

# Masterarbeit

Alexander Lauer

September 2, 2022

## **Abstract**

# Contents

<b>1 Preliminaries</b>	<b>3</b>
<b>2 partial consistency improving</b>	<b>4</b>
2.1 partial-conditions and -satisfiability . . . . .	4
2.2 minimal consistency improving . . . . .	4
<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 partial-conditions and -satisfiability

**Definition 2.1 (partial condition).** Let  $c$  be a condition over  $C_0$ . A partial condition of  $c$  over  $C'_0 \subseteq C_0$  is defined as:

1. *true* is the partial condition of *true* for every morphism.
2. if  $c = Q(a : C_0 \hookrightarrow C_1, d)$ , with  $Q \in \{\exists, \forall\}$ , a partial condition of  $c$  over  $C'_0$  is given by  $\exists(a' : C'_0 \hookrightarrow C'_1, d')$  with  $a'$  being a partial morphism of  $a$ ,  $C'_1 \subseteq C_1 \setminus a(C_0 \setminus C'_0)$  and  $d'$  is a partial condition of  $d$  over  $C'_1$ .
3. if  $c = d_1 \wedge d_2$  or  $c = d_1 \vee d_2$  the partial condition of  $c$  is given by  $d'_1 \wedge d'_2$  and  $d'_1 \vee d'_2$ , respectively, with  $d'_1$  and  $d'_2$  being partial conditions of  $d_1$  and  $d_2$  over  $C'_0$ .
4. if  $c = \neg d$  the partial condition of  $c$  is given by  $\neg d'$  with  $d'$  being a partial condition of  $d$  over  $C'_0$ .

A partial condition  $Q(a : C'_0 \hookrightarrow C'_1, d)$  of  $c$  over  $C'_0$  is called the closest partial condition of  $c$  over  $C'_0$  if  $C'_1 = C_1 \setminus a(C_0 \setminus C'_0)$  and  $d$  is the closest partial condition of  $d$  over  $C'_1$ .

We use the notation  $c' \leq c$  if  $c'$  is partial condition of  $c$  over  $C'_0 \subseteq C_0$  and  $c' < c$  if  $C'_i \subset C_i$  for any  $i$ .

**Definition 2.2 (partial satisfiability).** Let a condition  $c$  over  $C_0$  and a graph  $G$  be given. A morphism  $p_0 : C'_0 \rightarrow G$ , with  $C'_0 \subseteq C_0$ , partial satisfies  $c$ ,  $p_0 \models_p c$ , if a partial morphism  $p'_0 : C''_0 \rightarrow G$  of  $p_0$  satisfies a partial condition of  $c$  over  $C''_0$ .

Note, that  $p_0 \models c$  implies  $p_0 \models_p c$

### 2.2 minimal consistency improving

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

Let  $L$  be the set of subconditions of  $c$  with layer  $k$ . The condition up to layer  $k$  of  $c$  is the condition which is obtained by replacing each  $d = Q(a : P \hookrightarrow P', e) \in L$  by  $d = Q(a : P \hookrightarrow P', \text{true})$  if  $Q = \exists$  and by  $Q(a : P \hookrightarrow P', \text{false})$  if  $Q = \forall$  in  $c$ . If  $k = 0$ , the condition up to layer  $k$  of every condition  $c$  is *true*.

**Definition 2.4 (layered consistency).** Let  $G$  be a graph,  $p : C_0 \rightarrow G$  a morphism,  $c$  a condition over  $C_0$  and  $d$  the condition up to layer  $k$  of  $c$ . The morphism  $p$  satisfies  $c$  up to layer  $k$  if  $p \models d$ . The notation  $p \models_k c$  is used if  $p$  satisfies  $c$  up to layer  $k$ . The notation  $G \models_k c$  is used if  $c$  is a constraint and  $p \models_k c$  for  $p : \emptyset \rightarrow G$ .

**Lemma 2.5.** Let  $G$  be a graph  $p : C_0 \rightarrow G$  a morphism and  $c = Q(a_1 : C_0 \hookrightarrow C_1, \overline{Q}(a_2 : C_1 \hookrightarrow C_2, \dots))$  a condition over  $C_0$  in ANF with  $p \models_k c$ . If the subcondition  $d = \forall(a_k : C_{k-1} \hookrightarrow C_k, e)$  of  $c$  at layer  $k$  is universally bound,  $p \models Q(a_1 : C_0 \hookrightarrow C_1, \overline{Q}(a_2 : C_1 \hookrightarrow C_2, \dots \forall(a_k : C_{k-1} \hookrightarrow C_k, f)))$  for every condition  $f$  over  $C_k$ .

*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  $q \not\models f$  and  $q = q' \circ a_k$  cannot exist. It follows immediately that  $q \models \forall(a_k : C_{k-1} \hookrightarrow C_k, f)$ .  $\square$

**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  $c = Q(a_1 : C_0, \hookrightarrow C_1, \dots)$  be a condition in ANF over  $C_0$  and  $p : C_0 \rightarrow G$  a morphism with  $p \models_k c$ . Let  $d = Q'(a_{k+1} : C_k \hookrightarrow C_{k+1}, \overline{Q}'(a_{k+2} : C_{k+1} \hookrightarrow C_{k+2}, e))$  be the subcondition of  $c$  with layer  $k+1$ . There does exist a graph  $C_{k+1} \subseteq C' \subseteq C_{k+2}$  such that  $p \models Q(a_1 : C_0, \hookrightarrow C_1, \dots, Q'(a_{k+1} : C_k \hookrightarrow C_{k+1}, \overline{Q}'(a'_{k+2} : C_{k+1} \hookrightarrow C', f)))$  with  $f$  being a  $Q'$  bound condition over  $C'$  and  $a'_{k+2}$  being a partial morphism of  $a_{k+2}$ .*

*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  $Q' = \forall$  follows. 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  $p \models Q(a_1 : C_0, \hookrightarrow C_1, \dots, \forall(a_{k+1} : C_k \hookrightarrow C_{k+1}, \exists(a_{k+2} : C_{k+1} \hookrightarrow C_{k+1}, \text{true})))$ .  $\square$

**Definition 2.8 (biggest partially satisfying graph).** *Let  $G$  be a graph,  $c = Q(a_1 : C_0, \hookrightarrow C_1, \dots)$  a condition in ANF and  $p : C_0 \rightarrow G$  a morphism with  $p \models_k c$ . Let  $d = Q'(a_{k+1} : C_k \hookrightarrow C_{k+1}, \overline{Q}'(a_{k+2} : C_{k+1} \hookrightarrow C_{k+2}, e))$  be the subcondition of  $c$  with  $\text{lay}(d) = k+1$ . With lemma 2.7 there exists a graph  $C_{k+1} \subseteq C' \subseteq C_{k+2}$  and an  $Q'$  bound condition  $e$  over  $C'$  in ANF with  $p \models Q(a_1 : C_0, \hookrightarrow C_1, \dots, Q'(a_{k+1} : C_k \hookrightarrow C_{k+1}, \overline{Q}'(a_{k+2} : C_{k+1} \hookrightarrow C', e))) =: d_{C',e}$ . Let  $\mathcal{C}_{k,p,e}$  be the set of graphs  $C_{k+1} \subseteq C' \subseteq C_{k+2}$  with  $p \models d_{C',e}$ .*

*A graph  $C' \in \mathcal{C}_{k,p,e}$ , such that no  $C'' \in \mathcal{C}$  with  $C' \subseteq C''$  exists is called a biggest partially satisfying graph of  $c$  at level  $k$  with  $p$  and  $e$ . The set of these graphs is denoted by  $\mathcal{B}_{k,p,e}$ .*

**Definition 2.9.** *Let  $G$  be a graph and  $c = Q(a_1 : \emptyset \hookrightarrow C_1, \dots)$  a constraint in ANF, such that  $G \models_k c$  and  $G \not\models c$ . Let  $d = Q(a_1 : \emptyset \hookrightarrow C_1, \dots, \forall(a_{k+1} : C_k \hookrightarrow C_{k+1}, \exists(a_{k+2} : C_{k+1} \hookrightarrow C_{k+2}, \text{true})))$  be the condition up to layer  $k+2$  of  $c$ . The number of violations of  $c$  up to layer  $k+2$  in  $G$  is defined as the the number of morphisms  $q : C_{k+2} \rightarrow G$  that do not satisfy  $\exists(a_{k+2} : C_{k+1} \hookrightarrow C_{k+2}, \text{true})$ . This number is denoted by  $\text{nv}_c(k+2, G)$ .*

**Definition 2.10 (minimal consistency improving).** Let  $G$  be a graph,  $r$  a rule and  $c$  a constraint in ANF with  $G \models_k c$  and  $G \not\models c$ . With lemma 2.6 follows that the subcondition of  $c$  with layer  $k$  has to be existentially bound.

A transformation  $G \xrightarrow{r,m} H$  is called minimal consistency improving if either  $H \models c$  or  $\text{nv}_c(k+2, G) < \text{nv}_c(k+2, H)$  or every  $C' \in \mathcal{B}_{k,G,\text{true}}$  is also an element of  $\mathcal{C}_{k+2,H,e}$  and for at least one  $C'$  an  $C'' \in \mathcal{C}_{k+2,H,\text{true}}$  with  $C' \subseteq C''$  exists. The rule  $r$  is called minimal consistency improving if all of its applications are.

**Lemma 2.11.** 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  $q = q' \circ a_k$  must exists. It follows that  $q \models \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, \text{true})$  and with that  $p \models_\ell c$ .
2. The subcondition of  $c$  at layer  $k$  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 2.12.** 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.

The subcondition of  $c$  at layer  $j$  is existentially bound: With lemma 2.11 follows that  $H \models_{k+2} c$ . Let  $d = \exists(a_{k+2} : C_{k+1} \hookrightarrow C_{k+2}, e)$  be the subcondition of  $c$  with layer  $k+2$ . The set  $\mathcal{C}_{k+2,H,\text{true}}$  contains every graph  $C'$  with  $C_{k+1} \subseteq C' \subseteq C_{k+2}$ . Since  $G \not\models_{k+2} c$ ,  $C_{k+2} \notin \mathcal{C}_{k+2,G,\text{true}}$  and for every  $C'' \in \mathcal{B}_{k+2,G,\text{true}}$  it holds that  $C'' \in \mathcal{C}_{k+2,H,\text{true}}$  and  $C'' \subset C_{k+2} \in \mathcal{C}_{k+2,H,\text{true}}$ .

2. The subcondition of  $c$  at layer  $j$  is universally bound: With lemmas 2.5 and 2.6,  $H \models_{k+2} c$  follows and 1. can be applied.

□

### 3 application condition

**Construction 3.1.** Given a rule  $r = L \xleftarrow{L} K \xrightarrow{r} R$  and a constraint  $c$  in ANF with  $\text{nl}(c) = 1$ . Let  $G$  and  $H$  be graphs. We denote the set of all overlaps of  $G$  and  $H$  as  $\text{ol}(G, H)$ . We construct the following application conditions:

1. if  $c = \forall(a_0 : C_0 \rightarrow C_1, \text{false})$ :

$$\bigvee_{Q \in \text{ol}(L, C_1)} \left( \exists(a : L \hookrightarrow Q, \text{true}) \wedge \text{Left}(\neg \exists(b : R \rightarrow Q', \text{true})) \right)$$

with  $Q' \setminus (R \setminus (K \cup C_1)) = Q \setminus (L \setminus (K \cup C_1))$ .

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

$$\bigvee_{Q \in \text{ol}(L, C_1)} \bigvee_{Q' \subset Q} \left( \exists(a : L \rightarrow Q', \text{true}) \wedge \left( \bigvee_{\substack{\overline{Q} \subset Q \\ Q' \subset \overline{Q}}} \text{Left}(\exists(a : R \rightarrow \overline{Q}, \text{true})) \right) \right)$$

## References

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