})$ , assume that  $t$  does not satisfy the special increasing condition. Then there is no morphism  $p : C_{k+2} \hookrightarrow G$  such that  $\text{tr}_t \circ p$  is not total. So there is no overlap  $P$  of  $L$  and  $C_{k+2}$  with  $i_L^P(L \setminus K) \cap i_{C_{k+2}}^P(C_{k+2} \setminus C_{k+1}) \neq \emptyset$ . Since,  $k = \text{nl}(c) - 2$  it follows that  $\text{remv}(P, C') = \text{false}$  and  $m \not\models \text{incr}_k(C, \rho)$ . This is a contradiction.
2. Otherwise let  $P \in \text{ol}(L, C_{k+2})$ . We show that  $m \models \text{exv}(P, C) \wedge \text{remv}(P, C)$  implies that  $t$  satisfies the general increasing condition. If  $m \models \text{exv}(P, C)$ , there is 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  $Q \in \text{eol}(P, i_{C_{k+2}}^P, a_{k+2}^r)$ . Therefore,  $q := p \circ i_{C_{k+2}}^P \not\models \exists(a_{k+2}^r : C_{k+2} \hookrightarrow C', \text{true})$ .

If  $i_L^P(L \setminus K) \cap i_{C_{k+2}}^P(C_{k+2} \setminus C_{k+1}) \neq \emptyset$ ,  $\text{tr}_t \circ q$  is not total and the general increasing condition is satisfied.

Otherwise, if  $\text{tr}_t \circ q$  is total, and there is a morphism  $p : P' \hookrightarrow H$  with  $\text{tr}_t \circ q = p \circ i_{C_{k+1}}^{P'}$ . Since  $m \models \text{remv}(P, C)$ , all morphisms  $p \circ i_{C_{k+2}}^{P'}$  with  $n = p \circ i_R^Q$  satisfy  $\exists(C', \text{true})$ . Therefore,  $\text{tr}_t \circ q \models \exists(C', \text{true})$  and the general increasing condition is satisfied.

In summary,  $\rho$  is a consistency increasing rule at layer  $k$ . □

## 4.2 Basic Consistency-increasing and Consistency-maintaining Rules

The construction of the application conditions introduced in the previous section, as well as the constructed application conditions themselves, are very complex. For a certain set of rules, which we will call *basic consistency-increasing rules*, 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 way. The main idea is that these rules (a) are not able to delete occurrences of existentially bound graphs or insert occurrences of universally bound graphs and (b) are able to increase consistency at a certain layer. That is, given a basic increasing rule  $\rho$ , there exists a transformation  $t : G \Rightarrow_\rho H$  such that  $t$  is a consistency increasing transformation with respect to a constraint  $c$ .

To ensure that (a) is satisfied, we first 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 actually a subset of the set of consistency-maintaining rules up to layer, since these rules must be plain rules, whereas consistency-maintaining rules up to layer are allowed to have application conditions, i.e.  $\text{main}(\cdot, \cdot)$ .

**Definition 4.5 (basic consistency maintaining rule up to layer).** *Given a plain rule  $\rho$  and a constraint  $c$  in UANF. Then,  $\rho$  is called basic consistency maintaining rule up to layer  $-1 \leq k < \text{nl}(c)$  w.r.t.  $c$  if it is a direct consistency-maintaining rule at layer  $k$  w.r.t.  $c$ .*

**Example 4.3.** *Consider the rules `moveFeature`, `assignFeature` and `addDependency` given in Figure ?? and constraints  $c_1$  and  $c_2$  given in Figure ?. The rule `assignFeature` is a basic consistency maintaining rule up to layer 1 w.r.t.  $c_1$ , whereas `moveFeature` is not a basic consistency maintaining rule w.r.t.  $c_1$ . The rule `addDependency` is a basic consistency-maintaining rule up to layer  $-1$  w.r.t.  $c_2$ , but is not a basic consistency-maintaining rule up to layer 1 w.r.t.  $c_2$  since it can insert occurrences of  $C_3^2$ .*

Since there are infinitely many transformations via a plain rule  $\rho$ , it is impossible to check whether  $\rho$  is a basic consistency maintaining rule up to a level based on the definition above. Therefore, we present a characterisation of basic consistency-maintaining rules that relies only 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)  $\rho$  inserts an edge of  $C_j \setminus C_{j-1}$  connecting pre-existing nodes of  $C_j$ , since it is unclear whether this would create a new occurrence of  $C_j$ , or (b) if  $\rho$  inserts a node  $v$  of  $C_j$ , so 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 will not create an occurrence of  $C_j$ , since  $v$  is only connected to edges that have also been inserted by  $\rho$ .

Second, suppose  $\rho$  is able to delete occurrences of an existentially bound graph  $C_j$ . This is possible if (a)  $\rho$  deletes 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 is no transformation via  $\rho$  such that an occurrence of  $C_j$  is deleted by deleting that node, since the dangling edge condition would not be satisfied. A rule that satisfies these properties does not reduce the largest satisfied layer.

We also need to ensure that the number of violations is not increased. To do this, we have to check that  $\rho$  is not able to insert occurrences of the corresponding universally bound graph, as described above, and that  $\rho$  is not able to remove occurrences of any intermediate graph. This is only ensured if  $\rho$  does not remove any elements of  $C_{k+1} \setminus C_k$  when the set of intermediate graphs is given by  $\text{IG}(C_k, C_{k+1})$ .

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

**Lemma 4.6.** *Given a plain rule  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  and a constraint  $c$  in UANF. Let  $-1 \leq k < \text{nl}(c)$  be odd, then  $\rho$  is a basic consistency-maintaining rule up to layer  $k$  w.r.t.  $c$  if ?? and ?? hold for all  $k$ , and ?? holds if  $k < \text{nl}(c) - 2$ .*

1. *For each existentially bound graph  $C_j$  with  $2 \leq j \leq k + 1$  and each overlap  $P \in \text{ol}(L, C_j)$  with  $i_L^P(L \setminus K) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$ , the rule  $\rho$  is not applicable at match  $i_L^P$ .*
2. *For each universally bound graph  $C_j$  with  $1 \leq j \leq k + 2$  and each overlap  $P \in \text{ol}(R, C_j)$  with  $i_R^P(R \setminus K) \cap i_{C_j}^P(C_j) \neq \emptyset$ , the rule  $\rho^{-1}$  is not applicable at match  $i_r^P$ .*
3. *For all graphs  $P \in \text{ol}(L, C_{k+3})$  it holds that*

$$i_L^P(L \setminus K) \cap i_{C_{k+3}}^P(C_{k+3} \setminus C_{k+2}) = \emptyset.$$

*Proof.* Let  $\rho = L \xleftarrow{l} K \xrightarrow{r} R$  be a rule that satisfies the characterisations listed in Lemma ?? with  $-1 \leq k < \text{nl}(c)$  odd. Suppose  $\rho$  is not a direct consistency-maintaining rule up to layer  $k$  w.r.t.  $c$ . Then, there is a transformation  $t : G \Rightarrow_{\rho, m} H$  with  $k_{\max}(c, G) = k$  such that  $t$  is not direct consistency-maintaining w.r.t.  $c$ . Therefore, either the deleting, inserting, universally or existentially condition is not satisfied.

1. If the deleting condition is not satisfied, then  $k < \text{nl}(c) - 2$ . There is an occurrence  $p : C_{k+2} \hookrightarrow G$  such that  $p \models \text{IC}_0(\text{sub}_{k+2}(c), C')$  and  $\text{tr}_t \circ p \models \text{IC}_0(\text{sub}_{k_{\max}+2}(c), C')$  with  $C' \in \text{IG}(C_{k+2}, C_{k+3})$ . So an overlap  $P \in \text{ol}(L, C_{k+3})$  with  $i_L^P(L \setminus K) \cap i_{C_{k+3}}^P(C_{k+3} \setminus C_{k+2}) = \emptyset$  must exist. This is a contradiction.
2. If the inserting condition is not satisfied, there is an occurrence  $p : C_{k+2} \hookrightarrow H$  such that no morphism  $q : C_{k+2} \hookrightarrow G$  with  $p = \text{tr}_t \circ q$  exists and  $p \not\models \text{false}$  if  $\text{sub}_{k+2}(c) = \text{false}$  and  $p \not\models \text{IC}_0(\text{sub}_{k+2}(c), C_{j+3})$  otherwise. So there is an overlap  $P \in \text{ol}(R, C_{k+2})$  with  $i_R^P(R \setminus K) \cap i_{C_{k+2}}^P(C_{k+2}) \neq \emptyset$  such that  $\rho^{-1}$  is applicable at match  $i_R^P$ . This is a contradiction.
3. If the universally condition is not satisfied, there is an occurrence  $p : C_j \hookrightarrow H$  of an universally bound graph  $C_j$  with  $1 \leq j \leq k + 2$  such that no morphism  $q : C_j \hookrightarrow G$  with  $\text{tr}_t \circ q = p$  exists. So there is an overlap  $P \in \text{ol}(R, C_j)$  with  $i_R^P(R \setminus K) \cap i_{C_j}^P(C_j) \neq \emptyset$  such that  $\rho^{-1}$  is applicable at match  $i_R^P$ . This is a contradiction.
4. If the existentially condition is not satisfied, there is an occurrence  $p : C_j \hookrightarrow H$  of an existentially bound graph  $C_j$  with  $2 \leq j \leq k + 1$  such that  $\text{tr}_t \circ p$  is not total. So there is an overlap  $P \in \text{ol}(L, C_j)$  with  $i_L^P(L \setminus K) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$  such that the rule  $\rho$  is applicable at match  $i_L^P$ . This is a contradiction.

In summary,  $\rho$  is a basic consistency-maintaining rule up to layer  $k$ . □

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

This, at first sight, seems to be a restriction of the set of basic increasing rules, but the context of any rule  $\rho$  that satisfies all the properties of a basic increasing rule except that  $C_{k+2}$  is a subgraph of the left-hand side can be extended so that  $\rho$  is a basic increasing rule and the semantic of  $\rho$  is not increased. A method for deriving these rules will be presented later.

Basic increasing rules at layer  $k$  are called *deleting basic increasing rules* when  $p$  is removed and *inserting basic increasing rules* when an intermediate graph is inserted. For our repair process, we will introduce the restriction that deleting basic increasing rules may only delete edges but not nodes of  $C_{k+2}$ , since otherwise it is not possible to decide, given a rule set and a constraint, whether this rule set is able to repair an arbitrary graph based only on deleting basic increasing rules. For example, consider a rule that deletes a node from  $C_{k+2}$ . Then it is unknown whether this node is connected to nodes that do not belong to  $C_{k+2}$ , and it is unclear whether all occurrences of  $C_{k+2}$  could be destroyed by  $\rho$ , since the dangling edge condition might not be satisfied.

**Definition 4.7 (basic increasing rule).** *Given a constraint  $c$  in UANF and a direct consistency-maintaining rule  $\rho = (\text{ac}, L \xleftarrow{l} K \xrightarrow{r} R)$  up to layer  $-1 \leq k \leq \text{nl}(c) - 2$ , where  $k$  is odd. Then,  $\rho$  is a basic increasing rule w.r.t  $c$  at layer  $k$  if a morphism  $p : C_{k+2} \hookrightarrow L$ , called the increasing morphism, exists such that either ?? or ?? holds.*

1. Universally deleting:  $r \circ l^{-1} \circ p$  is not total. Then,  $\rho$  is called a deleting basic increasing rule.
2. Intermediate inserting: If  $k < \text{nl}(c) - 2$ , there is 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 4.4.** *Consider the rule `assignFeature` given in Figure ?? and constraint  $c_1$  given in Figure ??. Then, `assignFeature` is an inserting basic rule with  $C_2^2 \in \text{IG}(C_1^1, C_2^1)$  w.r.t.  $c_1$  but is not an inserting basic rule with respect to the constraint  $\forall(C_2^2, \exists(C_2^1, \text{true}))$  since the left-hand side of `assignFeature` does not contain an occurrence of  $C_2^2$ .*

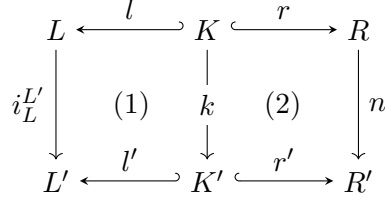


Figure 16: Pushout diagram for the construction of basic increasing rules.

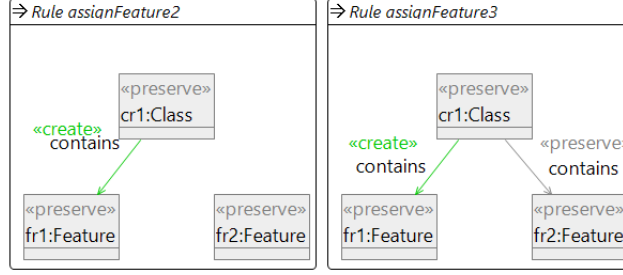


Figure 17: Derived rules of `assignFeature` and `c1`.

As mentioned above, given a direct consistency-maintaining rule  $\rho$ , we can derive basic increasing rules that are applicable when  $\rho$  is applicable by extending the context of that rule so that it contains an occurrence of the graph  $C_{k+2}$ .

**Definition 4.8 (derived rules).** *Given a constraint  $c$  in UANF and a rule  $\rho = (\text{ac}, L \xleftarrow{l} K \xrightarrow{r} R)$ . The set of derived rules from  $\rho$  at level  $0 \leq k \leq \text{nl}(c) - 2$ , where  $k$  is odd, contains rules characterised as follows: Let*

$$\mathbf{G} := \begin{cases} \{C_{k+2}\} & \text{if } k = \text{nl}(c) - 2 \text{ is existentially bound} \\ \text{IG}(C_{k+2}, C_{k+3}) & \text{otherwise.} \end{cases}$$

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

$$L' \xleftarrow{l'} K' \xrightarrow{r'} R' \text{ is universally deleting or intermediate inserting}$$

*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 4.5.** *Consider the rule `assignFeature` given in Figure ?? and constraint  $c_1$  given in Figure ?. The set of derived rules from  $\rho$  at layer 1 is given in Figure ??.*

Obviously, a rule  $\rho'$  contained in the set of rules derived from a rule  $\rho$  is only applicable to a match  $m'$  if  $\rho$  is applicable to a match  $m$  with  $m = m' \circ i$ , where  $i$  is the inclusion of the left side of  $\rho$  in the left side of  $\rho'$ . Therefore, given a set of rules  $\mathcal{R}$ , extending  $\mathcal{R}$  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 to extend a given set of rules by as many basic increasing rules as possible without extending the expressiveness of that set.

**Lemma 4.9.** *Given 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 direct maintaining rule up to layer  $-1 \leq k \leq \text{nl}(c) - 2$ , where  $k$  is odd, which is universally deleting and intermediate inserting according to Definition ???. Then every rule contained in the set of derived rules of  $\rho$  at layer  $k$  is a basic increasing rule.*

*Proof.* Let  $\rho' = (\text{ac}', L' \xleftarrow{l'} K' \xrightarrow{r'} R')$  be one of these derived rules. Since  $\rho'$  deletes and inserts exactly the same elements as  $\rho$  and  $m', \models \text{ac}' \iff m' \circ i_L^{L'} \models \text{ac}$ ,  $\rho'$  is a direct consistency maintaining rule up to layer  $k$  and is universally deleting or intermediate inserting according definition ???. It follows that  $\rho'$  is a basic increasing rule at layer  $k$ .  $\square$

In transformations via a rule  $\rho$  such that the match intersects an occurrence of a universally bound graph  $C_{k+2}$ ,  $\rho$  can be replaced by a derived rule of  $\rho$  at level  $k$ .

**Lemma 4.10.** *Given a constraint  $c$  in UANF and a rule  $\rho = (\text{ac}, L \xleftarrow{l} K \xrightarrow{r} R)$ . Then, for each transformation*

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

*such that an occurrence  $p : C_{k+2} \hookrightarrow G$  of a universally bound graph  $C_{k+2}$  with  $p(C_{k+2}) \cap m(L) \neq \emptyset$  exists, there exists a transformation*

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

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

*Proof.* Since  $p(C_{k+2}) \cap m(L) \neq \emptyset$  there is an overlap  $P \in \text{ol}(C_{k+2}, L)$  such that there exists a morphism  $q : P \hookrightarrow G$  with  $m = q \circ i_L^P$  and  $p = q \circ i_{C_{k+2}}^P$ . Since  $t$  exists, there is a derived rule  $\rho' = (\text{ac}', L' \xleftarrow{l'} K' \xrightarrow{r'} R')$  of  $\rho$  at layer  $k$ , where  $L' = P$ . We set  $m' = q$ , since  $m = m' \circ i_L^{L'} \models \text{ac}$ , it follows that  $m' = \text{ac}'$ , and since  $\rho$  removes and inserts the same elements as  $\rho'$ , there is the transformation  $t' : G \Longrightarrow_{\rho', m'} H$ .  $\square$

This allows us to replace consistency-increasing transformations via a direct consistency-maintaining rule  $\rho$  at layer  $k$  by a rule derived from  $\rho$  at layer  $k$ , i.e. a basic increasing rule at layer  $k$ .

### 4.3 Application Conditions for Basic Rules

Let us now introduce the application conditions for basic increasing rules. Since basic rules are consistency-maintaining at a certain layer  $k$  it suffices to check whether  $m \circ i \exists(C_{k+3}, \text{true})$  if  $\rho$  is a deleting rule, and whether  $m \circ i \models \exists(C', \text{true})$  if  $\rho$  is an inserting rule, where  $m$  is the match of the transformation and  $i$  is the increasing morphism of  $\rho$ .

**Definition 4.11 (application conditions for basic increasing rules).** *Given 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  $-1 \leq k \leq \text{nl}(c) - 2$ , where  $k$  is odd. The basic application condition of  $\rho$  w.r.t.  $c$  at layer  $-1 \leq j \leq \text{nl}(c) - 2$  is given by*

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

with

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

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

These application conditions are much easier to construct and smaller than those constructed by Definition ???. Note that in the case of an inserting rule  $\rho$  which inserts an intermediate graph  $C$ , the application condition only checks whether the increasing morphism does not satisfy  $\exists(C, \text{true})$ . But an application of this rule could also lead to a consistency increasing transformation w.r.t.  $c$  if the increasing morphism satisfies  $\exists(C, \text{true})$  and another intermediate graph  $C'$  is inserted. To check this, conditions similar to those constructed via Definition ??? must be constructed. At first sight this seems 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 an inserting basic increasing rule with  $C'$ , so that this rule, equipped with the corresponding basic application condition, can be used to perform this consistency-increasing transformation. For example, consider the rule **assignFeature** and constraint  $c_1$  given in Figure ??, there is 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 is also 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 ??? are indeed direct consistency increasing rules at layer.

**Theorem 4.4.** *Given 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  $-1 \leq k \leq \text{nl}(c) - 2$ , where  $k$  is odd.*

*Then,  $\rho' = (\text{ac} \wedge \text{basic}_k(\rho), L \xleftarrow{l} K \xrightarrow{r} R)$  is a direct consistency increasing rule at layer  $k$ .*

$$\begin{aligned}
\text{basic}_{-1}(\text{assignFeature}) &= \neg\exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \downarrow \text{contains} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \end{array}, \text{true} \right) \wedge \neg\exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \downarrow \text{contains} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array}, \text{true} \right) \\
\text{basic}_{-1}(\text{assignFeature3}) &= \neg\exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \\ \swarrow \text{contains} \quad \searrow \text{contains} \\ \boxed{\text{rf1} = \text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array}, \text{true} \right) \wedge \neg\exists \left( \begin{array}{c} \boxed{\text{rc1} = \text{cc1:Class}} \quad \boxed{\text{rf1} = \text{cf3:Feature}} \\ \swarrow \text{contains} \quad \searrow \text{contains} \\ \boxed{\text{cf1:Feature}} \quad \boxed{\text{cf2:Feature}} \end{array}, \text{true} \right)
\end{aligned}$$

Figure 18: Application condition for **assignFeature** and **assignFeature3** with  $c_1$  at layer 1.

*Proof.* Given a graph  $G$  with  $k_{\max} = k$ . We show that each transformation  $t : G \Rightarrow_{\rho', m} H$  is direct consistency increasing w.r.t.  $c$ . Since,  $\rho'$  is a basic increasing rule at layer  $k$ ,  $\rho'$  is also a consistency maintaining transformation at layer  $k$  and  $t$  satisfies the deletion, inserting, universally and existentially conditions. Therefore, we only need to show that  $t$  satisfies the special increasing or general increasing condition respectively.

1. If  $\rho'$  is a deleting rule,  $r \circ l^{-1} \circ i$  is not total, where  $i$  is the increasing morphism of  $\rho'$ . If  $k = \text{nl}(c) - 2$ , the transformation satisfies the special increasing condition, since one occurrence of  $C_{k+2}$  is removed. If  $k < \text{nl}(c) - 2$ , since  $m \models \text{basic}_k(\rho)$ , the morphism  $m \circ i$  does not satisfy  $\exists(C_{k+3}, \text{true})$ . Since this occurrence is destroyed,  $t$  satisfies the general increasing condition.
2. If  $\rho'$  is an inserting rule with  $C' \in \text{IG}(C_{k+2}, C_{k+3})$ , then  $k \leq \text{nl}(c) - 2$ . The morphism  $m \circ i$  does not satisfy  $\exists(C', \text{true})$ , since  $m \models \text{basic}_k(\rho)$ . But it holds that  $\text{tr}_t \circ m \circ i \models \exists(C, \text{true})$  and therefore  $t$  satisfies the general increasing condition.

In summary,  $\rho'$  is a basic direct consistency increasing rule at layer  $k$  w.r.t.  $c$ .  $\square$

**Example 4.6.** Again, consider the rule **assignFeature**, its derived rule **assignFeature3** and  $c_1$ . The application condition for these rules at layer  $-1$  w.r.t.  $c_1$  is given in Figure ??.

## 5 Rule-based Graph Repair

In the following, we present our rule-based graph repair approach. First, we propose a graph repair process for a constraint in UANF, and second, a repair process for a set of constraints in UANF, both based on a given set of rules  $\mathcal{R}$ . In addition, we need to make further assumptions for these constraints and sets of constraints, namely that they are *circular conflict free*, in order to guarantee that our approach terminates.

### 5.1 Conflicts within Conditions

During a repair process, inserting elements of an existentially bound constraint  $C_j$  could also insert new occurrences of universally bound graphs  $C_i$ . This insertion is unproblematic if  $i > k_{\max} + 2$ , but if  $i \leq k_{\max} + 2$  it could lead either to the insertion of new

violations or to a reduction of the largest satisfied layer. Additionally, removing elements of a universally bound graph  $C_j$  may destroy occurrences of an existentially bound graph  $C_i$ . Again, this can lead to the insertion of new violations or a reduction of the largest satisfied layer.

We will now introduce the notion of *conflicts within conditions*, which states that  $C_j$  has a conflict with  $C_i$  if and only if one of the cases described above can occur. Note that conflicts can only occur between existentially and universally bound graphs, and vice versa. There cannot be a conflict between two existentially bound or two universally bound graphs, since the insertion of elements cannot destroy occurrences of existentially bound graphs, and the removal of elements cannot insert new occurrences of universally bound graphs.

**Definition 5.1 (conflicts within conditions).** *Given a condition  $c$  in UANF. A existentially bound graph  $C_k$  has a conflict with an universally bound graph  $C_j$  if there is a transformation  $t : G \Rightarrow_\rho H$  with  $\rho = C_{k-1} \xleftarrow{\text{id}} C_{k-1} \xrightarrow{a_{k-1}} C_k$  such that*

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

*A universally bound graph  $C_k$  has a conflict with an existentially bound graph  $C_j$  if there is a transformation  $t : G \Rightarrow_\rho H$  with  $\rho = C_k \xleftarrow{a_{k-1}^r} C \xrightarrow{\text{id}} C$  for any  $C \in \text{IG}(C_{k-1}, C_k)$  such that*

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

Additionally, we introduce *conflicts graphs*, which represent the conflicts within a condition via a graph. With these we are able to define *transitive conflicts*, *circular conflicts* and their absence, which will be a necessary property for the termination of our repair process. Intuitively, as the name suggests, a condition  $c$  contains a circular conflict if a graph  $C_k$  has a conflict with itself or if there exists a sequence  $C_k = C_{j_1}, \dots, C_{j_n} = C_k$  of graphs such that  $C_{j_i}$  has a conflict with  $C_{j_{i+1}}$ . We can check this property by checking whether the conflict graph contains cycles.

**Definition 5.2 (conflict graph, circular conflicts).** *Let a condition  $c$  in UANF be given. The conflict graph of  $c$  is constructed in the following way. For every  $0 \leq k < \text{nl}(c)$  there is a node labelled  $k$ . If there is a conflict between  $C_k$  and  $C_j$ , there is an edge  $e$  with  $\text{src}(e) = k'$  and  $\text{tar}(e) = j'$  if either  $k = k'$  or  $k = k' + 1$ , either  $j = j'$  or  $j = j' + 1$  and  $j' \neq k'$ .*

*A graph  $C_k$  has a transitive conflict with  $C_j$  if there exists a path from  $k$  to  $j$ . A graph  $C_k$  has a circular conflict if  $C_k$  has a transitive conflict with itself. A condition  $c$  is called circular conflict free if  $c$  does not contain a circular conflict.*

In other words, a condition  $c$  is *circular conflict free* if its conflict graph is acyclic.

**Example 5.1.** *Consider constraint  $c_3$  and the transformations  $t_1$  and  $t_2$  shown in Figure ???. Transformation  $t_1$  shows that  $C_1$  has a conflict with  $C_2$  because the rule  $\rho = C_1 \xleftarrow{\text{id}}$   $C_1 \xrightarrow{a_1} C_2$  has been applied and there is a newly inserted occurrence of  $C_1$  that does not*



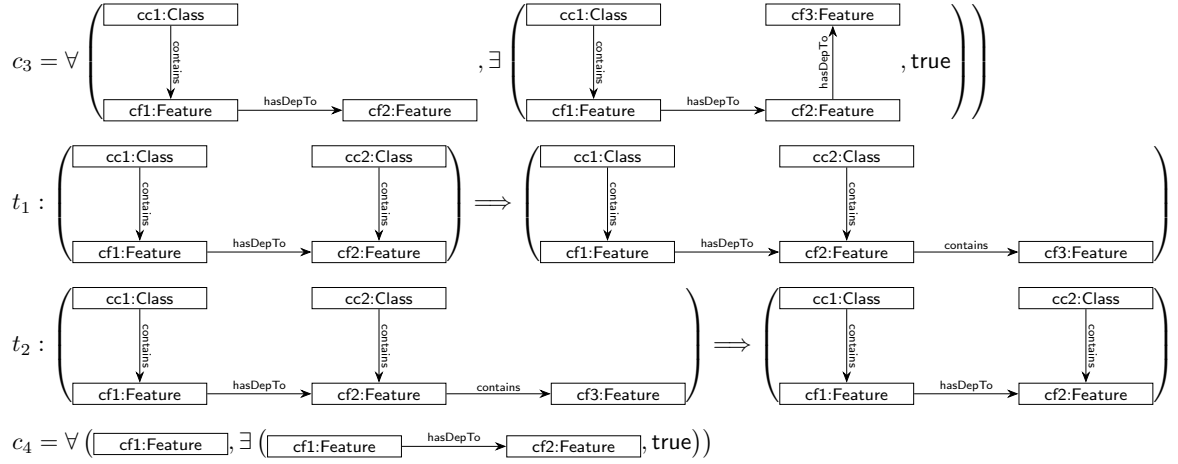


Figure 19: Constraint  $c_3$  and the transformation that show the existence of conflicts between  $C_1$  and  $C_2$  and  $C_2$  and  $C_1$ .

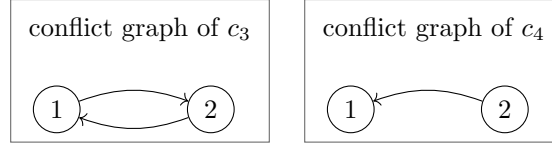


Figure 20: Conflict graphs of  $c_3$  and  $c_4$ .

satisfy  $\exists(C_2, \text{true})$ . Transformation  $t_2$  shows that  $C_2$  has a conflict with  $C_1$ , since the rule  $C_2 \xleftarrow{a_1} C_1 \xrightarrow{\text{id}} C_1$  has been applied and one occurrence of  $C_1$  has been destroyed. So  $c_3$  contains a circular conflict, the conflict graph of  $c_3$  is shown in Figure ??.

In general, the statement “ $C_j$  has a conflict with  $C_k$ ” does not imply that “ $C_k$  has a conflict with  $C_j$ ” as shown by constraint  $c_4$  given in Figure ??. The conflict graph of  $c_4$  is also shown in Figure ??. It can be seen that  $c_4$  is a circular conflict free constraint.

We will now present two characterisations of conflicts. One based on overlaps and the other based on rules. For  $C_k$ , which is existentially bound and  $C_j$ , which is universally bound, the overlap-based characterisation checks whether for each overlap of  $C_k$  and  $C_j$ , such that the overlap morphisms restricted to  $C_k \setminus C_{k+1}$  and  $C_j$  overlap, the rule that only deletes  $C_k \setminus C_{k-1}$  is applicable. If this is not possible, there does not exist a transformation as described in Definition ??. If  $C_k$  is universally bound and  $C_j$  is existentially bound, the characterisation checks whether for each overlap of  $C_k$  and  $C_j$  such that the elements of  $C_k \setminus C_{k-1}$  and  $C_j \setminus C_{j-1}$  overlap, a rule is applicable that only removes elements of  $C_k \setminus C_{k-1}$ . Again, if this is not possible, there is no transformation as described in definition ??.

**Lemma 5.3.** *Given a constraint  $c$  in UANF.*

1. Let  $C_k$  be an existentially and  $C_j$  a universally bound graph of  $c$ . Then,  $C_k$  has a

conflict with  $C_j$ , if and only there is an overlap  $P \in \text{ol}(C_k, C_j)$  with

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

and the rule  $\rho = C_k \xleftrightarrow{a_{k-1}} C_{k-1} \xrightarrow{\text{id}} C_{k-1}$  is applicable at match  $i_{C_k}^P$ .

2. Let  $C_k$  be an universally and  $C_j$  be an existentially bound graph of  $c$ . Then,  $C_k$  has a conflict with  $C_j$  if there is an overlap  $P \in \text{IG}(C_k, C_j)$  with

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

and a rule  $\rho = C_k \xleftrightarrow{a_{k-1}^r} C \xrightarrow{\text{id}} C$  with  $C \in \text{IG}(C_{k-1}, C_k)$  and  $i_{C_k}^P(C_k \setminus C) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$  is applicable at match  $i_{C_k}^P$ .

*Proof.* Given a condition  $c$  in UANF.

1. “ $\implies$ ”: Let  $C_k$  be an existentially bound graph that has a conflict with an universally bound graph  $C_j$ . Then, there is a transformation  $t : G \implies_\rho H$  with  $\rho = C_{k-1} \xleftrightarrow{\text{id}} C_{k-1} \xleftrightarrow{a_{k-1}} C_k$  such that a new occurrence  $p$  of  $C_j$  is inserted. Since only elements of  $C_k \setminus C_{k-1}$  are inserted, it holds that  $p(C_j) \cap n(C_k \setminus C_{k-1}) \neq \emptyset$ , with  $n$  being the co-match of  $t$ . The graph  $p(C_j) \cup n(C_k)$  is the overlap we are looking for, and the rule  $\rho^{-1} = C_k \xleftrightarrow{a_{k-1}} C_{k-1} \xrightarrow{\text{id}} C_{k-1}$  must be applicable at the match  $n$ .  
“ $\Leftarrow$ ”: Let  $C_k$  be an existentially and  $C_j$  an universally bound graph such that there exists an overlap  $P \in \text{ol}(C_k, C_j)$  with  $i_{C_k}^P(C_k \setminus C_{k-1}) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$  so that the rule  $\rho = C_k \xleftrightarrow{a_{k-1}} C_{k-1} \xrightarrow{\text{id}} C_{k-1}$  is applicable at match  $i_{C_k}^P$ . Then the inverse transformation of  $t : P \implies_{\rho, i_{C_k}^P} H$  is the transformation we are looking for 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, there is a transformation  $t : G \implies_\rho H$  with  $\rho = C_k \xleftrightarrow{a_{k-1}^r} C \xrightarrow{\text{id}} C$  and  $C \in \text{IG}(C_{k-1}, C_k)$  such that  $\text{tr}_t \circ p$  is no total for an occurrence  $p : C_j \hookrightarrow G$ . The graph  $p(C_j) \cup m(C_k)$  is the overlap we are looking for and  $i_{C_k}^P(C_k \setminus C) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$  must 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 there is an overlap  $P \in \text{ol}(C_k, C_j)$  with  $i_{C_k}^P(C_k \setminus C_{k-1}) \cap i_{C_j}^P(C_j \setminus C_{j-1}) \neq \emptyset$  so that a rule  $\rho = C_k \xleftrightarrow{a_{k-1}^r} C \xrightarrow{\text{id}} C$  with  $C \in \text{IG}(C_{k-1}, C_k)$  is applicable at match  $i_{C_k}^P$ . Then, the transformation of  $t : P \implies_{\rho, i_{C_k}^P} H$  is the transformation we are looking for and  $C_k$  has a conflict with  $C_j$ .

□

Our second characterisation of conflicts is based on the notion of basic maintaining rules.

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

1. *Let  $C_k$  be an existentially and  $C_j$  be an universally bound graph of  $c$ . Then,  $C_k$  has a conflict with  $C_j$  if and only if the rule  $\rho = C_{k-1} \xleftarrow{\text{id}} C_{k-1} \xrightarrow{a_{k-1}} C_k$  is not a basic consistency maintaining rule up to layer  $-1$  w.r.t.  $\forall(a_{j-1} \circ \dots \circ a_0 : C_0 \hookrightarrow C_j, \text{false})$ .*
2. *Let  $C_k$  be an universally and  $C_j$  be an existentially bound graph of  $c$ . Then,  $C_k$  has a conflict with  $C_j$  if and only if each rule  $\rho = C_k \xleftarrow{a_{k-1}^r} C \xrightarrow{\text{id}} C$  with  $C \in \text{IG}(C_{k-1}, C_k)$  is not a basic consistency maintaining rule up to layer  $-1$  w.r.t.  $\exists(a_{j-1} \circ \dots \circ a_0 : C_0 \hookrightarrow C_j, \text{true})$ .*

*Proof.* 1. Let  $C_k$  be an existentially and  $C_j$  an universally bound graph of  $c$ .

“ $\implies$ ”: Assume that  $C_k$  has a conflict with  $C_j$ . Therefore, there does exist a transformation  $t : G \implies_\rho H$  with  $\rho = C_{k-1} \xleftarrow{\text{id}} C_{k-1} \xrightarrow{a_{k-1}} C_k$  such that a new occurrence  $p : C_j \hookrightarrow H$  has been inserted. Then,  $t$  does not satisfy the universally condition and  $\rho$  is not a basic maintaining rule up to layer  $-1$ .

“ $\impliedby$ ”: Assume that  $\rho = C_{k-1} \xleftarrow{\text{id}} C_{k-1} \xrightarrow{a_{k-1}} C_k$  is not a basic maintaining rule up to layer  $-1$  w.r.t.  $\forall(a_{j-1} \circ \dots \circ a_0 : C_0 \hookrightarrow C_j, \text{false})$ . Since this constraint only contains universally bound graphs, there must exist a transformation  $t : G \implies_\rho$  that does not satisfy the universally condition. Therefore, a new occurrence of  $C_j$  has been inserted by  $t$  and with Definition ?? follows that  $C_k$  has a conflict with  $C_j$ .

2. Let  $C_k$  be an universally and  $C_j$  be an existentially bound graph of  $c$  and  $c' = \exists(a_{j-1} \circ \dots \circ a_0 : C_0 \hookrightarrow C_j, \text{true})$ .

“ $\implies$ ”: Assume that  $C_k$  has a conflict with  $C_j$ . Therefore, there is a transformation  $t : G \implies_\rho H$  with  $\rho = C_k \xleftarrow{a_{k-1}^r} C \xrightarrow{\text{id}} C$ , for a  $C \in \text{IG}(C_{k-1}, C_k)$  such that an occurrence of  $C_j$  has been destroyed. Then,  $t$  does not satisfy the existentially condition. Therefore,  $\rho$  is not a basic consistency maintaining rule w.r.t.  $c$  up to layer  $-1$ .

“ $\impliedby$ ”: Assume that  $\rho = C_k \xleftarrow{a_{k-1}^r} C \xrightarrow{\text{id}} C$  is not a basic increasing rule w.r.t.  $c$  up to layer  $-1$ . Therefore, there is transformation  $t : G \implies_\rho H$  that does not satisfy the existentially condition and an occurrence of  $C_j$  has been removed by  $t$ . It follows that  $C_k$  has a conflict with  $C_j$ . □

## 5.2 Repairing rule Sets

Given a set of rules and a constraint, it is unclear whether or not it is possible to repair a graph using the rules of that set. Therefore, we introduce the notion of *repairing rule*

sets, which is a characterisation of rule sets that are able to repair a graph w.r.t. a circular conflict free constraint. First, we introduce the notion of *repairing sequences*. A repairing sequence is a sequence of rule applications that either destroys an occurrence of a universal or inserts an occurrence of an existentially bound graph, and is applicable to each occurrence of these graphs. To ensure that these sequences are applicable to every occurrence, it is necessary to ensure that no nodes of these occurrences are removed and that the left-hand side of the first rule of the repairing sequence is contained in that occurrence. In other words, every repairing sequence of  $C_k$  starts with a transformation originating in  $C_k$  if  $C_k$  is universally bound and  $C_{k-1}$  if  $C_k$  is existentially bound.

**Definition 5.5 (repairing sequence).** *Let a constraint  $c$  in UANF and a set of rules  $\mathcal{R}$  be given.*

1. *If  $C_k$  is existentially bound, a sequence of transformations*

$$C_{k-1} = G_0 \xrightarrow{\rho_1, m_1} G_1 \xrightarrow{\rho_2, m_2} \dots \xrightarrow{\rho_n, m_n} G_n$$

*with  $\rho_i \in \mathcal{R}$  is called a repairing sequence of  $C_k$  if  $G_n \models_k c$ ,  $\text{tr}_{t_n} \circ \dots \circ \text{tr}_{t_1} \circ \text{id}_{C_{k-1}}$  is total and the concurrent rule of this sequence is a basic consistency maintaining rule w.r.t.  $\forall(C_j, \text{false})$  for all universally bound graphs  $C_j$  such that  $C_k$  has no conflict with  $C_j$ .*

2. *If  $C_k$  is universally bound, a sequence of transformations*

$$C_k = G_0 \xrightarrow{\rho_1, m_1} G_1 \xrightarrow{\rho_2, m_2} \dots \xrightarrow{\rho_n, m_n} G_n$$

*with  $\rho_i \in \mathcal{R}$  is called a repairing sequence of  $C_k$  if  $G_n \models_k c$ , for each node  $v \in V_{G_0}$  there does exist a node  $v' \in V_{G_n}$  with  $v' = \text{tr}_{t_n}(\dots \text{tr}_{t_1}(v))$  and the concurrent rule of this sequence is a basic consistency maintaining rule w.r.t.  $\forall(C_j, \text{true})$  for universally bound graphs  $C_j$ .*

In both cases the insertion of additional elements, i.e.  $G_n \neq C_{k+1}$  if  $C_k$  is existentially bound and  $G_n \neq C$  for all  $C \in \text{IG}(C_{k-1}, C_k)$  if  $C_k$  is universally bound, could lead to the insertion of universally bound graphs. For an existentially bound graph, this can happen if there is an overlap with a universally bound graph in a similar way as shown in figure ???. To ensure that this does not happen, we need the additional condition that the concurrent rule is a basic consistency maintaining rule with respect to certain constraints. If  $G_n = C_{k+1}$  if  $C_k$  is existentially or  $G_n = C$  with  $C \in \text{IG}(C_{k-1}, C_k)$  if  $C_k$  is universally bound, this condition is not needed as the following Theorem shows.

**Theorem 5.1.** *Let a constraint  $c$  in UANF and a set of rules  $\mathcal{R}$  be given.*

1. *If  $C_k$  is existentially bound and there is sequence*

$$C_{k-1} \xRightarrow{\rho_1, m_1} \dots \xRightarrow{\rho_n, m_n} C_k$$

*with  $\rho_i \in \mathcal{R}$  such that  $\text{tr}_{t_n} \circ \dots \circ \text{tr}_{t_1} \circ \text{id}_{C_{k-1}}$  is total and  $C_k \models_k c$ . Then, this is a repairing sequence for  $C_k$ .*

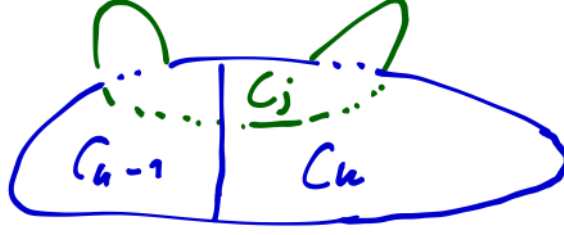


Figure 21: Scheme of an overlap of existentially bound graph  $C_k$  and universally bound graph  $C_j$  that could lead to an insertion of  $C_j$  via repairing sequences.

2. If  $C_k$  is universally bound and there is a sequence

$$C_k \Rightarrow_{\rho_1, m_1} \dots \Rightarrow_{\rho_n, m_n} C$$

with  $\rho_i \in \mathcal{R}$ ,  $C \in \text{IG}(C_{k-1}, C_k)$ ,  $C \models_k c$  and for each node  $v \in V_{G_0}$  there does exist a node  $v' \in V_{G_n}$  with  $v' = \text{tr}_{t_n}(\dots \text{tr}_{t_1}(v))$ . Then, this is a repairing sequence for  $C_k$ .

*Proof.* 1. If  $C_k$  is existentially bound, the concurrent rule is given by  $\rho = C_{k-1} \xleftarrow{\text{id}} C_k \xrightarrow{a_k} C_k$ . Let  $C_j$  be a universally bound graph such that  $C_k$  has no conflict with  $C_j$  and  $\rho$  is not a basic consistency maintaining rule w.r.t.  $\forall(C_j, \text{true})$ . With Lemma ?? follows immediately that  $C_k$  has a conflict with  $C_j$ , this is a contradiction.

2. If  $C_k$  is universally bound, the concurrent rule is given by  $\rho = C_k \xleftarrow{a_{k-1}^r} C \xrightarrow{\text{id}} C$ . Then,  $\rho$  is a basic consistency maintaining rule w.r.t.  $\forall(C_j, \text{true})$  for all universally bound graphs since  $\rho$  does not insert any elements. □

**Definition 5.6 (repairing rule set).** Let a set of rules  $\mathcal{R}$  and a circular conflict free constraint  $c$  in UANF be given. Then,  $\mathcal{R}$  is called a repairing rule set of  $c$  if there does exist a repairing sequence for each existentially bound graph of  $c$  and, if  $\text{nl}(c)$  is odd, i.e.  $c$  ends with a condition of the form  $\forall(C_{\text{nl}(c)}, \text{false})$ ,  $\mathcal{R}$  contains a repairing sequence for  $C_{\text{nl}(c)}$ .

Note that there cannot exist a repairing sequence for a universally bound graphs  $C_k$  such that  $C_k \setminus C_{k-1}$  does not contain any edges. Therefore, there is no repairing set for all constraints of the form  $\forall(C_1, \text{false})$  such that  $E_{C_1} = \emptyset$ .

**Theorem 5.2.** Let a circular conflict free constraint  $c$  in UANF and a repairing set  $\mathcal{R}$  of  $c$  be given. Then, for each graph  $G$  with  $G \not\models c$ , there is a sequence of transformations

$$G = G_0 \Rightarrow_{\rho_1, m_1} \dots \Rightarrow_{\rho_n, m_n} G_n$$

with  $\rho_i \in \mathcal{R}$  such that  $G_n \models c$ .

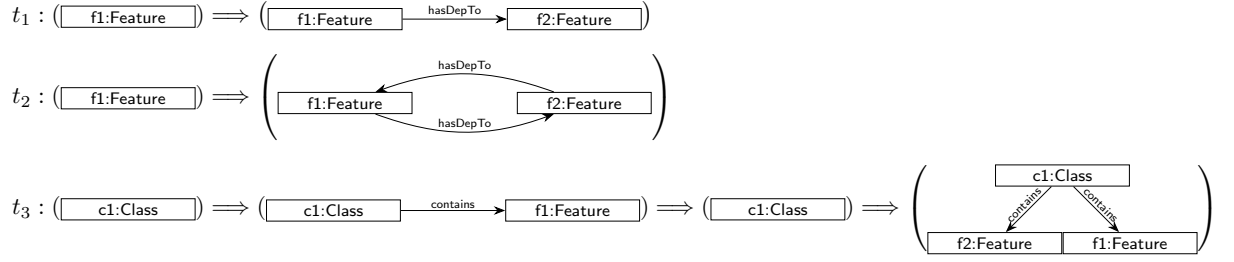


Figure 22: Repairing sequences for  $c_1$  and  $c_4$ .

We will postpone the proof of this Theorem, as it follows immediately from the termination of our repair process.

**Example 5.2.** Consider the constraints  $c_1$ ,  $c_4$  and the sequences shown in Figure ?? . The first sequence is not a repairing sequence for the existentially bound graph of  $c_4$ , since  $G_1 \not\models_1 c_4$  and therefore a rule set containing only this rule is not a repairing set w.r.t  $c_4$ . The second sequence is a repairing sequence for the existentially bound graph of  $c_4$ , since the last graph satisfies  $c_4$  and the existentially bound graph has a conflict with the universally bound graph. Therefore, the condition for the concurrent rule is also satisfied, and a rule set containing this rule is a repairing set w.r.t.  $c_4$ .

The third sequence is a repairing sequence for  $c_1$  since the last graph satisfies  $c_1$  and the sequence satisfies the criteria given in Theorem ?? . Note that this sequence consists of two applications of the same rule. A set of rules containing this rule is a repairing set w.r.t.  $c_1$ .

### 5.3 Rule-based Graph Repair for one Constraint

In the following, we present our graph repair process for a circular conflict-free constraint in UANF. The process is shown in Algorithm ?? and proceeds as follows. The algorithm starts by finding all occurrences of  $C_{k_{\max}+2}$  that do not satisfy  $\text{cut}_0(\text{sub}_{k_{\max}+2}(c))$  (line 2). This condition is equal to **false** if  $k_{\max}+2 = \text{nl}(c) - 2$  and equal to  $\exists(C_{k_{\max}+2}, \text{true})$  otherwise. If  $P$  is empty, it must follow that  $G \models_{k_{\max}+2} c$ , and so we will apply repairing sequences at occurrences contained in this set. It may be sufficient to repair only some of these occurrences. Since we do not know which of these are likely to increase the consistency, we choose one at random (line 3). For example, for existentially bound constraints  $d$ , i.e. their equivalent constraint in UANF is  $\forall(\emptyset, d)$ , there may exist occurrences of  $C_{k_{\max}+2}$  whose repair will never lead to an increase of the largest satisfied layer.

There are two ways to repair the selected occurrence, either by destroying it, or by inserting elements such that the occurrence satisfies  $\text{cut}_0(\text{sub}_{k_{\max}+2}(c))$ . The algorithm chooses one of these options (line 4) and applies the appropriate repair sequence (lines 5–11). Note that there may be no repair sequence for  $C_{k_{\max}+2}$ , since this graph is universally bound. If this is the case, we use the repairing sequence for  $C_{k_{\max}+3}$ . This must exist because  $C_{k_{\max}+3}$  is existentially bound.

If the repairing sequence for  $C_{k_{\max}+2}$  was applied, occurrences of existentially bound graphs may have been destroyed. Note that these can only be occurrences of graphs  $C_i$  such that  $C_{k_{\max}+2}$  has a conflict with  $C_i$ . This could lead to a reduction of the largest satisfied layer. Therefore the algorithm finds all these destroyed occurrences, in particular it finds all occurrences  $p$  of universally bound graphs  $C_i$  such that an occurrence  $q$  of  $C_{i+1}$  with  $p = q \circ a_j$  has been removed (line 7). If the repairing sequence for  $C_{k_{\max}+3}$  has been applied, occurrences of universally bound graphs may have been inserted. Again, these can only be occurrences of graphs  $C_i$  such that  $C_{k_{\max}+2}$  has a conflict with  $C_i$  and this could lead to a decrease of the largest satisfied layer. Again, the algorithm finds all inserted occurrences of universally bound graphs (line 10). If the largest satisfied layer has not been reduced, the algorithm chooses the next occurrence in  $P$ .

Otherwise, the largest satisfied layer must be restored. To do this, the occurrences contained in  $M$  must be repaired. Repairing these occurrences may again result in the insertion of existentially bound graphs or the removal of universally bound graphs. These occurrences are added to  $H$ , and this process is repeated until the largest satisfied layer is restored, i.e.  $H \models_{k_{\max}} c$  (line 12 – 25). The whole process is repeated until a graph satisfying  $c$  is derived.

From this it is clear why  $c$  must be circular conflict free. For a constraint with circular conflicts, during the restoring phase a new occurrence of  $C_{k_{\max}+2}$  can be inserted and an occurrence of  $C_{k_{\max}+3}$  can be removed. In certain cases this could lead to an infinite loop, so there is no guarantee that this algorithm will terminate. For example, consider the constraint  $c_3$  given in Figure ???. The set of rules used for the transformations  $t_1$  and  $t_2$  in figure ??? forms a repairing set. During a repair process using Algorithm ??, where the starting graph is the first graph of  $t_1$ , it is possible for the Algorithm ?? to enter an infinite loop by alternately applying  $t_1$  and  $t_2$ .

Optimisation of the repair algorithm in terms of the number of elements inserted or deleted can be achieved by using partial repairing sequences where possible. For example, consider the repairing sequence

$$C_k \Longrightarrow C_1 \Longrightarrow \dots \Longrightarrow C_{k+1}$$

with  $C_1 \in \text{IG}(C_k, C_{k+1})$ . For an occurrence  $p$  of  $C_k$ , which already satisfies the condition  $\exists(C_1, \text{true})$ , it may be sufficient to apply only the sequence

$$C_1 \Longrightarrow \dots \Longrightarrow C_{k+1}$$

at  $p$ . But then you have to check that no occurrences of existentially bound graphs have been destroyed, and that no occurrences of universally bound graphs  $C_i$  such that  $C_k$  has no conflict with  $C_i$  have been inserted. If this is the case, the transformations must be reversed and another (partial) repairing sequence must be used. Even though this would lead to an optimisation in terms of the number of elements inserted and deleted, it would lead to an increase in runtime due to the reversal of transformations.

For any circular conflict-free constraint, Algorithm ?? will always terminate according to the following Theorem.

---

**Algorithm 1:** Repair for one circular conflict free constraint

---

**Data:** A graph  $G$ , a circular conflict free constraint  $c$  in UANF and a repairing set  $\mathcal{R}$  for  $c$ .

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

```
1 while  $G \not\models c$  do
2    $P \leftarrow \{q : C_{k_{\max}+2} \hookrightarrow H \mid q \not\models \text{cut}_0(\text{sub}_{k_{\max}+2}(c))\}$ ;
3   Choose  $p \in P$  uniformly at random ;
4   Choose  $r \in \{0, 1\}$  uniformly at random;
5   if  $r = 0$  and  $\mathcal{R}$  contains a repairing sequence for  $C_{k_{\max}+2}$  then
6     Apply the repairing sequence for  $C_{k_{\max}+2}$  at match  $p$  and let  $H$  be the
        derived graph ;
7      $M \leftarrow \{q : C_j \hookrightarrow H \mid j \text{ odd and } \neg \exists q' : C_j \hookrightarrow G(\text{tr} \circ q' = q)\}$ ;
8   else
9     Apply the repairing sequence for  $C_{k_{\max}+3}$  at match  $p$  and let  $H$  be the
        derived graph ;
10     $M \leftarrow \{q : C_j \hookrightarrow H \mid j \text{ odd and } \exists q' : C_{j+1} \hookrightarrow G(q =$ 
         $q' \circ a_j \wedge \text{tr} \circ q' \text{ is not total})\}$ ;
11  end
12  while  $H \not\models_{k_{\max}(c,G)} c$  do
13    Choose  $p : C_j \hookrightarrow H \in M$  uniformly at random ;
14    Choose  $r \in \{0, 1\}$  uniformly at random ;
15    if  $r = 0$  and  $\mathcal{R}$  contains a repairing sequence for  $C_j$  then
16      Apply the repairing sequence for  $C_j$  at match  $p$  and let  $H'$  be the
        derived graph ;
17       $M' \leftarrow \{q : C_i \hookrightarrow H' \mid i \text{ odd and } \neg \exists q' : C_i \hookrightarrow H(\text{tr} \circ q' = q)\}$  ;
18    else
19      Apply the repairing sequence for  $C_{j+1}$  at match  $p$  and let  $H'$  be the
        derived graph ;
20       $M' \leftarrow \{q : C_i \hookrightarrow H' \mid i \text{ odd and } \exists q' : C_{i+1} \hookrightarrow G(q =$ 
         $q' \circ a_j \wedge \text{tr} \circ q' \text{ is not total})\}$ ;
21    end
22     $M \leftarrow (M \setminus \{p\}) \cup M'$  ;
23     $H \leftarrow H'$ ;
24  end
25   $G \leftarrow H$ ;
26 end
27 return  $G$ ;
```

---



**Theorem 5.3.** *Given  $n$  graph  $G$ , a circular conflict free condition  $c$  in UANF and a repairing set  $\mathcal{R}$  of  $c$ . Then, Algorithm ?? with input  $G, c$  and  $\mathcal{R}$  terminates and returns a graph  $H$  with  $H \models c$ .*

*Proof.* If Algorithm ?? terminates, it returns a graph that satisfies  $c$ . Therefore, it is sufficient to show that Algorithm ?? terminates. Since  $G$  is finite, the set  $P$  must also be finite. If a repairing sequence has been applied, the set  $M$  contains only occurrences of graphs  $C_j$  such that  $C_{k_{\max}+2}$  has a (transitive) conflict with  $C_j$ , since the repairing sequence is not able to destroy or insert occurrences of  $C_i$  such that  $C_{k_{\max}+2}$  has no (transitive) conflict with  $C_i$ . Since  $G$  is finite,  $|M|$  must also be finite.

If the derived graph does not satisfy  $\text{cut}_{k_{\max}(c,G)}(c)$ , we need to restore the largest satisfied layer. Since the largest satisfied layer only decreases if an occurrence of an existentially bound graph is destroyed or an occurrence of universally bound graphs is inserted, and  $M$  contains all these occurrences, we only need to consider the occurrences contained in  $M$ . Applying repairing sequences to occurrences  $p : C_j \hookrightarrow H \in M$  could again lead to the insertion of universally bound graphs or the removal of existentially bound graphs. The set  $M'$  contains all these occurrences, and again these are only occurrences of  $C_i$  such that  $C_j$  has a (transitive) conflict with  $C_i$ . Since  $c$  is free of circular conflict,  $M'$  cannot contain any occurrences of  $C_{k_{\max}+2}$ , otherwise  $C_j$  would have a (transitive) conflict with  $C_{k_{\max}+2}$  and therefore  $C_{k_{\max}+2}$  has a circular conflict. Therefore no occurrences of  $C_{k_{\max}+3}$  are destroyed and no occurrences of  $C_{k_{\max}+2}$  are inserted. In addition,  $C_{k_{\max}+2}$  has a (transitive) conflict with  $C_i$ , and repairing any  $p$  in  $M'$  will not lead to insertion of an occurrence of  $C_{k_{\max}+2}$  or removal of an occurrence of  $C_{k_{\max}+3}$ .

Since  $c$  is circular conflict free, there must exist graphs  $C_i$ , such that  $C_i$  has no conflict with any other graph  $C_{i'}$  and  $C_{k_{\max}+2}$  has a (transitive) conflict with  $C_i$ . Therefore, the application of repairing sequences at occurrences of these graphs will not lead to the insertion or removal of any universally or existentially bound graph, respectively. Since  $c$  is finite, the number of graphs  $C_i$  such that  $C_{k_{\max}+2}$  has a (transitive) conflict with  $C_i$  is finite. Since  $|M'|$  is also finite, after a finite number of applications of repairing sequences,  $M'$  contains only occurrences of graphs that have no conflicts. After a repairing sequence has been applied to all these occurrences,  $M'$  is empty and  $H \models_{k_{\max}(c,G)} c$ , since all occurrences  $p$  of  $C_j$  which have either been inserted or an occurrence  $q$  of  $C_{j+1}$  with  $p = a_j \circ q$  has been removed satisfy  $\exists(C_{j+1}, \text{true})$ .

Therefore, after a finite number of iterations, the set  $P$  is empty and the largest satisfied layer has been increased by at least 1. It follows that after a finite number of iterations  $G \models c$ . Then Algorithm ?? terminates and returns  $G$ .  $\square$

**Example 5.3.** *Consider constraint  $c = \forall(C_2^2, \exists(C_2^1, \text{true}))$  which is composed of the graphs shown in Figure ???. This constraint is circular conflict free and a repairing set for  $c$  is given in Figure ???. There does exist a repairing sequence for  $C_2^2$  via the rule **remove** and a repairing sequence for  $C_2^1$  via the rule **insert**. Using the rule set  $\{\text{remove}, \text{insert}\}$ , Algorithm ?? could return one of the graphs  $G_1, G_2$  or  $G_3$  given in Figure ??, depending on the repairing sequences used.*

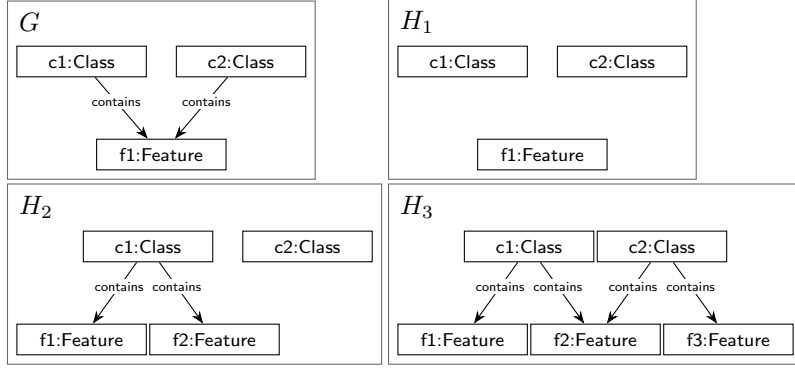


Figure 23: Possible outputs of the repairing process for  $G$  and  $\forall(C_2^2, \exists(C_2^1, \text{true}))$  using the rule set  $\{\text{removeFeature}, \text{createFeature}\}$ .

If  $\mathcal{R}$  is not a repairing set for a constraint  $c$  in UANF, it is unclear whether a graph  $G$  can be repaired with this set of rules. The only approach is a brute-force approach using consistency-maintaining or consistency-increasing rules up to layer  $k_{\max}$ . This can be done by equipping the rules of  $\mathcal{R}$  with maintaining, increasing or basic application conditions. In this case, we can repair  $G$  by brute-force if there is a sequence

$$G \Rightarrow_{\rho_1} \dots \Rightarrow_{\rho_n} H$$

of transformations with  $\rho_1, \dots, \rho_n \in \mathcal{R}$  such that each transformation is consistency-maintaining or consistency-increasing w.r.t.  $c$  and  $H \models c$ . If no such sequence exists, this approach cannot repair  $G$ , even if  $G$  can be repaired with rules of  $\mathcal{R}$ .

#### 5.4 Rule-based Graph Repair for multiple Constraints

We will now present our rule-based repair approach for a set of constraints in UANF.

**Definition 5.7 (satisfaction of constraint sets).** *Let  $\mathcal{C}$  be a set of constraints. A graph  $G$  satisfies  $\mathcal{C}$ , denoted by  $G \models \mathcal{C}$ , if  $G \models \bigwedge_{c \in \mathcal{C}} c$ . The set  $\mathcal{C}$  is called satisfiable if there exists a graph  $G$  with  $G \models \mathcal{C}$ .*

To guarantee that a set of constraints can be repaired by a set of rules, we need to extend the notion of repairing sets such that a set of rules is called a *repairing set* for a set of constraints if it is a repairing set for every constraint in the constraint set.

**Definition 5.8 (repairing set for a set of constraints).** *Given a set of constraints  $\mathcal{C}$  and a set of rules  $\mathcal{R}$ . Then  $\mathcal{R}$  is called a repairing set for  $\mathcal{C}$  if  $\mathcal{R}$  is a repairing set for all constraints  $c \in \mathcal{C}$ .*

We also extend the notion of conflicts to *conflicts between constraints*. Intuitively, a constraint  $c$  has a conflict with another constraint  $c'$  if one of its graphs has a conflict with a graph of  $c'$ .

---

**Algorithm 2:** Repair for a circular constraint-conflict free set of constraints

---

**Data:** A graph  $G$ , circular constraint-conflict free set of constraints  $\mathcal{C}$  and a repairing set  $\mathcal{R}$  for  $\mathcal{C}$ .

**Result:** A graph  $H$  with  $H \models \bigwedge_{c \in \mathcal{C}} c$ .

```
1  $(c_1, \dots, c_n) \leftarrow$  topological ordering of  $\mathcal{C}$  ;  
2 for  $i \leftarrow 1$  to  $n$  do  
3   | Repair  $c_i$  in  $G$  with Algorithm ??, let  $H$  be the retuned graph ;  
4   |  $G \leftarrow H$  ;  
5 end  
6 return  $G$ ;
```

---

**Definition 5.9 (conflict between constraints).** Let the constraints  $c, c'$  in UANF and a set of rules  $\mathcal{R}$  be given. Then  $c$  has a conflict with  $c'$  w.r.t.  $\mathcal{R}$  if a repairing sequence

$$C_k = G_0 \Rightarrow_{\rho_1, m_1} \dots \Rightarrow_{\rho_n, m_n} G_n$$

exists for a graph  $C_k$  of  $c$  such that the concurrent rule of that sequence is not a basic consistency maintaining rule w.r.t.  $\forall(C_j, \text{false})$  or  $\exists(C_j, \text{true})$  for any universally or existentially bounded graph  $C_j$  of  $c'$ .

The following lemma is a useful statement for proving the correctness of our repair approach. It states that applying a repairing sequence to a constraint  $c$  cannot destroy the satisfaction of  $c'$  if  $c$  has no conflict with  $c'$ .

**Lemma 5.10.** Let two constraints  $c$  and  $c'$  in UANF be given such that  $c$  has no conflict with  $c'$  w.r.t. to a set of rules  $\mathcal{R}$ . Then the concurrent rule  $\rho$  of any repairing sequence for  $c$  is a  $c'$ -preserving rule.

*Proof.* Suppose  $\rho$  is not a  $c'$ -preserving rule. Then there exists a transformation  $t : G \Rightarrow_{\rho, m} H$  such that  $G \models c'$  and  $H \not\models c'$ . Therefore, either a universally bound graph of  $c'$  has been inserted or an existentially bound graph of  $c'$  has been removed. It follows that  $\rho$  is not a basic maintaining rule w.r.t.  $\forall(C_j, \text{false})$  for all universally bound graphs  $C_j$  of  $c'$  or  $\rho$  is not a basic maintaining rule w.r.t.  $\exists(C_j, \text{true})$  for all existentially bound graphs  $C_j$  of  $c'$ , which is a contradiction.  $\square$

The *conflict graph* for a set of constraints and *circular conflicts* for it are defined in a similar way to the conflict graph and circular conflicts for a constraint. A set of constraints is called *circular conflict free* if each of its constraints is circular conflict free and there is no sequence  $c = c_0, \dots, c_n = c$  such that  $c_i$  has a conflict with  $c_{i+1}$  for all  $0 \leq i < n$ . In other words, the conflict graph of this set is acyclic.

**Definition 5.11 (conflict graphs, circular conflicts).** Given a set of constraints  $\mathcal{C}$  in UANF. The conflict graph of  $\mathcal{C}$  is constructed in the following way. For each constraint  $c \in \mathcal{C}$  there is a node. If there is a conflict between  $c$  and  $c'$ , there exists an edge  $e$  with  $\text{src}(e) = c$  and  $\text{tar}(e) = c'$ .

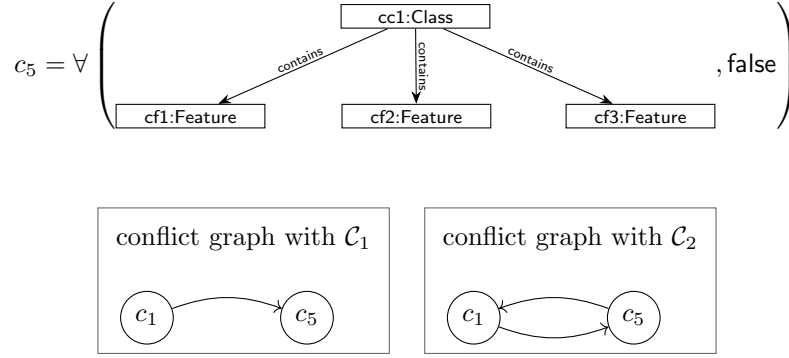


Figure 24: Constraints  $c_5$  and conflicts graphs of the constraint set  $\{c_1, c_5\}$  with the rule sets  $\mathcal{C}_1 = \{\text{resolve}, \text{createFeatures}\}$  and  $\mathcal{C}_2 = \{\text{resolve2}, \text{createFeatures}\}$ .

A constraint  $c$  has a transitive conflict with  $c'$  if the conflict graph of  $\mathcal{C}$  contains a path from  $c$  to  $c'$ . A constraint  $c$  has a circular conflict if  $c$  has a transitive conflict with itself. A set of constraints  $\mathcal{C}$  is called circular conflict free if every constraint in  $\mathcal{C}$  is circular conflict free and  $\mathcal{C}$  contains no circular conflicts.

**Example 5.4.** Consider the rules `resolve`, `resolve2`, `createFeatures` and constraints  $c_1$  and  $c_5$  given in figures ?? and ??. The constraint set  $\mathcal{C} = \{c_1, c_5\}$  is a multiplicity which states that “Each node of type **Class** is connected to exactly two nodes of type **Feature**”. With the rule set  $\mathcal{R}_1 = \{\text{resolve}, \text{createFeatures}\}$ , there is only one conflict in  $\mathcal{C}$ ;  $c_1$  has a conflict with  $c_5$ , since applying `createFeatures` could lead to inserting the universally bound graph of  $c_5$ . With the rule set  $\mathcal{R}_2 = \{\text{resolve2}, \text{createFeatures}\}$  there are two conflicts. Again, there is a conflict of  $c_1$  with  $c_5$ , but also a conflict of  $c_5$  with  $c_1$ , since applying `resolve` can destroy an occurrence of the existentially bound graph of  $c_1$ .

Therefore, our approach can repair with the rule set  $\mathcal{R}_1$  but not with  $\mathcal{R}_2$  because in this case,  $\mathcal{C}$  is not circular conflict free.

Our repair process exploits the fact that the conflict graph of a circular conflict-free set of constraints in UANF is acyclic. In particular, our approach uses the *topological ordering* of this conflict graph.

**Definition 5.12 (topological ordering of a graph).** Given is a graph  $G$ . A sequence  $(v_1, \dots, v_n)$  of nodes of  $G$  is called a topological ordering of  $G$  if no edge  $e \in E_G$  exists with  $\text{src}(e) = v_i$ ,  $\text{tar}(e) = v_j$  and  $i \geq j$ . The topological ordering of a circular conflict-free set of constraints  $\mathcal{C}$  is the topological order of its conflict graph.

It is well known that every directed acyclic graph has a topological ordering, and therefore every conflict graph of a circular conflict-free set of constraints also has a topological ordering.

The repair process is given in Algorithm ?? and proceeds as follows. First, the topological ordering of the constraint set is determined (line 1). Then Algorithm ?? is

used to repair each constraint of  $\mathcal{C}$  in the order of the topological ordering (lines 2 – 4). This ensures that the satisfaction of a constraint that has already been repaired is not destroyed by the repair of another constraint.

**Theorem 5.4.** *Given a graph  $G$ , a satisfiable, circular conflict-free set of constraints in UANF,  $\mathcal{C}$ , and a set of rules  $\mathcal{R}$ . Then Algorithm ?? terminates and returns a graph  $H$  with  $H \models \mathcal{C}$ .*

*Proof.* Since  $\mathcal{C}$  is finite and every  $c \in \mathcal{C}$  is circular conflict free, Algorithm ?? terminates for every  $c \in \mathcal{C}$ . Therefore Algorithm ?? will also terminate. It remains to show that the returned graph satisfies  $\mathcal{C}$ . Let  $(c_1, \dots, c_n)$  be a topological ordering of  $\mathcal{C}$ . Then no constraint  $c_j$  with  $j \neq 1$  has a conflict with  $c_1$ , and by Lemma ?? it follows that the concurrent rule of every repair sequence for every  $c_i$  with  $2 \leq i \leq n$  is a  $c_1$ -preserving rule. In general, the concurrent rule of each repairing sequence for  $c_j$  is a  $c_i$ -preserving rule if  $i < j$ . Note that in Algorithm ?? each repairing sequence can be replaced by its concurrent rule. After one iteration it holds that  $G \models c_1$ . Suppose that after  $m$  iterations it holds that  $G \models c_i$  for all  $1 \leq i \leq m$ . In iteration  $m + 1$ ,  $c_{m+1}$  is repaired by Algorithm ?. Since each concurrent rule of each repairing sequence of  $c_{m+1}$  is a  $c_i$ -preserving rule for all  $1 \leq i \leq m$ , the application of repairing sequence can be replaced by its concurrent rule, and Algorithm ?? applies only these concurrent rules, it follows that  $H \models c_i$  for all  $1 \leq i \leq m + 1$ . Therefore, after  $n$  iterations,  $H \models c_i$  for all  $1 \leq i \leq n$  and the returned graph  $G$  satisfies  $\mathcal{C}$ .  $\square$

## 6 Related Work

In this section, we summarise other concepts for rule-based graph repair.

### Iterative Development of Consistency-Preserving Rule-Based Refactorings:

Becker et al. [?] introduced an interactive approach to construct consistency-preserving transformation based on their invariant checker introduced in [?] for so-called *well-formedness constraints*. These are constraints of the form  $\neg\exists(c_1, \text{true})$  or  $\neg\exists(C_1, \neg\exists C_2, \text{true})$ . Given a consistent graph, a well-formedness constraint and a refactoring specification, which is a set of rules in the single-pushout approach [?], the invariant checker constructs all minimal counterexamples that lead to a non-consistency preserving transformation via rules of the refactoring specification. If there are no such counterexamples, then any transformation is consistency-preserving. This approach is designed to be fully interactive, requiring the user to revise the refactoring specification until no counterexamples are returned.

### Ensuring Consistency of Conditional Graph Grammars:

Heckel and Wagner [?] have presented an approach to construct consistency-preserving application conditions for rules in the single-pushout approach and a constraint of the form  $\forall(C_1, \exists(C'_1, \text{true})) \wedge \dots \wedge \forall(C_n, \exists(C'_n, \text{true}))$ . Although the constructed application conditions are not presented as nested-conditions, they can be transformed into nested conditions of the form

$$\forall(C_1, \exists(C_1^1, \text{true}) \vee \dots \vee \exists(C_1^{k_1}, \text{true})) \wedge \dots \wedge \forall(C_n, \exists(C_n^1, \text{true}) \vee \dots \vee \exists(C_n^{k_n}, \text{true})).$$

**Sustaining and Improving graduated Graph Consistency:** Kosiol et al. [?] have introduced the notions of (direct) consistency-sustaining and (direct) consistency-improving transformations as already introduced in section ?? . This approach is designed for rules in the double-pushout approach and nested-conditions in ANF. They have introduced a method to construct consistency-sustaining application conditions, a sufficient criterion for consistency-sustaining transformations and a necessary criterion for consistency-improving transformations, which have been implemented and evaluated.

**Constructing optimized constraint-preserving application conditions for model transformation rules:** Nassar et al. [?] have introduced a method to construct consistency-sustaining and consistency-preserving application conditions in the framework of  $\mathcal{M}$ -adhesive categories. Due to some optimisations, these application conditions are less restrictive and less complex than those described in [?] and [?]. They have introduced the notion of *weakest application conditions*. As the name suggests, a weakest application condition is implied by any other application condition with the same property. For example, a weakest consistency-preserving application condition is implied by every other consistency-preserving application condition. The construction of the application conditions has been implemented as an eclipse plug-in called *OCL2AC*, which is able to construct consistency-guaranteeing, weakest consistency-preserving or consistency-sustaining application conditions.

**Rule-based Graph Repair:** Sandmann and Habel [?] have introduced a repair process for so-called *proper constraints* based on so-called *repair programs*. A constraint in ANF is called *proper* if it ends with  $\exists(C, \text{true})$  or is of the form  $\exists(C_1, \forall(C_2, \text{false}))$  or  $\forall(C_1, \text{false})$ . The authors describe a method for inductively constructing a repair program consisting of rules in the double-pushout approach that, when applied to a non-consistent graph, returns a consistent graph. They also introduce graph repair given a set of rules  $\mathcal{R}$ . The approach can be used to repair a graph with rules from  $\mathcal{R}$  if there is a repair program such that for every rule in the repair program there is an equivalent rule in  $\mathcal{R}$ .

**Rule-based Repair of EMF Models:** Nassar et al. [?, ?] introduced a repair approach for models of the *eclipse modeling framework* (EMF) [?]. In particular, this approach is able to repair multiplicities of a given EMF metamodel. Multiplicities can be described as nested conditions of the form  $\forall(C_1, \exists(C_2, \text{true}))$  and  $\forall(C_1, \text{false})$ . The approach was implemented using two Eclipse plugins in Henshin [?]. One plugin derives rules for the repair process from a metamodel. The other plugin is an implementation of the repair process.

## 7 Conclusion

In summary, we have introduced a new concept of consistency, which leads to the new notions of the notions of consistency-maintaining and consistency-increasing transformations and rules, which is finer-grained, compared to the already existing notions of consistency preserving, consistency guaranteeing, consistency sustaining and consistency improving. Finer-grained, in the sense that the smallest increases and decreases of consistency can be detected, namely the insertion or removal of one graph element. We have compared our notions with the already existing notions mentioned above. The notions of consistency-maintaining and consistency-improving are related to the notions of consistency-preserving and consistency-guaranteeing, and are generally not related to the notions of consistency-sustaining and consistency-improving. Only in special cases, for specific constraints, are these notions related or even identical.

Furthermore, we have introduced application conditions such that a rule equipped with this condition is consistency-maintaining or consistency-increasing. In particular, we introduce two types of application conditions. First, application conditions for general rules, and second, application conditions for a special set of rules, which we call basic-maintaining and basic-increasing rules, respectively, where the application conditions for basic rules are less complex and smaller in size compared to the general application conditions.

We have introduced a rule-based graph repair process based on our new notions of consistency-maintaining and consistency-increasing for a particular type of constraints in ANF, which we call circular conflict-free constraints, using a given rule set. We present characterisations of these circular conflict free constraints and a characterisation of a rule set that is able to repair a constraint using our repair process. In addition, we extend the notion of conflicts in constraints to the notion of conflicts between constraints, and present a rule-based graph repair process for a satisfiable circular conflict free set of constraints, which is based on the repair process for one constraint.

Future work is to extend the notions of consistency-maintaining and consistency-increasing transformations for all types of nested constraints, and a rule-based repair process for all satisfiable nested constraints. Although we have presented characterisations for circular conflict-free constraints, it remains unclear for which practical applications our approach is suitable. This may require implementation and further evaluation of the repair process, characterisations and construction of application conditions.

## References

- [1] T. Arendt, E. Biermann, S. Jurack, C. Krause, and G. Taentzer. [Henshin: advanced concepts and tools for in-place EMF model transformations](#). In *International Conference on Model Driven Engineering Languages and Systems*, pages 121–135. Springer, 2010.
- [2] B. Becker, D. Beyer, H. Giese, F. Klein, and D. Schilling. Symbolic invariant verification for systems with dynamic structural adaptation. In *Proceedings of the 28th international conference on Software engineering*, pages 72–81, 2006.
- [3] B. Becker, L. Lambers, J. Dyck, S. Birth, and H. Giese. [Iterative development of consistency-preserving rule-based refactorings](#). In *International Conference on Theory and Practice of Model Transformations*, pages 123–137. Springer, 2011.
- [4] H. Ehrig, K. Ehrig, U. Prange, and G. Taentzer. [Fundamentals of algebraic graph transformation](#). *Monographs in theoretical computer science. An EATCS series*. Springer, 2006.
- [5] H. Ehrig, A. Habel, and L. Lambers. Parallelism and concurrency theorems for rules with nested application conditions. *Electronic Communications of the EASST*, 26, 2010.
- [6] A. Habel and K.-H. Pennemann. Nested constraints and application conditions for high-level structures. In *Formal methods in software and systems modeling*, pages 293–308. Springer, 2005.
- [7] 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.
- [8] R. Heckel and A. Wagner. [Ensuring consistency of conditional graph grammars—a constructive approach](#). *Electronic Notes in Theoretical Computer Science*, 2:118–126, 1995.
- [9] J. Kosiol, D. Strüder, G. Taentzer, and S. Zschaler. [Sustaining and improving graduated graph consistency: A static analysis of graph transformations](#). *Science of Computer Programming*, 214:102729, 2022.
- [10] N. Nassar, J. Kosiol, T. Arendt, and G. Taentzer. [Constructing optimized constraint-preserving application conditions for model transformation rules](#). *Journal of Logical and Algebraic Methods in Programming*, 114:100564, 2020.
- [11] N. Nassar, J. Kosiol, and H. Radke. [Rule-based repair of EMF models: formalization and correctness proof](#). In *Electronic Pre-Proc. Intl. Workshop on Graph Computation Models*, 2017.



- [12] N. Nassar, H. Radke, and T. Arendt. [Rule-based repair of EMF models: An automated interactive approach](#). In *International Conference on Theory and Practice of Model Transformations*, pages 171–181. Springer, 2017.
- [13] D. Plump. Confluence of graph transformation revisited. In *Processes, Terms and Cycles: Steps on the Road to Infinity*, pages 280–308. Springer, 2005.
- [14] C. Sandmann and A. Habel. [Rule-based graph repair](#). *arXiv preprint arXiv:1912.09610*, 2019.
- [15] D. Steinberg, F. Budinsky, E. Merks, and M. Paternostro. [EMF: eclipse modeling framework](#). Pearson Education, 2008.