Masterarbeit

Alexander Lauer

September 9, 2022

Abstract

Contents

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* $\mathrm{sub}(G_1)$.

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* $\mathrm{ol}(G, G')$.

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* $\mathrm{ol}_p(G, C')$.

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 c. 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 nl *of a condition is defined as* $\mathrm{nl}(\mathsf{true} = 0$ *and* $\mathrm{nl}(\exists(a : P \rightarrow Q, d)) := \mathrm{nl}(d) + 1$.

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*

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

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_1' 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 $p_0 : C_0' \to G$, with $C_0' \subseteq C_0$, partial satisfies c, $p_0 \models_p c$, if a partial morphism $p_0' : C_0'' \to G$ 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 (**Layer of a subcondition**). *Let c be a condition and d a subcondition of c. The layer of d is defined as $\mathrm{lay}(d) := \mathrm{nl}(c) - \mathrm{nl}(d) - 1$.*

Definition 2.4 (**substitution at layer**). *Let c be a condition, such that the subcondition of c with layer $0 \leq k \leq \mathrm{nl}(c)$ 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, $\mathrm{sub}(k, c, e)$, *is recursively defined as:*

1. *If $k = 0$:*
$$\mathrm{sub}(0, c, e) := e$$
2. *If $k > 0$:*
$$\mathrm{sub}(k, c, e) := Q\big(a : C_0 \hookrightarrow C_1, \mathrm{sub}(k-1, d, e)\big)$$

Definition 2.5 (**Condition up to layer**). *Let c be a condition and d be the subcondition of c at layer $0 \leq k \leq \mathrm{nl}(c)$. The* condition up to layer k *of c,* $\mathrm{cond}(k, c)$ *is defined as*

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

Definition 2.6 (**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 satisfies* $\operatorname{cond}(k, c)$.

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

Lemma 2.7. *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:*

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

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 $q : G_{k-1} \to G$ be a morphism such that $q \models \forall(a_k : C_{k-1} \hookrightarrow C_k, \mathsf{false})$. 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 $q' : C_k \to G$ with $q = q' \circ a_k$. Hence, for every condition f over C_k a morphism $q' : C_k \to G$ with $q \not\models f$ and $q = q' \circ a_k$ cannot exist. It follows immediately that $q \models \forall(a_k : C_{k-1} \hookrightarrow C_k, f)$. □

Lemma 2.8. *Let G be a graph, $p : C_0 \to G$ a morphism and c a condition over C_0 in ANF with $p \models_k c$. If the subcondition d of c with* $\operatorname{lay}(d) = k$ *is universally bound,*

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

Proof. Follows immediately by using lemma 2.7 and setting f to the subcondition of c with layer $k + 1$. □

Lemma 2.9. *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*

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

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 $f = \mathsf{true}$. Let $q : C_k \to G$ with $p = q \circ a_k \circ \ldots \circ a_1$ and $q \circ \ldots \circ a_\ell$ satisfying the condition up to $\ell - k$ of the subcondition of c at layer ℓ for all $0 \leq \ell \leq k$. This morphism must exists since $p \models_k c$ and $p \not\models c$. Let $q' : C_{k+1} \to G$ 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 $q' \models \exists(a'_{k+2} : C_{k+1} \hookrightarrow C', \mathsf{true})$ and therefore $p \models \operatorname{sub}\big(k + 1, c, Q(a_{k+2} : C_{k+1} \hookrightarrow C', f)\big)$. □

Definition 2.10 (**biggest partially satisfying graph**)**.** *Let G be a graph, $c = Q(a_1 : C_0, \hookrightarrow C_1, \ldots)$ a condition in ANF and $p : C_0 \to G$ 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* $\mathrm{lay}(d) = k+1)$*. With lemma 2.9 there exists a graph $C_{k+1} \subseteq C' \subseteq C_{k+2}$ and an Q' bound condition e over C' in ANF with $p \models Q(a_1 : C_0, \hookrightarrow C_1, \ldots Q'(a_{k+1} : C_k \hookrightarrow C_{k+1}, \overline{Q}(a_{k+2} : C_{k+1} \hookrightarrow C', e)) =: d_{C',e}$. 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 $C'' \in \mathrm{C}$ 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 $\mathcal{B}_{k,e}$*

Definition 2.11. *Let G be a graph and c a constraint in ANF, such that $G \models_k c$ and $G \not\models c$. Let $d = \forall(a_{k+1} : C_k \hookrightarrow C_{k+1}, e)$ be the condition up to layer $k+2$ of of the condition at layer $k+1$ of c. The number of violations of c up to layer $k+2$ in G is defined as the the number of morphisms $q : C_{k+1} \to G$ that do not satisfy e, with C' being a smallest graph such that $C \subset C'$ for at least one $C \in \mathcal{B}_{k,\mathsf{true}}$ if $G \not\models_{k+2} c$ and 0 otherwise. This number is denoted by* $\mathrm{nvc}(k+2, G)$.

Definition 2.12 (**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*

$$\mathrm{nvc}(k, H) < \mathrm{nvc}(k, G).$$

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

Lemma 2.13. *Let G be $p : C_0 \to G$ 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 2.7 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 ℓ. Since $\ell < j_1$, a morphism $q : C_\ell \to G$ with $q \models d$ must exists and therefore a morphism $q' : C_{\ell+1} \to G$ with $q = q' \circ a_k$ must exists. If follows that $q \models \exists(a_\ell : C_\ell \hookrightarrow C_{\ell+1}, \mathsf{true})$ and with that $p \models_\ell c$.

2. The subcondition of c at layer k is universally bound: With lemma 2.7 follows that $p \models_{k+1} c$. Since c is in ANF 1. can be applied to $k+1$.

□

Lemma 2.14. *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 2.8 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 $\mathrm{nvc}(k+1, G) > 0$. If the subcondition of c at layer j is universally bound, $H \models c$ follows with lemma 2.8 and lemma 2.13 $H \models_{k+2} c$. Therefore $\mathrm{nvc}(k+2, H) = 0$. Otherwise the subcondition of c at layer j is existentially bound and therefore $G \models_{k+2} c$ and $\mathrm{nvc}(k+2, H) = 0$ follows immediately. □

Definition 2.15 (**direct minimal consistency improving**). *Let G be a graph, r a plain rule and c a constraint in ANF with $G \models_k$ 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*

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

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 $r = L \overset{l}{\hookleftarrow} K \overset{r}{\hookrightarrow} R$ be a plain rule, C a graph and C' an overlap of C and L. An overlap G of R and C is called an* overlap shifted graph *of C' if*

$$G = C' \setminus (L \setminus (K \cap C)) \cup R \setminus (K \cap C) \dot{\cup} R \setminus (K \setminus (K \cap C))$$

Definition 3.3 (**minimal consistency improving application condition**). *Let $r = L \overset{l}{\hookleftarrow} K \overset{r}{\hookrightarrow} R$ 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 ap_k of the condition at layer k of c and $C_{k+1} \subset C' \subseteq C_{k+2}$ is defined as:*

$$\begin{aligned} \mathrm{ap}(k, C') := &\Big(\bigvee_{P \in \mathrm{ol}(L, C_{k+1})} \mathrm{nex}(P, C') \wedge (\mathrm{rep}(P, C') \vee \mathrm{del}(P, C'))\Big) \wedge \\ &\Big(\bigwedge_{P \in \mathrm{ol}(L, C_{k+1})} \mathrm{ex}(P, C')\Big) \wedge \Big(\bigwedge_{P \in \mathrm{ol}(L, C_{k+1})} \mathrm{rem}(P, C')\Big) \end{aligned} \tag{3.1}$$

with

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

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

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.*

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

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

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

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

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

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

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

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* $\operatorname{ap}(k, C')$ *with $C' \notin \mathcal{B}_{k,\textsf{true}}$, such that a $C \in \mathcal{B}_{k,\textsf{true}}$ 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 $r = L \xhookleftarrow{l} K \xhookrightarrow{r} R$ 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* $\operatorname{ap}(k, C')$ *with $C' \notin \mathcal{B}_{k,\textsf{true}}$, such that a $C \in \mathcal{B}_{k,\textsf{true}}$ with $C \subset C'$ exists. The following simplifications apply:*

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

Proof. □

References

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