

# Masterarbeit

Alexander Lauer

September 12, 2022

## **Abstract**

# Contents

<b>1 Preliminaries</b>	<b>3</b>
<b>2 partial consistency improving</b>	<b>4</b>
2.1 conditions up to layer . . . . .	4
2.2 minimal consistency improving . . . . .	6
<b>3 application condition</b>	<b>7</b>

# 1 Preliminaries

**Definition 1.1 (subgraph).** Let  $G_1$  and  $G_2$  be graphs. The graph  $G_2$  is called a subgraph of  $G_1$  if an injective morphism  $f : G_2 \rightarrow G_1$  exists. We use the notation  $G_2 \subseteq G_1$  if  $G_2$  is a subgraph of  $G_1$  and  $G_2 \subset G_1$  if  $f$  is not bijective. The set of all subgraphs of  $G_1$  is denoted by  $\text{sub}(G_1)$ .

**Definition 1.2 (overlap).** Let  $G$  and  $G'$  be graphs. A graph  $H$  is called an overlap of  $G$  and  $G'$  if morphisms  $p : G \hookrightarrow H$  and  $p' : G' \hookrightarrow H$  such that  $p$  and  $p'$  are jointly surjective. The set of all overlaps of  $G$  and  $G'$  is denoted by  $\text{ol}(G, G')$ .

**Definition 1.3 (overlap at morphism).** Let  $C, G$  and  $C'$  with  $C \subset C'$  be graphs and  $p : C \hookrightarrow G$  a morphism. A graph  $H$  is called an overlap of  $G$  and  $C'$  at  $p$  if a morphism  $p' : C' \hookrightarrow H$  with  $p'|_C = p$  exists. The set of all overlaps of  $G$  and  $C'$  at  $p$  is denoted by  $\text{ol}_p(G, C')$ .

**Definition 1.4 (partial morphism).** Let  $f : G_1 \rightarrow G_2$  and  $g : G_3 \rightarrow G_4$  be morphisms. The morphism  $g$  is called a partial morphism of  $f$  if  $G_3 \subseteq G_1$ ,  $G_4 \subseteq G_2$  and  $f|_{G_3} = g$ .

**Definition 1.5 (nested graph condition).** A graph condition over a graph  $C_0$  is inductively defined as follows:

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

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

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

**Definition 1.6 (alternating quantifier normal form (ANF)[1]).** A graph condition  $c$  is in alternating normal form (ANF) if it is of the form

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

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

## 2 partial consistency improving

### 2.1 conditions up to layer

**Definition 2.1 (Layer of a subcondition).** Let  $c$  be a condition and  $d$  a subcondition of  $c$ . The layer of  $d$  is defined as  $\text{lay}(d) := \text{nl}(c) - \text{nl}(d) - 1$ .

**Definition 2.2 (substitution at layer).** Let  $c$  be a condition, such that the subcondition of  $c$  with layer  $0 \leq k \leq \text{nl}(c)$  is an condition over  $C_k$ . Let  $e$  be a condition over  $C_k$ . The substitution of  $c = Q(a : C_0 \hookrightarrow C_1, d)$  at layer  $k$  with  $e$ ,  $\text{sub}(k, c, e)$ , is recursively defined as:

1. If  $k = 0$ :

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

2. If  $k > 0$ :

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

**Definition 2.3 (Condition up to layer).** Let  $c$  be a condition and  $d$  be the subcondition of  $c$  at layer  $0 \leq k \leq \text{nl}(c)$ . The condition up to layer  $k$  of  $c$ ,  $\text{cond}(k, c)$  is defined as

$$\text{cond}(k, c) := \begin{cases} \text{sub}(k, c, \text{true}) & , \text{if } k = 0 \vee d \text{ is existentially bound} \\ \text{sub}(k, c, \text{false}) & , \text{if } d \text{ is universally bound.} \end{cases}$$

**Definition 2.4 (Satisfaction up to layer).** Let  $G$  be a graph and  $c$  be a condition over  $C_0$ . A morphism  $p : C_0 \hookrightarrow G$  satisfies  $c$  up to layer  $k$ ,  $p \models_k c$ , if

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

A graph  $G$  satisfies a constraint  $c$  up to layer  $k$ ,  $G \models_k c$ , if  $q : \emptyset \hookrightarrow G$  satisfies  $\text{cond}(k, c)$ .

**Lemma 2.5.** Let  $G$  be a graph  $p : C_0 \hookrightarrow G$  a morphism and  $c$  a condition over  $C_0$  in ANF with  $p \models_k c$ . If the subcondition  $d = Q(a_k : C_{k-1} \hookrightarrow C_k, e)$  of  $c$  at layer  $k$  is universally bound, then for any condition  $f$  over  $C_k$  holds:

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

*Proof.* Let  $k$  be the smallest number such that  $p \models_k c$  and the subcondition of  $c$  with layer  $k$  is universally bound, let  $d = \forall(a_k : C_{k-1} \hookrightarrow C_k, e)$  be this subcondition. Let  $q : G_{k-1} \rightarrow G$  be a morphism such that  $q \models \forall(a_k : C_{k-1} \hookrightarrow C_k, \text{false})$ . This must exist, since  $p \models_k c$  and  $k$  is the smallest number such that  $p \models_k v$  and the subcondition of  $c$  with layer  $k$  is universally bound.

Therefore, there does not exist a morphism  $q' : C_k \rightarrow G$  with  $q = q' \circ a_k$ . Hence, for every condition  $f$  over  $C_k$  a morphism  $q' : C_k \rightarrow G$  with  $q \neq f$  and  $q = q' \circ a_k$  cannot exist. It follows immediately that  $q \models \forall(a_k : C_{k-1} \hookrightarrow C_k, f)$ .  $\square$

**Lemma 2.6.** Let  $G$  be a graph,  $p : C_0 \rightarrow G$  a morphism and  $c$  a condition over  $C_0$  in ANF with  $p \models_k c$ . If the subcondition  $d$  of  $c$  with  $\text{lay}(d) = k$  is universally bound,

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

*Proof.* Follows immediately by using lemma 2.5 and setting  $f$  to the subcondition of  $c$  with layer  $k + 1$ .  $\square$

**Lemma 2.7.** Let  $c$  be a condition in ANF over  $C_0$  and  $p : C_0 \hookrightarrow G$  a morphism with  $p \models_k c$ . Let  $d = Q(a_{k+2} : C_{k+1} \hookrightarrow C_{k+2}, e)$  be the subcondition of  $c$  with layer  $k + 2$ . There does exist a graph  $C_{k+1} \subseteq C' \subseteq C_{k+2}$  such that

$$p \models \text{sub}(k+1, c, Q(a_{k+2} : C_{k+1} \hookrightarrow C', f))$$

with  $f$  being a  $\overline{Q}$  bound condition over  $C'$ .

*Proof.* If  $p \models c$ , we can choose  $C' = C_{k+2}$  and  $f = e$ .

If  $p \not\models c$ , there does not exists a  $j$  with  $p \models_j c$  and the subcondition of  $c$  with layer  $j$  is universally bound and  $Q = \exists$  follows immediately. We choose  $C' = C_{k+1}$  and  $f = \text{true}$ . Let  $q : C_k \rightarrow G$  with  $p = q \circ a_k \circ \dots \circ a_1$  and  $q \circ \dots \circ a_\ell$  satisfying the condition up to  $\ell - k$  of the subcondition of  $c$  at layer  $\ell$  for all  $0 \leq \ell \leq k$ . This morphism must exists since  $p \models_k c$  and  $p \not\models c$ . Let  $q' : C_{k+1} \rightarrow G$  be a morphism with  $q = q' \circ a_{k+1}$ . Since  $C' = C_{k+1}$ , the morphism  $a'_{k+2}$  has to be the identity and therefore  $q' = q' \circ a'_{k+2}$ . It follows that  $q' \models \exists(a'_{k+2} : C_{k+1} \hookrightarrow C', \text{true})$  and therefore  $p \models \text{sub}(k+1, c, Q(a_{k+2} : C_{k+1} \hookrightarrow C', f))$ .  $\square$

**Definition 2.8 (partial condition).** Let  $c$  be a condition in ANF over  $C_0$ . Let  $d$  be the subcondition of  $c$  at layer  $k + 1$ . The partial condition of  $c$  at layer  $k$  with  $C'$ ,  $\text{part}(k, c, C')$  is defined as:

1. If  $d$  is universally bound, let  $e = \exists(a : C_{k+1} \hookrightarrow C_{k+2}, f)$  be the subcondition of  $c$  at layer  $k + 2$  with  $C_{k+1} \subseteq C' \subseteq C_{k+2}$ :

$$\text{part}(k, c, C') := \text{sub}(k+2, c, \exists(a : C_{k+1} \hookrightarrow C', \text{true}))$$

2. If  $d = \exists(a : C_k \hookrightarrow C_{k+1}, f)$  is existentially bound with  $C_k \subseteq C' \subseteq C_{k+1}$ :

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

**Definition 2.9 (biggest partially satisfying condition).** Let  $G$  be a graph,  $c$  a condition over  $C_0$  and  $p : C_0 \hookrightarrow G$  a morphism with  $p \models_k c$ .

A partial condition  $c = \text{part}(k, c, C')$  with  $p \models c$  is a biggest partially satisfying condition if there does not exist a graph  $C'' \subset C'$  with  $p \models \text{part}(k, c, C'')$ .

The set of graphs  $C'$  such that  $\text{part}(k, c, C')$  is a biggest partially satisfying condition is denoted by  $\mathcal{G}_{k,c}$ .

## 2.2 minimal consistency improving

**Definition 2.10 (number of violations).** Let  $G$  be a graph and  $c$  a constraint in ANF, such that  $G \models_k c$  and  $G \not\models c$ . The number of violations  $nvc(k, G)$  at layer  $k$  in  $G$  is set to 0 if there does exist a  $j > k$  with  $G \models_j c$  and otherwise:

1. If  $k = 0$  and  $c = \exists(a : C_0 \hookrightarrow C_1, e)$  is existentially bound:

$$nvc(k, G) := |\{C' \mid G \not\models \text{part}(0, c, C') \wedge C_0 \subseteq C' \subseteq C_1\}|$$

2. If  $k \neq 0$ , let  $d = \forall(a_k : C_{k+1} \hookrightarrow C_{k+2}, e)$  be the subcondition of  $c$  at layer  $k + 1$ .

$$nvc(k, G) := \sum_{C \in \mathcal{G}_{k,c}} \sum_{C \subset C'} |\{q \mid q : C_{k+2} \hookrightarrow G \wedge q \not\models \text{part}(1, e, C')\}|$$

**Definition 2.11 (minimal consistency improving).** Let  $G$  be a graph,  $r$  a rule and  $c$  a constraint in ANF with  $G \models_k c$  and  $G \not\models c$ .

A transformation  $G \xrightarrow{r,m} H$  is called minimal consistency improving, if

$$nvc(k, H) < nvc(k, G).$$

A rule  $r$  is called minimal consistency improving, if all of its applications are.

**Lemma 2.12.** Let  $G$  be  $p : C_0 \rightarrow G$  a morphism,  $c$  a constraint in ANF over  $C_0$  with  $p \models_k c$ . Then  $p \models_j c$  for all  $j < k$  such that the subcondition of  $c$  at layer  $j$  is existentially bound.

*Proof.* 1. The subcondition of  $c$  at layer  $k$  is existentially bound: If an  $j < k$  with  $p \models_j c$  exists such that the subcondition of  $c$  at layer  $j$  is universally bound, let  $j_1$  be the smallest of these. With lemma 2.5 follows that  $p \models_{j_2} c$  for all  $j_1 < j_2$ . Let  $\ell < j_1$ , such that the subcondition of  $c$  at layer  $j$  is existentially bound and let  $d = \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, e)$  be the condition up to layer  $j_1 - \ell$  of the subcondition of  $c$  at layer  $\ell$ . Since  $\ell < j_1$ , a morphism  $q : C_\ell \rightarrow G$  with  $q \models d$  must exists and therefore a morphism  $q' : C_{\ell+1} \rightarrow G$  with  $q = q' \circ a_\ell$  must exists. It follows that  $q \models \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, \text{true})$  and with that  $p \models_\ell c$ .

2. The subcondition of  $c$  at layer  $k$  is universally bound: With lemma 2.5 follows that  $p \models_{k+1} c$ . Since  $c$  is in ANF 1. can be applied to  $k + 1$ .

□

**Lemma 2.13.** Let  $G$  be a graph,  $r$  a rule and  $c$  a constraint in ANF with  $G \not\models c$ . Let  $k$  be the biggest number, such that  $G \models_k c$ . A transformation  $G \xrightarrow{r,m} H$  is minimal consistency improving if  $G \models_j c$  and  $k < j$ .

*Proof.* Since  $G \not\models c$ , with lemma 2.6 follows that the subcondition of  $c$  at layer  $k$  has to be existentially bound and since  $k$  is the biggest number such that  $G \models_k c$  it follows that  $\text{nvc}(k+1, G) > 0$ . If the subcondition of  $c$  at layer  $j$  is universally bound,  $H \models c$  follows with lemma 2.6 and lemma 2.12  $H \models_{k+2} c$ . Therefore  $\text{nvc}(k+2, H) = 0$ . Otherwise the subcondition of  $c$  at layer  $j$  is existentially bound and therefore  $G \models_{k+2} c$  and  $\text{nvc}(k+2, H) = 0$  follows immediately.  $\square$

**Definition 2.14 (direct minimal consistency improving).** Let  $G$  be a graph,  $r$  a plain rule and  $c$  a constraint in ANF with  $G \models_k c$  and  $G \not\models c$ . Let  $d = \forall(a_k : C_k \hookrightarrow C_{k+1}, e)$  be the condition at layer  $k$  of  $c$ . A transformation  $t : G \xrightarrow{r,m} H$  is called direct minimal consistency improving if the transformation is minimal consistency improving and

$$\begin{aligned} \forall p : C_{k+1} \hookrightarrow G((p \models e \wedge \text{tr}_t \circ p \text{ is total}) \implies \text{tr}_t \circ p \models p) \wedge \\ \forall p' : C_{k+1} \hookrightarrow H(\neg \exists p : C \hookrightarrow G(p' = \text{tr}_t \circ p) \implies p' \models d). \end{aligned} \quad (2.1)$$

### 3 application condition

**Definition 3.1 (extended overlap).** Let  $G$  and  $C_0 \subset C_1$  be graphs and  $C'$  an overlap of  $G$  and  $C_0$  with overlap morphisms  $p : G \hookrightarrow C'$  and  $q : C_0 \hookrightarrow C'$ . An overlap  $C''$  of  $G$  and  $C_1$  is called the extended overlap of  $C'$  with  $C''$  if  $C' \subset C''$  and a morphism  $q' : C_1 \hookrightarrow C''$  with  $q'|_{C_0} = q$  exists.

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

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

Let  $r = L \xleftarrow{l} K' \xrightarrow{r} R$  be the rule with

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

The graph  $H$ , of the transformation  $G \xrightarrow{r,p} H$  is called the overlap shifted graph of  $C'$ .

**Definition 3.3 (minimal consistency improving application condition).** Let  $r = L \xleftarrow{l} K \xrightarrow{r} R$  be a plain rule and  $c$  a constraint in ANF. Let  $d = \forall(a_k : C_k \hookrightarrow C_{k+1}, \exists(b : C_{k+1} \hookrightarrow C_{k+2}, e))$  be the condition at layer  $k$  of  $c$ . The application condition  $\text{ap}_k$  of the condition at layer  $k$  of  $c$  and  $C_{k+1} \subset C' \subseteq C_{k+2}$  is defined as:

$$\begin{aligned} \text{ap}(k, C') := & \left( \bigvee_{P \in \text{ol}(L, C_{k+1})} \text{nex}(P, C') \wedge (\text{rep}(P, C') \vee \text{del}(P, C')) \right) \wedge \\ & \left( \bigwedge_{P \in \text{ol}(L, C_{k+1})} \text{ex}(P, C') \right) \wedge \left( \bigwedge_{P \in \text{ol}(L, C_{k+1})} \text{rem}(P, C') \right) \end{aligned} \quad (3.2)$$

with

1. Let  $Q$  be the extended overlap of  $P$  with  $C'$ .

$$\text{nex}(P, C') := \exists(a : L \hookrightarrow P, \neg \exists(b : P \hookrightarrow Q, \text{true}))$$

2. Let  $Q$  be the extended overlap of  $P$  with  $C'$ ,  $Q'$  the overlap shifted graph of  $Q$  and  $P'$  the overlap shifted graph of  $P$ .

$$\text{rep}(P, C') := \text{Left}(\forall(a : L \hookrightarrow P', \neg \exists(b : P' \hookrightarrow Q', \text{true})), r)$$

3. Let  $P'$  be the overlap shifted graph of  $P$ .

$$\text{del}(P, C') := \text{Left}(\neg \exists(a : R \hookrightarrow P', \text{true}), r)$$

4. Let  $Q$  be the extended overlap of  $P$  with  $C'$  and  $Q'$  the overlap shifted graph of  $Q$ .

$$\text{ex}(P, C') := \exists(a : L \hookrightarrow Q, \text{true}) \implies \text{Left}(\exists(a : R \hookrightarrow Q', \text{true}), r)$$

5.  $Q$  be the extended overlap of  $P$  with  $C'$  and  $Q'$  the overlap shifted graph of  $Q$ .

$$\text{rem}(P, C') := \exists(a : L \hookrightarrow Q, \text{true}) \implies \text{Left}(\exists(b : R \hookrightarrow Q', \text{true}), r)$$

**Lemma 3.4.** Let  $G$  be a graph,  $c$  a constraint in ANF, such that  $G \models_k c$  and  $r$  a plain rule. Let  $d = \forall(a_k : C_k \hookrightarrow C_{k+1}, \exists(b : C_{k+1} \hookrightarrow C_{k+2}, e))$  be the condition at layer  $k$  of  $c$ . Then, every application of  $r$ , equipped with  $\text{ap}(k, C')$  with  $C' \notin \mathcal{B}_{k,\text{true}}$ , such that a  $C \in \mathcal{B}_{k,\text{true}}$  with  $C \subset C'$  exists is minimal consistency improving.

*Proof.*

□

**Lemma 3.5.** Let  $G$  be a graph,  $c$  a constraint in ANF, such that  $G \models_k c$ , and  $r = L \xleftarrow{l} K \xrightarrow{r} R$  a plain rule. Let  $d = \forall(a_k : C_k \hookrightarrow C_{k+1}, \exists(b : C_{k+1} \hookrightarrow C_{k+2}, e))$  be the condition at layer  $k$  of  $c$  and  $\text{ap}(k, C')$  with  $C' \notin \mathcal{B}_{k,\text{true}}$ , such that a  $C \in \mathcal{B}_{k,\text{true}}$  with  $C \subset C'$  exists. The following simplifications apply:

1. Let  $P \in \text{ol}(L, C_{k+1})$ . If an injective morphism  $p : P \hookrightarrow G$  does not exist,  $\text{nex}(P, C')$ ,  $\text{rep}(P, C')$ ,  $\text{del}(P, C')$  and  $\text{ex}(P, C')$  can be replaced by *false*.
2. If  $(L \setminus K) \cap C_{k+1} = \emptyset$ , every  $\text{del}(P, C')$  can be replaced by *false*.
3. If  $(R \setminus K) \cap C' = \emptyset$ , every  $\text{rep}(P, C')$  can be replaced by *false*.
4. If 1. and 2. apply,  $\text{ap}(k, C')$  can be replaced by *false*.
5. If  $(R \setminus K) \cap C_{k+1} = \emptyset$ , every  $\text{rem}(P, C')$  can be replaced by *true*.

*Proof.*

□

## References

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