Rewrite Rules for a Solver for Restricted Intensional Sets

Maximiliano Cristiá
Universidad Nacional de Rosario

Gianfranco Rossi and Andrea Fois
Università di Parma

June 26, 2019

Abstract

This document lists in a compact way all the rewrite rules used in the constraint solver for $\mathcal{L}_{\mathcal{RIS}}$, a constraint language which provides both extensional and intensional finite sets (namely, restricted intensional sets), along with basic operations on them. The constraint solver for $\mathcal{L}_{\mathcal{RIS}}$ takes the form of a rewrite system acting on $\mathcal{L}_{\mathcal{RIS}}$ formulas, i.e., quantifier-free conjunctions and disjunctions of positive and negative $\mathcal{L}_{\mathcal{RIS}}$ constraints.

1 Basic definitions and notation

$\mathcal{L}_{\mathcal{RIS}}$ is parametric w.r.t. a first-order theory $\mathcal{X}$, along with its first-order logic language $\mathcal{L}_{\mathcal{X}}$ and its (complete) solver $SAT_{\mathcal{X}}$. $\mathcal{X}$ is assumed to contain at least the predicate symbols $=_{\mathcal{X}}$ and $\neq_{\mathcal{X}}$ which are interpreted as the identity and not identity in the interpretation domain of $\mathcal{X}$, respectively.

$\mathcal{L}_{\mathcal{RIS}}$ terms are called *set terms*, while $\mathcal{L}_{\mathcal{X}}$ terms are called *non-set terms*. Set terms can have the following forms: *a*) the empty set, noted $\emptyset$; *b*) *extensional sets*, noted $\{x \sqcup A\}$, where x, called *element part*, is a non-set term, and A, called *set part*, is a set term; and *c*) *restricted intensional sets* (RIS), noted $\{c : D \mid F \bullet P[c]\}$, where c, called *control term*, is a non-set term; D, called *domain*, is an extensional set term (either variable or not); F, called *filter*, is a $\mathcal{X}$-formula; and P, called *pattern*, is a non-set term containing c. Variables which range over set terms (resp., non-set terms) are called *set variables* (resp., *non-set variables*).

A RIS term is a *variable-RIS* if its domain is a variable; otherwise it is a *non-variable RIS*. Moreover, we say that a RIS term is *simple* if its control term is identical to its pattern (i.e., it has the form $\{c : D \mid F \bullet c\}$).

The set of primitive predicate symbols Π for $\mathcal{L}_{\mathcal{RIS}}$ is $\{=, \neq, \in, \notin, un, \parallel\}$. A π-*constraint* is any $\mathcal{L}_{\mathcal{RIS}}$ literal based on π. A *rewrite rule for* π, $\pi \in \Pi$, is a rewrite rule of the form:

$$\phi \rightarrow \Phi$$

where ϕ is a π-constraint and Φ is a $\mathcal{L}_{\mathcal{RIS}}$ formula. If Φ has more than one disjunct then the rule is non-deterministic. Conjunctions occurring in Φ have higher precedence than disjunctions.

A *rewriting procedure* for π-constraints consists of the collection of all the rewrite rules for π-constraints. The following subsections present the rewriting procedure for all primitive $\mathcal{L}_{\mathcal{RIS}}$ constraints. For each rewriting procedure, the solver selects rules in the order they are presented. The first rule whose left-hand side matches the input π-constraint c is used to rewrite c. If no rules applies to c, then c is left unchanged (i.e., c is *irreducible*).

Notational conventions

- t, u, c, d (possibly subscripted) stand for any *non-set* term; A, B, C, D stand for any *set* term (extensional set or RIS, variable or not); $\bar{A}$, $\bar{B}$, $\bar{C}$, $\bar{D}$, represent either *set variables* or *variable-RIS*. X, Y, N are set variables (not RIS), while x, y, n are non-set variables. $\emptyset$ represents an empty set or a RIS with empty domain. $\boldsymbol{X}$ represents a set variable X or a variable-RIS whose domain is X or a RIS whose domain is a set term containing X as its innermost variable set part (e.g., $\{t_0, ..., t_n \sqcup \boldsymbol{X}\}$ stands for $\{t_0, ..., t_n \sqcup X\}$, or $\{t_0, ..., t_n \sqcup \{X \mid F \bullet P\}\}$, or $\{t_0, ..., t_n \sqcup \{\{u_0, ..., u_m \sqcup X\} \mid F \bullet P\}\}$).
- variable names n and N (possibly with sub and superscripts) are used to denote *fresh* variables (non-set variables and set variables, respectively);

- $\{t_1, t_2, \dots, t_n \sqcup t\}$ (resp., $\{t_1, t_2, \dots, t_n\}$) is a shorthand for $\{t_1 \sqcup \{t_2 \sqcup \cdots \{t_n \sqcup t\} \cdots\}\}$ (resp., $\{t_1 \sqcup \{t_2 \sqcup \cdots \{t_n \sqcup \emptyset\} \cdots\}\}$)

- $t_1 \equiv t_2$ (resp., $t_1 \not\equiv t_2$), for any terms t_1 and t_2, means that t_1 is (resp., is not) of the same (syntactic) form as t_2;

- $vars(t_1, \dots, t_n)$ denotes the set of variables occurring in $t_1, \dots, t_n$;

- s^σ, where s is any term (or formula) and $\sigma = \{x_1 \mapsto t_1, \dots, x_n \mapsto t_n\}$ is a *substitution*, denotes the term obtained by applying the substitution σ to s.

- The notation $F(t) \wedge n = P(t)$, used in the rewriting of a constraint involving the RIS $\{c : D \mid F \bullet P\}$, is a shorthand for the constraint generated by the following procedure:

  let σ be $\{x_1 \mapsto n_1, ..., x_k \mapsto n_k\}$, $\{x_1, ..., x_k\} = vars(c)$ and n_i fresh variables;
  $\phi \leftarrow SAT_{\mathcal{X}}(c^\sigma = t)$;
  if $\phi \neq false$
  then return $SAT_{\mathcal{X}}(\phi \wedge F^\sigma \wedge n = P^\sigma)$;
  else return $false$;

  Similarly, the notation $\neg F(t)$ is a shorthand for the constraint generated by the following procedure:

  let σ be $\{x_1 \mapsto n_1, ..., x_k \mapsto n_k\}$, $\{x_1, ..., x_k\} = vars(c)$ and n_i fresh variables;
  $\phi \leftarrow SAT_{\mathcal{X}}(c^\sigma = t)$;
  if $\phi \neq false$
  then return $SAT_{\mathcal{X}}(\phi \wedge \neg F^\sigma)$;
  else return $true$;

  Note that when the control term c is a variable, say x, then $SAT_{\mathcal{X}}(c^\sigma = t)$ is surely true and the effect of the above procedures amounts to simply call $SAT_{\mathcal{X}}$ on a formula based on F and P where all occurrences of (the renamed variant of) x are replaced by t.

- $\{D \mid F \bullet P\}$ is a shorthand for $\{c : D \mid F \bullet P\}$, for some control term c.

2 Equality $(=)$

1. $\emptyset = \emptyset \to true$
2. $X = X \to true$
3. If A is not a variable and $A \not\equiv \{\{d \sqcup D\} \mid F \bullet P\}$: [symmetrical case (rules 4, 5, 6)]
   $A = X \to X = A$
4. If $X \in vars(t_0, ..., t_n)$:
   $X = \{t_0, ..., t_n \sqcup A\} \to false$
5. $X = \{t_0, ..., t_n \sqcup \boldsymbol{X}\} \to X = \{t_0, ..., t_n \sqcup \boldsymbol{X}^{\{X \mapsto N\}}\}$ [special case]
6. If X occurs in other constraints in the input formula and $A \not\equiv \{\{d \sqcup D\} \mid F \bullet P\}$:
   $X = A \to X = A$ and substitute X by A in the rest of the formula.
7. $\{t \sqcup A\} = \emptyset \to false$
8. $\emptyset = \{t \sqcup A\} \to false$ [symmetrical case (rule 7)]
9. $\{t_0, ..., t_m \sqcup \boldsymbol{X}\} = \{u_0, ..., u_n \sqcup \boldsymbol{X}\} \to$ [special case]
   $t_0 = u_j \wedge \{t_1, ..., t_m \sqcup \boldsymbol{X}\} = \{u_0, ..., u_{j-1}, u_{j+1}, ..., u_n \sqcup \boldsymbol{X}\}$
   $\vee\ t_0 = u_j \wedge \{t_0, ..., t_m \sqcup \boldsymbol{X}\} = \{u_0, ..., u_{j-1}, u_{j+1}, ..., u_n \sqcup \boldsymbol{X}\}$
   $\vee\ t_0 = u_j \wedge \{t_1, ..., t_m \sqcup \boldsymbol{X}\} = \{u_0, ..., u_n \sqcup \boldsymbol{X}\}$
   $\vee\ \boldsymbol{X} = \{t_0 \sqcup N\} \wedge \{t_1, ..., t_m \sqcup \boldsymbol{X}^{\{X \mapsto N\}}\} = \{u_0, ..., u_n \sqcup \boldsymbol{X}^{\{X \mapsto N\}}\}$
10. $\{t \sqcup A\} = \{u \sqcup B\} \to$
    $t = u \wedge A = B \vee t = u \wedge \{t \sqcup A\} = B \vee t = u \wedge A = \{u \sqcup B\} \vee A = \{u \sqcup N\} \wedge \{t \sqcup N\} = B$
11. $\{\{d \sqcup D\} \mid F \bullet P\} = \emptyset \to \neg F(d) \wedge \{D \mid F \bullet P\} = \emptyset$
12. $\{\{d \sqcup D\} \mid F \bullet P\} = B \to$
    $F(d) \wedge n = P(d) \wedge \{n \sqcup \{D \mid F \bullet P\}\} = B$
    $\vee\ \neg F(d) \wedge \{D \mid F \bullet P\} = B$
13. $\{X \mid F \bullet P\} = \{u_0, u_1, ..., u_n \sqcup \boldsymbol{X}\} \to$ [special case]
    $X = \{n \sqcup N\} \wedge F(n) \wedge u_0 = P(n) \wedge \{N \mid F \bullet P\} = \{u_1, ..., u_n \sqcup \boldsymbol{X}^{\{X \mapsto N\}}\}$
14. $\{X \mid F \bullet P\} = \{t \sqcup A\} \to X = \{d \sqcup N\} \wedge F(d) \wedge t = P(d) \wedge \{N \mid F \bullet P\} = A$
15. *i.* $A = \{\{d \sqcup D\} \mid F \bullet P\} \to \{\{d \sqcup D\} \mid F \bullet P\} = A$ [symmetrical case (rules 11, 12)]
    ii. $\{t \sqcup A\} = \{X \mid F \bullet P\} \to \{X \mid F \bullet P\} = \{t \sqcup A\}$ [symmetrical case (rules 13, 14)]
16. $t_1 = t_2 \to t_1 =_{\mathcal{X}} t_2$

Solved forms

1. $X = A$, A is not a RIS term and X does not occur in A nor in other constraints.
2. $X = \{Y \mid F \bullet P\}$, and X does not occur in other constraints
3. $\{X \mid F \bullet P\} = \emptyset$ or $\emptyset = \{X \mid F \bullet P\}$
4. $\{X \mid F \bullet P\} = \{Y \mid G \bullet Q\}$

3 Inequality ($\neq$)

1. $\emptyset \neq \emptyset \rightarrow false$
2. $X \neq X \rightarrow false$
3. If A is neither a variable nor a RIS [symmetrical case (rules 4, 5)]
   $A \neq X \rightarrow X \neq A$
4. If $X \notin vars(t_1, ..., t_n)$: [special case]
   $X \neq \{t_1, ..., t_n \sqcup \boldsymbol{X}\} \rightarrow (t_1 \notin X \vee ... \vee t_n \notin X)$
5. If $X \in vars(t_1, ..., t_n)$:
   $X \neq \{t_1, ..., t_n \sqcup A\} \rightarrow true$
6. $\emptyset \neq \{t \sqcup A\} \rightarrow true$
7. $\{t \sqcup A\} \neq \emptyset \rightarrow true$ [symmetrical case (rule 6)]
8. $\{t \sqcup A\} \neq \{u \sqcup B\} \rightarrow$
   $n \in \{t \sqcup A\} \wedge n \notin \{u \sqcup B\}$
   $\vee\ n \notin \{t \sqcup A\} \wedge n \in \{u \sqcup B\}$
9. $\{D \mid F \bullet P\} \neq A \rightarrow$
   $n \in \{D \mid F \bullet P\} \wedge n \notin A$
   $\vee\ n \notin \{D \mid F \bullet P\} \wedge n \in A$
10. $A \neq \{D \mid F \bullet P\} \rightarrow \{D \mid F \bullet P\} \neq A$ [symmetrical case (rule 9)]
11. $t_1 \neq t_2 \rightarrow t_1 \neq_{\mathcal{X}} t_2$

Solved forms

1. $X \neq \emptyset$
2. $X \neq \{t \sqcup A\}$, and X does not occur in $\{t \sqcup A\}$ nor as the domain of a RIS which is the argument of a $=$ or $\notin$ or *un* constraint in the input formula.

4 Set membership ($\in$)

1. $t \in \emptyset \to \mathit{false}$
2. $t \in \{u \sqcup A\} \to t = u \vee t \in A$
3. $t \in X \to X = \{t \sqcup N\}$
4. $t \in \{D \mid F \bullet P\} \to d \in D \wedge F(d) \wedge t = P(d)$

Solved forms
none

5 Not set membership ($\notin$)

1. $t \notin \emptyset \to \mathit{true}$
2. $t \notin \{u \sqcup A\} \to t \neq u \wedge t \notin A$
3. If $X \in \mathit{vars}(t)$:
   $t \notin X \to \mathit{true}$
4. $t \notin \{\{d \sqcup D\} \mid F \bullet P\} \to$
   $F(d) \wedge n = P(d) \wedge t \neq n \wedge t \notin \{D \mid F \bullet P\}$
   $\vee \neg F(d) \wedge t \notin \{D \mid F \bullet P\}$

Solved forms

1. $t \notin X$, and X does not occur in t
2. $t \notin \{X \mid F \bullet P\}$

6 Disjointness ($\parallel$)

1. $X \parallel X \rightarrow X = \emptyset$
2. $A \parallel \emptyset \rightarrow true$
3. $\emptyset \parallel A \rightarrow true$ [symmetrical case (rule 2)]
4. $A \parallel \{t \sqcup B\} \rightarrow t \notin A \wedge A \parallel B$
5. $\{t \sqcup B\} \parallel A \rightarrow t \notin A \wedge A \parallel B$ [symmetrical case (rule 4)]
6. $A \parallel \{\{d \sqcup D\} \mid F \bullet P\} \rightarrow$
   $F(d) \wedge n = P(d) \wedge n \notin A \wedge A \parallel \{D \mid F \bullet P\}$
   $\vee \neg F(d) \wedge A \parallel \{D \mid F \bullet P\}$
7. $\{\{d \sqcup D\} \mid F \bullet P\} \parallel A \rightarrow$ [symmetrical case (rule 7)]
   $F(d) \wedge n = P(d) \wedge n \notin A \wedge \{D \mid F \bullet P\} \parallel A$
   $\vee \neg F(d) \wedge \{D \mid F \bullet P\} \parallel A$

Solved forms

1. $\bar{A} \parallel \bar{B}$, and $\bar{A} \not\equiv \bar{B}$ when $\bar{A}$ and $\bar{B}$ are set variables.

7 Union (un)

1. $un(X, X, B) \to X = B$
2. $un(A, B, \emptyset) \to A = \emptyset \wedge B = \emptyset$
3. $un(\emptyset, A, B) \to B = A$
4. $un(A, \emptyset, B) \to B = A$ [symmetrical case (rule 3)]
5. $un(\{t \sqcup C\}, A, \bar{B}) \to$
   $\{t \sqcup C\} = \{t \sqcup N_1\} \wedge \bar{B} = \{t \sqcup N\} \wedge t \notin A \wedge un(N_1, A, N)$
   $\vee \{t \sqcup C\} = \{t \sqcup N_1\} \wedge \bar{B} = \{t \sqcup N\} \wedge A = \{t \sqcup N_2\} \wedge un(N_1, N_2, N)$
6. $un(A, \{t \sqcup C\}, \bar{B}) \to$ [symmetrical case (rule 5)]
   $\{t \sqcup C\} = \{t \sqcup N_1\} \wedge \bar{B} = \{t \sqcup N\} \wedge t \notin A \wedge un(N_1, A, N)$
   $\vee \{t \sqcup C\} = \{t \sqcup N_1\} \wedge \bar{B} = \{t \sqcup N\} \wedge A = \{t \sqcup N_2\} \wedge un(N_1, N_2, N)$
7. $un(A, B, \{t \sqcup C\}) \to$
   $\{t \sqcup C\} = \{t \sqcup N\} \wedge A = \{t \sqcup N_1\} \wedge un(N_1, B, N)$
   $\vee \{t \sqcup C\} = \{t \sqcup N\} \wedge B = \{t \sqcup N_1\} \wedge un(A, N_1, N)$
   $\vee \{t \sqcup C\} = \{t \sqcup N\} \wedge A = \{t \sqcup N_1\} \wedge B = \{t \sqcup N_2\} \wedge un(N_1, N_2, N) \wedge \{t \sqcup C\} = \{t \sqcup N\}$
8. If at least one of A, B, C is not a variable nor a variable-RIS:
   $un(A, B, C) \to$
   $un(T_1(A), T_2(B), T_3(C)) \wedge K_1(A) \wedge K_2(B) \wedge K_3(C)$

   where T_i is a set-valued function and K_i is a constraint-valued function defined as follows:

   - $T_i(S) = S$, if S is either a variable or the empty set or an extensional set term or a variable-RIS
   - $T_i(\{\emptyset \mid F \bullet P\}) = \emptyset$
   - $T_i(\{\{d \sqcup D\} \mid F \bullet P\}) = N_i$
   - $K_i(S) = true$, if S is either a variable or the empty set or an extensional set term or a variable-RIS or a RIS with empty domain
   - $K_i(\{\{d \sqcup D\} \mid F \bullet P\}) =$
     $F(d) \wedge n = P(d) \wedge N_i = \{n \sqcup \{D \mid F \bullet P\}\}$
     $\vee \neg F(d) \wedge N_i = \{D \mid F \bullet P\}$

Solved forms

1. $un(\bar{A}, \bar{B}, \bar{C})$, and $\bar{A} \not\equiv \bar{B}$ when $\bar{A}$ and $\bar{B}$ are set variables.

8 Elimination of $\neq$-constraints

The following rewrite rule is applied to the irreducible formula returned at the end of the rewriting loop performed by the solver, in order to eliminate all inequalities of the form $X \neq A$, where X is the domain of a variable-RIS, still possibly occurring in some constraints of the computed formula.

1. If X is the argument of a *un*-constraint or the domain of a variable-RIS occurring as the argument of a $=$ or $\notin$ or *un* constraint in the input formula:
   $X \neq A \rightarrow (x \in X \wedge x \notin A) \vee (x \in A \wedge x \notin X)$

9 Rule applicability tables

The following tables show in a compact way which rewriting rules are applied for each possible combination of the different constraint arguments.

It turns out that all possible constraint literals are either in solved form or are rewritten by some rewrite rule. On the other hand, each rewrite rule is applied to a constraint literal of some kind.

9.1 $=$ and $\neq$ constraints

$=$	$\emptyset$	$\{b \sqcup B\}$	X_2	$\{\emptyset \mid G \bullet H\}$	$\{\{e \sqcup E\} \mid G \bullet H\}$	$\{\bar{E} \mid G \bullet H\}$
$\emptyset$	1	8	3	1	15	SF3
$\{a \sqcup A\}$	7	10 (9)	3	7	15	15
X_1	6 (SF1)	6 (4, 5, SF2)	6 (2, SF2)	6(SF1)	15	6(SF1)
$\{\emptyset \mid F \bullet P\}$	1	8	3	1	15	SF3
$\{\{d \sqcup D\} \mid F \bullet P\}$	11	12	12	11	12	12
$\{X \mid F \bullet P\}$	SF3	14(13)	3	SF3	15	SF4

$\neq$	$\emptyset$	$\{b \sqcup B\}$	X_2	$\{\emptyset \mid G \bullet H\}$	$\{\{e \sqcup E\} \mid G \bullet H\}$	$\{\bar{E} \mid G \bullet H\}$
$\emptyset$	1	6	3	1	10	10
$\{a \sqcup A\}$	7	8	3	7	10	10
X_1	SF1	5 (4, SF1)	SF1	10	10	10
$\{\emptyset \mid F \bullet P\}$	1	6	9	1	9	9
$\{\{d \sqcup D\} \mid F \bullet P\}$	9	9	9	9	9	9
$\{X \mid F \bullet P\}$	9	9	9	9	9	9

9.2 $\in$ and $\notin$ constraints

$t \in A_1$	$\emptyset$	$\{b \sqcup B\}$	X_2	$\{\emptyset \mid G \bullet H\}$	$\{\{e \sqcup E\} \mid G \bullet H\}$	$\{\bar{E} \mid G \bullet H\}$
t	1	2	3	1	4	4

$t \notin A_1$	$\emptyset$	$\{b \sqcup B\}$	X_2	$\{\emptyset \mid G \bullet H\}$	$\{\{e \sqcup E\} \mid G \bullet H\}$	$\{\bar{E} \mid G \bullet H\}$
t	1	2	SF1 (3)	1	4	SF1

9.3 $\parallel$ constraint

$\parallel$	$\emptyset$	$\{b \sqcup B\}$	X_2	$\{\emptyset \mid G \bullet H\}$	$\{\{e \sqcup E\} \mid G \bullet H\}$	$\{\bar{E} \mid G \bullet H\}$
$\emptyset$	2	3	3	2	3	3
$\{a \sqcup A\}$	2	4	5	2	5	5
X_1	2	4	SF1(1)	2	6	SF1
$\{\emptyset \mid F \bullet P\}$	2	3	3	2	3	3
$\{\{d \sqcup D\} \mid F \bullet P\}$	2	4	7	2	6	7
$\{X \mid F \bullet P\}$	2	4	SF1	2	6	SF1

9.4 *un* constraint

un	all cases
$un(A, B, \emptyset)$	2
$un(A, B, \{\emptyset \mid F \bullet P\})$	2
$un(A, B, \{a \sqcup A\})$	7
$un(A, B, \{\{d \sqcup D\} \mid F \bullet P\})$	8

$un(A_1, A_2, X_3)$	$\emptyset$	$\{b \sqcup B\}$	X_2	$\{\emptyset \mid G \bullet H\}$	$\{\{e \sqcup E\} \mid G \bullet H\}$	$\{\bar{E} \mid G \bullet H\}$
$\emptyset$	3	3	3	3	3	3
$\{a \sqcup A\}$	4	5	5	4	5	5
X_1	4	6	SF1	4	8	SF1
$\{\emptyset \mid F \bullet P\}$	3	3	3	3	3	3
$\{\{d \sqcup D\} \mid F \bullet P\}$	4	6	8	4	8	8
$\{X \mid F \bullet P\}$	4	6	SF1	4	8	SF1

$un(A_1, A_2, \{X' \mid F' \bullet G'\})$	$\emptyset$	$\{b \sqcup B\}$	X_2	$\{\emptyset \mid G \bullet H\}$	$\{\{e \sqcup E\} \mid G \bullet H\}$	$\{\bar{E} \mid G \bullet H\}$
$\emptyset$	3	3	3	3	3	3
$\{a \sqcup A\}$	4	5	5	4	5	5
X_1	4	6	SF1	4	8	SF1
$\{\emptyset \mid F \bullet P\}$	3	3	3	3	3	3
$\{\{d \sqcup D\} \mid F \bullet P\}$	4	6	8	4	8	8
$\{X \mid F \bullet P\}$	4	6	SF1	4	8	SF1