Technical Report CS-01-18

Contract-based Query Rewriting under Privacy Constraints

Hannes Grunert

University of Rostock
Faculty of Computer Science and Electrical Engineering
Institute of Computer Science
Database Research Group

INSTITUT
FÜR
INFORMATIK

Additional material for VLIoT @ VLDB 2018

Contract-based Query Rewriting under Privacy Constraints

Hannes Grunert
Database Research Group
University of Rostock
18051 Rostock
hg(at)informatik.uni-rostock.de

1 Introduction and Motivation

Smart environments consist of multiple, heterogeneous devices. While powerful servers can handle any type of Big Data analysis, small devices can only handle a small subset of operators. Our approach, Query Rewriting by Contract, takes a complex query and finds fragments of the query, which can be executed on such restricted devices. This leads to data minimization in terms of privacy, but without altering the actual result of the analysis.

As a us case, imagine that your living room is equipped with an audio system and a smart assistant by *Provider A*. Additionally, your smartphone has its own assistant, e. g. the service offered by *Provider G*. Now, imagine you are driving back from work and you want to hear your favorite music at home. You pull out your phone and ask *Provider G* to do the following:

"OK G, tell A to play the music of the componist, which shortest track on an album is shorter than 15 seconds and which longest track on the same album is longer than 10 minutes. The name of the componist should not be empty."

G records your speech, analyzes its on locally on your phone and sends the detected words to *G*'s Cloud Center. In the Cloud, *G* generates the following SQL query, together with general instructions for *A*:

```
SELECT Composers.name AS Composer, Albums.name AS Album
FROM Composers JOIN Tracks
ON(Composers.id=Tracks.composer)
JOIN Albums ON(Albums.id=Tracks.album)
WHERE Composers.name !=  ''
GROUP BY Albums.id, Composers.name, Albums.name
HAVING min(length) < 15*1000
AND max(length) >= 10*60*1000
```

G sends this to your home, where *A* is waiting. She retrieves ans interprets *G*'s instructions, generates her own instructions for *A*'s Cloud Center. *A* checks the query against your own music archive and finds out, which music you want to hear. The result is returned to your smart home, where your audio systems waits until you arrive back home and starts playing your favorite songs.

In the next section, we give an overview on our rule set for Query Rewriting by Contract. The performance measurements for each rule are given in Section 3. We will come back to the example above in Section 4 and show, how these rules can be applied on a larger qery.

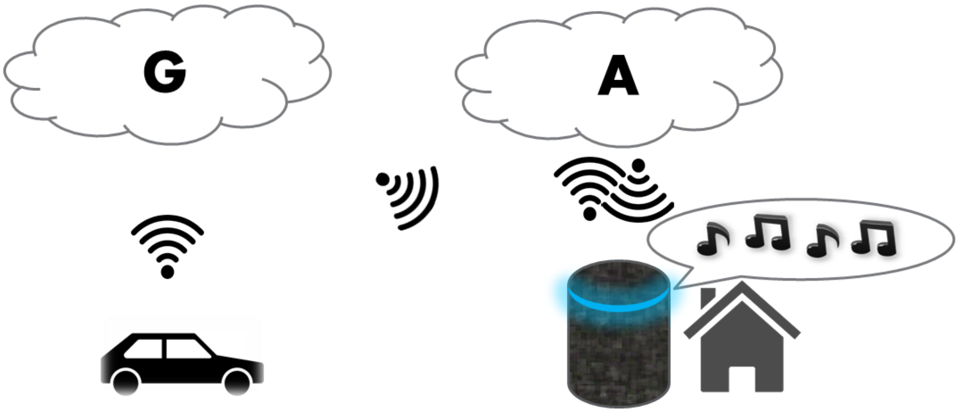


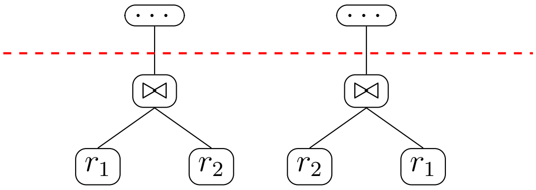
Figure 1: Usage of a smart environment

2 Rule set

2.1 Classic Rules

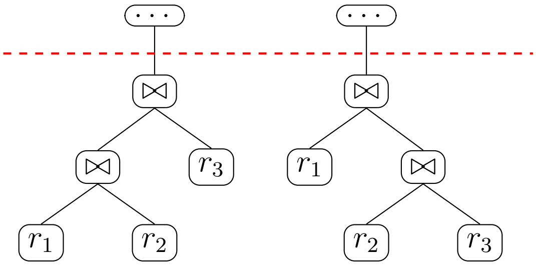
K0-1: Commutativity of joins

Rule: $r_1 \bowtie r_2 \equiv r_2 \bowtie r_1$
Invariants: —
Preconditions: —
Postconditions: —
Remark: Only used for internal swapping.



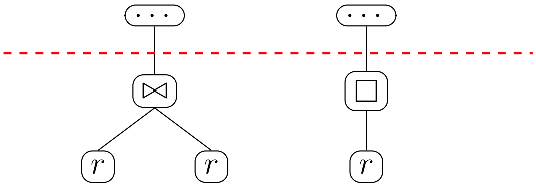
K0-2: Associativity of joins

Rule: $(r_1 \bowtie r_2) \bowtie r_3 \equiv r_1 \bowtie (r_2 \bowtie r_3)$
Invariants: —
Preconditions: —
Postconditions: —
Remark: Only used for internal swapping.



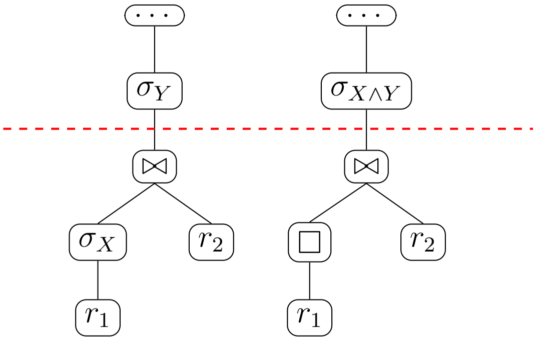
K01: Elimination of redundant operations

Rule: $r \bowtie r \equiv r$
Invariants: —
Preconditions: $\bowtie$ is not supported
Postconditions: —



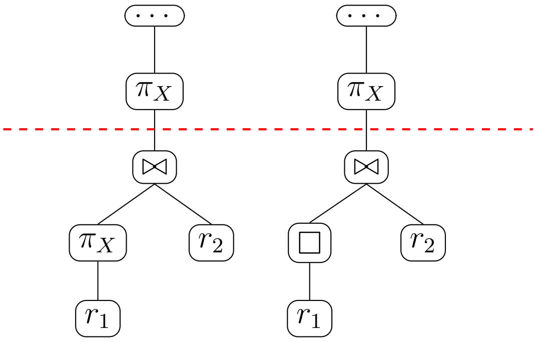
K02: Commutativity of selection and join

Rule: $\sigma_Y(\sigma_X(r_1) \bowtie r_2) \sqsubseteq_K \sigma_{X \wedge Y}(r_1 \bowtie r_2)$
Invariants: $attr(X) \subseteq r_1$
Preconditions: σ_X is not supported
Postconditions: —



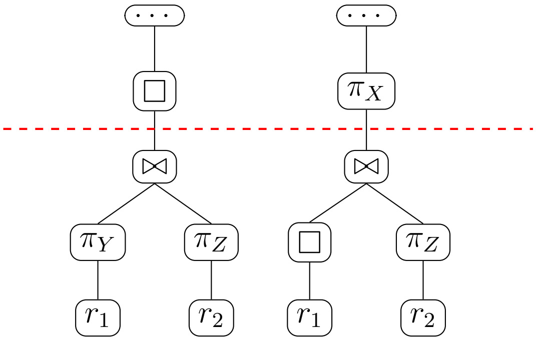
K03: Commutativity of projection and join – V1

Rule: $\pi_X(\pi_X(r_1) \bowtie r_2) \sqsubseteq_K \pi_X(r_1 \bowtie r_2)$
Invariants: $attr(X) \subseteq r_1$
Preconditions: π_X is not supported
Postconditions: —



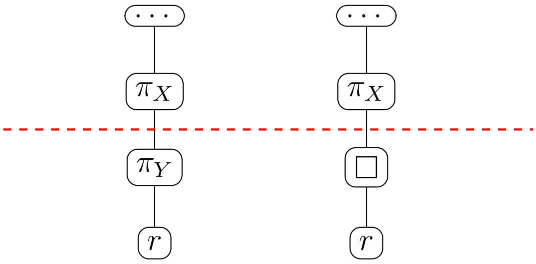
K04: Commutativity of projection and join – V2

Rule: $\pi_Y(r_1) \bowtie \pi_Z(r_2) \sqsubseteq_K \pi_X(r_1 \bowtie \pi_Z(r_2))$
Invariants: $X := Y \cup Z$
Preconditions: π_Y is not supported,
$Y \cap Z \neq \emptyset \vee R(r_1) \cap R(r_2) = \emptyset$
Postconditions: —



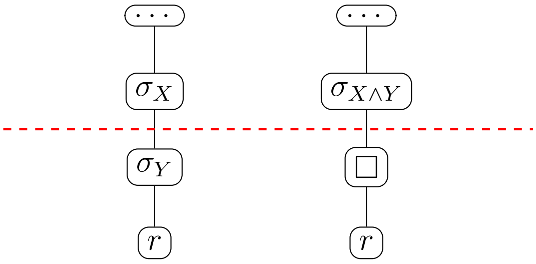
K05: Domination of the outer projection

Rule: $\pi_X(\pi_Y(r)) \sqsubseteq_K \pi_X(r)$
Invariants: $X \subseteq Y$
Preconditions: π_Y is not supported
Postconditions: —



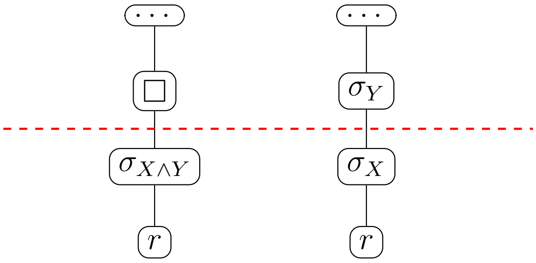
K06: Commutativity of two selections – V1

Rule: $\sigma_X(\sigma_Y(r)) \sqsubseteq_K \sigma_{X \wedge Y}(r)$
Invariants: —
Preconditions: σ_Y is not supported
Postconditions: —



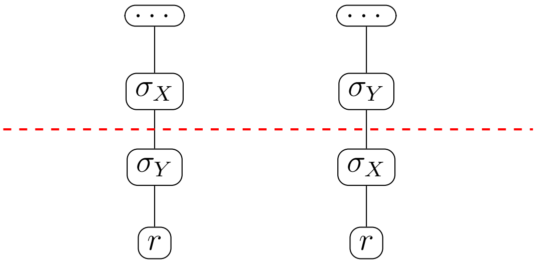
K07: Commutativity of two selections – V2

Rule: $\sigma_{X \wedge Y}(r) \sqsubseteq_K \sigma_Y(\sigma_X(r))$
Invariants: —
Preconditions: σ_Y or $\wedge$ are not supported
Postconditions: σ_X is supported

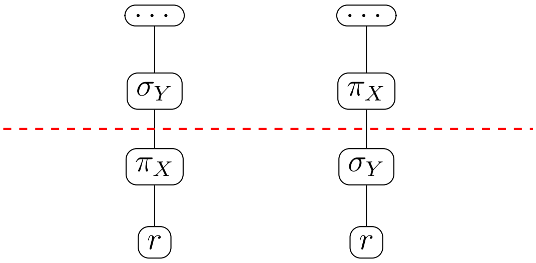


K08: Commutativity of selections – V3

Rule: $\sigma_X(\sigma_Y(r)) \sqsubseteq_K \sigma_Y(\sigma_X(r))$
Invariants: —
Preconditions: σ_Y is not supported
Postconditions: σ_X is supported

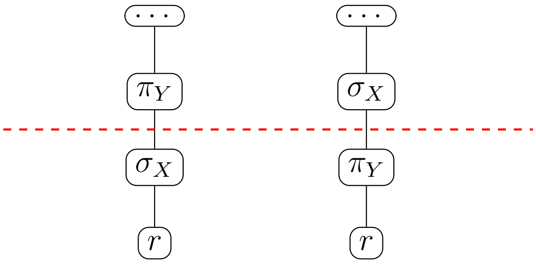


K09: Commutativity of selection and projection – V1



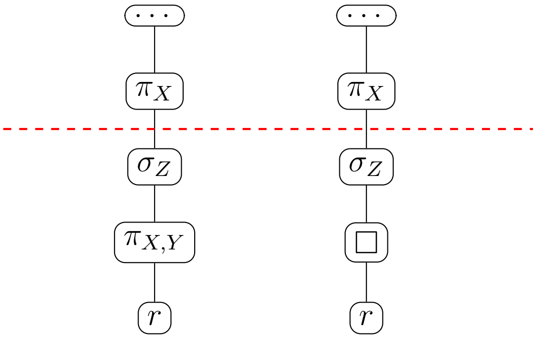
Rule: $\sigma_Y(\pi_X(r)) \sqsubseteq_K \pi_X(\sigma_Y(r))$
Invariants: —
Preconditions: π_X is not supported
Postconditions: σ_Y is supported

K10: Commutativity of selection and projection – V2



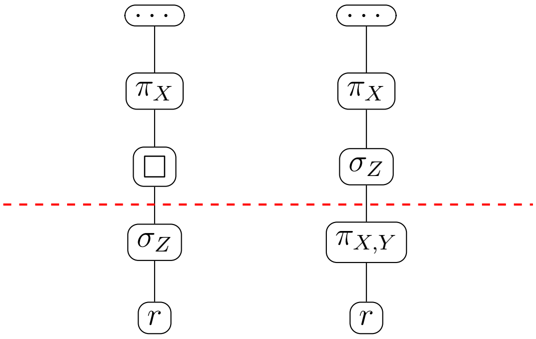
Rule: $\pi_Y(\sigma_X(r)) \sqsubseteq_K \sigma_X(\pi_Y(r))$
Invariants: —
Preconditions: σ_X is not supported,
$attr(X) \subseteq Y$
Postconditions: π_Y is supported

K11: Commutativity of selection and projection – V3



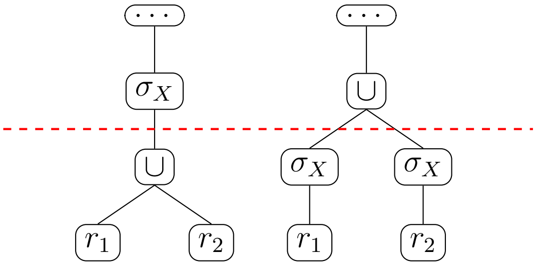
Rule: $\pi_X(\sigma_Z(\pi_{X,Y}(r))) \sqsubseteq_K \pi_X(\sigma_Z(r))$
Invariants: —
Preconditions: $\pi_{X,Y}$ is not supported,
$attr(Z) \subseteq X \cup Y$
Postconditions: —

K12: Commutativity of selection and projection – V4



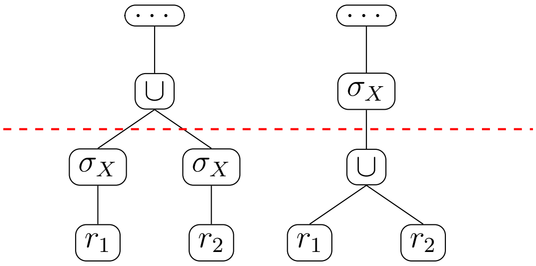
Rule: $\pi_X(\sigma_Z(r)) \sqsubseteq_K \pi_X(\sigma_Z(\pi_{X,Y}(r)))$
Invariants: —
Preconditions: σ_Z is not supported
Postconditions: $\pi_{X,Y}$ is supported,
$Y := attr(Z) \backslash X$

K13: Commutativity of selection and set union – V1



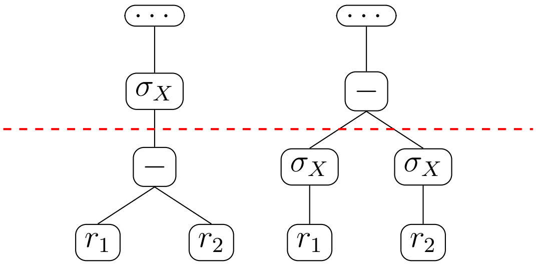
Rule: $\sigma_X(r_1 \cup r_2) \sqsubseteq_K \sigma_X(r_1) \cup \sigma_X(r_2)$
Invariants: —
Preconditions: $\cup$ is not supported
Postconditions: σ_X is supported

K14: Commutativity of selection and set union – V2



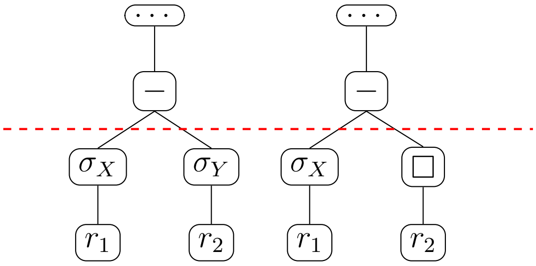
Rule: $\sigma_X(r_1) \cup \sigma_X(r_2) \sqsubseteq_K \sigma_X(r_1 \cup r_2)$
Invariants: —
Preconditions: σ_X is not supported
Postconditions: $\cup$ is supported

K15: Commutativity of selection and set difference – V1



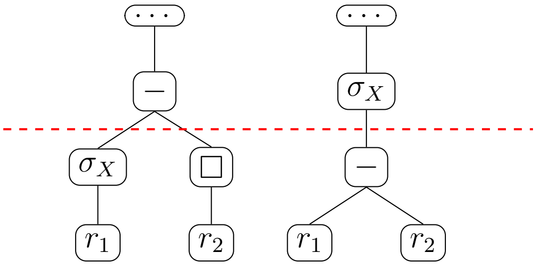
Rule: $\sigma_X(r_1 - r_2) \sqsubseteq_K \sigma_X(r_1) - r_2$
Invariants: —
Preconditions: $-$ is not supported
Postconditions: σ_X is supported

K16: Commutativity of selection and set difference – V2



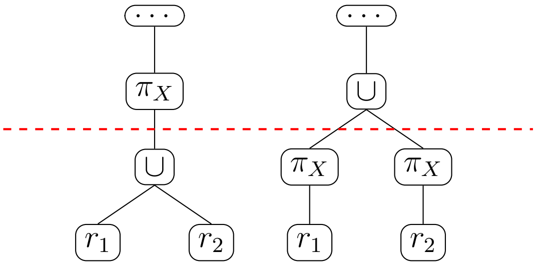
Rule: $\sigma_X(r_1) - \sigma_X(r_2) \sqsubseteq_K \sigma_X(r_1) - r_2$
Invariants: —
Preconditions: σ_X is not supported in the right subtree
Postconditions: —

K17: Commutativity of selection and set difference – V3



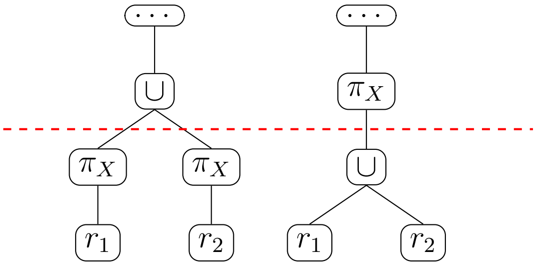
Rule: $\sigma_X(r_1) - r_2 \sqsubseteq_K \sigma_X(r_1 - r_2)$
Invariants: —
Preconditions: σ_X is not supported
Postconditions: $-$ is supported

K18: Commutativity of projection and set union – V1



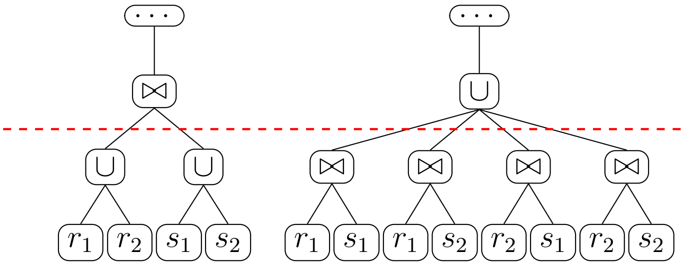
Rule: $\pi_X(r_1 \cup r_2) \sqsubseteq_K \pi_X(r_1) \cup \pi_X(r_2)$
Invariants: —
Preconditions: $\cup$ is not supported
Postconditions: π_X is supported

K19: Commutativity of projection and set union – V2



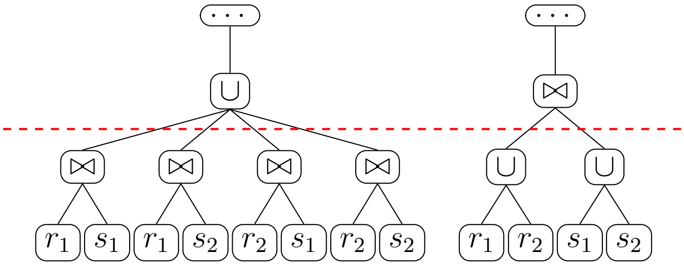
Rule: $\pi_X(r_1) \cup \pi_X(r_2) \sqsubseteq_K \pi_X(r_1 \cup r_2)$
Invariants: —
Preconditions: π_X is not supported
Postconditions: $\cup$ is supported, $R(r_1) \equiv R(r_2)$

K20: Distributivity of set union and join – V1



Rule: $(R_1 \cup R_2) \bowtie (S_1 \cup S_2) \sqsubseteq_K (R_1 \bowtie S_1) \cup (R_1 \bowtie S_2) \cup (R_2 \bowtie S_1) \cup (R_2 \bowtie S_2)$
Invariants: —
Preconditions: $\cup$ is not supported
Postconditions: $\bowtie$ is supported

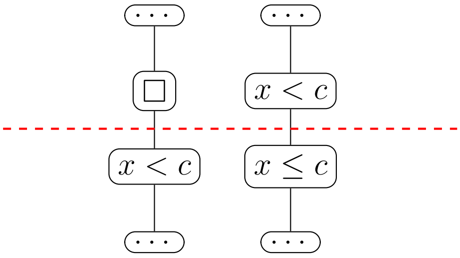
K21: Distributivity of set union and join – V2



Rule: $(r_1 \bowtie s_1) \cup (r_1 \bowtie s_2) \cup (r_2 \bowtie s_1) \cup (r_2 \bowtie s_2) \sqsubseteq_K (r_1 \cup r_2) \bowtie (s_1 \cup s_2)$
Invariants: —
Preconditions: $\bowtie$ is not supported
Postconditions: $\cup$ is supported

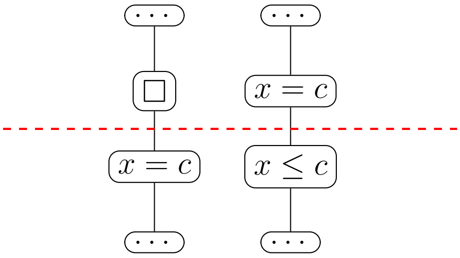
2.2 LAC1 + LAC2

L01: less-than → less-than-or-equal-to (attribute-constant)



Rule: $x < c \sqsubseteq_K x \leq c$
Invariants: —
Preconditions: $<$ is not supported
Postconditions: $\leq$ is supported

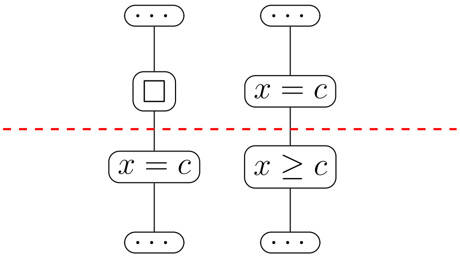
L02: equal-to → less-than-or-equal-to (attribute-constant)



Rule: $x = c \sqsubseteq_K x \leq c$
Invariants: —
Preconditions: $=$ is not supported
Postconditions: $\leq$ is supported

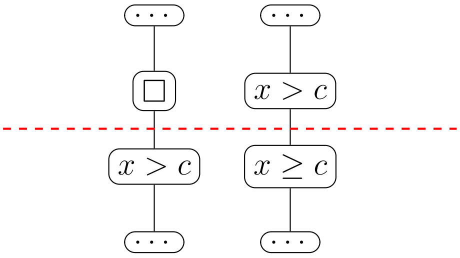
L03: equal-to → greater-than-or-equal-to (attribute-constant)

Rule: $x = c \sqsubseteq_K x \geq c$
Invariants: —
Preconditions: $=$ is not supported
Postconditions: $\geq$ is supported



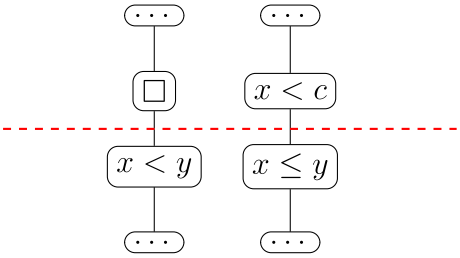
L04: greater-than → greater-than-or-equal-to (attribute-constant)

Rule: $x > c \sqsubseteq_K x \geq c$
Preconditions: $>$ is not supported
Postconditions: $\geq$ is supported



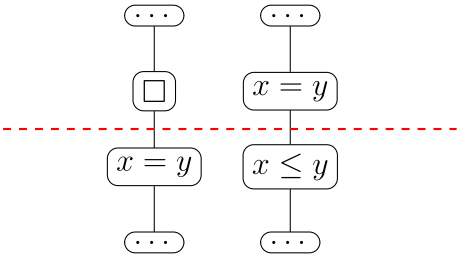
L05: less-than → less-than-or-equal-to (attribute-attribute)

Rule: $x < y \sqsubseteq_K x \leq y$
Invariants: —
Preconditions: $<$ is not supported
Postconditions: $\leq$ is supported



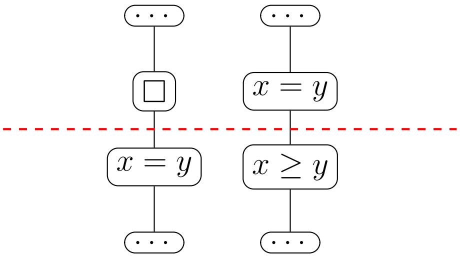
L06: equal-to → less-than-or-equal-to (attribute-attribute)

Rule: $x = y \sqsubseteq_K x \leq y$
Invariants: —
Preconditions: $=$ is not supported
Postconditions: $\leq$ is supported



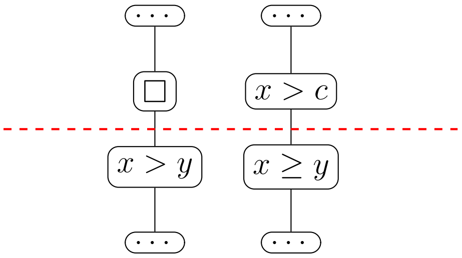
L07: equal-to → greater-than-or-equal-to (attribute-attribute)

Rule: $x = y \sqsubseteq_K x \geq y$
Invariants: —
Preconditions: $=$ is not supported
Postconditions: $\geq$ is supported



L08: greater-than → greater-than-or-equal-to (attribute-attribute)

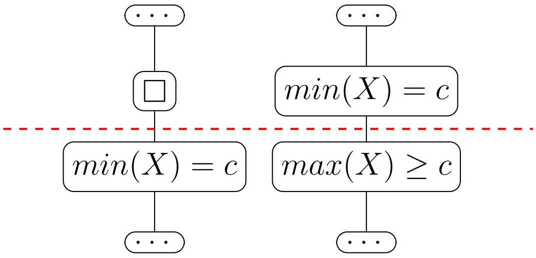
Rule: $x > y \sqsubseteq_K x \geq y$
Invariants: —
Preconditions: $>$ is not supported
Postconditions: $\geq$ is supported



2.3 Aggregate Constraints

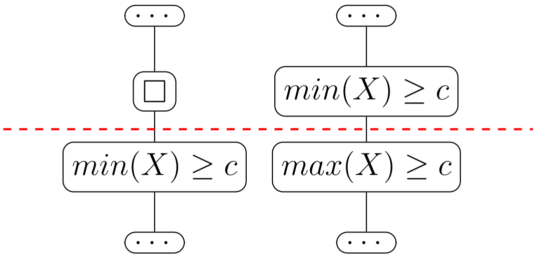
A01: minimum-equals → maximum-greater-than-or-equals-to (attribute-constant)

Rule: $min(X) = c \sqsubseteq_K max(X) \geq c$
Invariants: —
Preconditions: min or $=$ are not supported
Postconditions: max und $\geq$



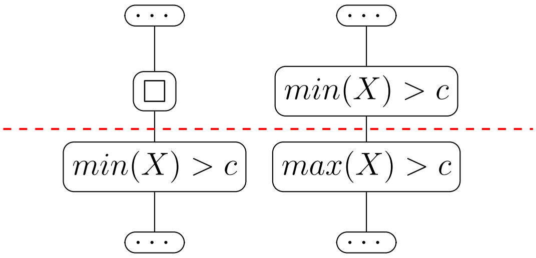
A02: minimum-greater-than-or-equals-to → maximum-greater-than-or-equals-to (attribute-constant)

Rule: $min(X) \geq c \sqsubseteq_K max(X) \geq c$
Invariants: —
Preconditions: min is not supported
Postconditions: max is supported



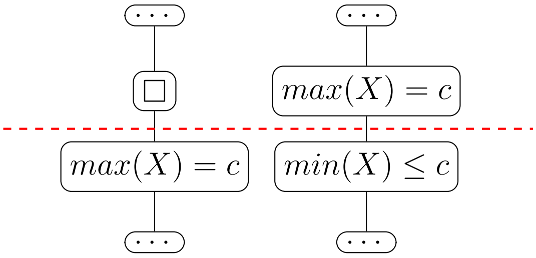
A03: minimum-greater-than → maximum-greater-than (attribute-constant)

Rule: $min(X) > c \sqsubseteq_K max(X) > c$
Invariants: —
Preconditions: min is not supported
Postconditions: max is supported



A04: maximum-equals → minimum-less-than-or-equals-to (attribute-constant)

Rule: $max(X) = c \sqsubseteq_K min(X) \leq c$
Invariants: —
Preconditions: max or $=$ are not supported
Postconditions: min und $\leq$ is supported



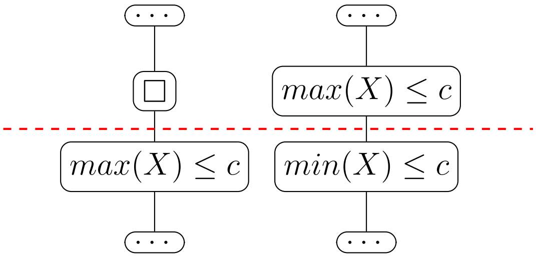
A05: maximum-less-than-or-equals-to → minimum-less-than-or-equals-to (attribute-constant)

Rule: $max(X) \leq c \sqsubseteq_K min(X) \leq c$

Invariants: —

Preconditions: max is not supported

Postconditions: min is supported



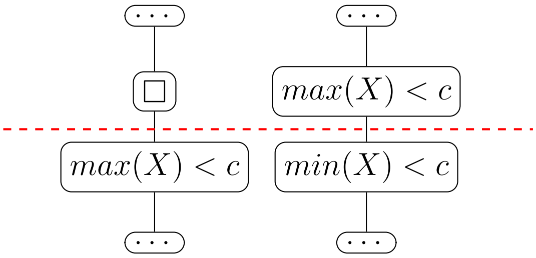
A06: maximum-less-than → minimum-less-than (attribute-constant)

Rule: $max(X) < c \sqsubseteq_K min(X) < c$

Invariants: —

Preconditions: max is not supported

Postconditions: min is supported



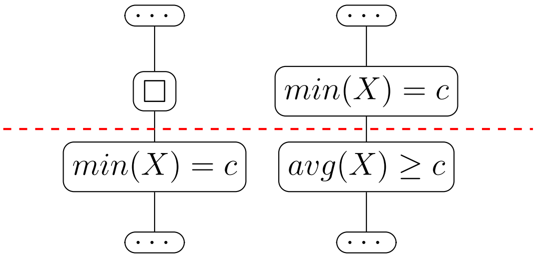
A07: minimum-equals → average-greater-than-or-equals-to (attribute-constant)

Rule: $min(X) = c \sqsubseteq_K avg(X) \geq c$

Invariants: —

Preconditions: min or $=$ are not supported

Postconditions: avg und $\geq$ is supported



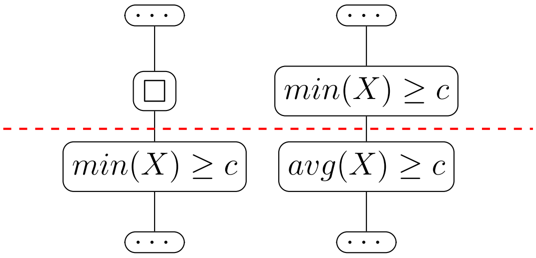
A08: minimum-greater-than-or-equals-to → average-greater-than-or-equals-to (attribute-constant)

Rule: $min(X) \geq c \sqsubseteq_K avg(X) \geq c$

Invariants: —

Preconditions: min is not supported

Postconditions: avg is supported



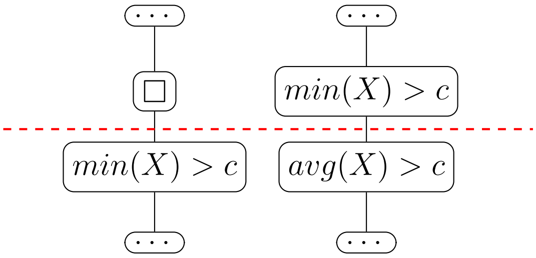
A09: minimum-greater-than → average-less-than (attribute-constant)

Rule: $min(X) > c \sqsubseteq_K avg(X) > c$

Invariants: —

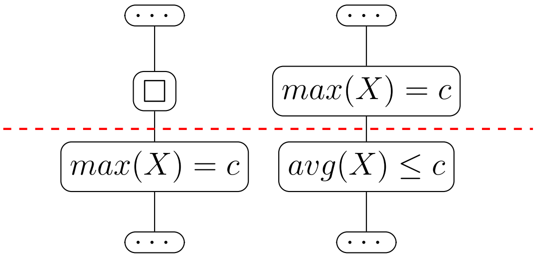
Preconditions: min is not supported

Postconditions: avg und $>$ is supported



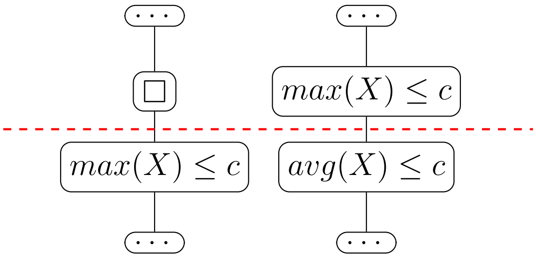
A10: maximum-equals → average-less-than-or-equals-to (attribute-constant)

Rule: $max(X) = c \sqsubseteq_K avg(X) \leq c$
Invariants: —
Preconditions: max or $=$ are not supported
Postconditions: avg und $\leq$ is supported



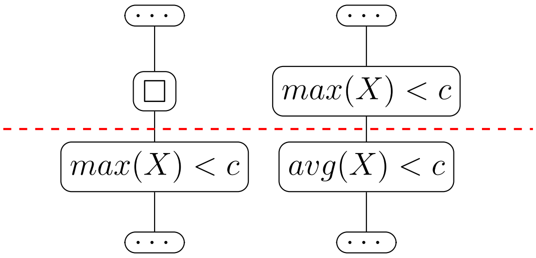
A11: maximum-less-than-or-equals-to → average-less-than-or-equals-to (attribute-constant)

Rule: $max(X) \leq c \sqsubseteq_K avg(X) \leq c$
Invariants: —
Preconditions: max is not supported
Postconditions: avg is supported



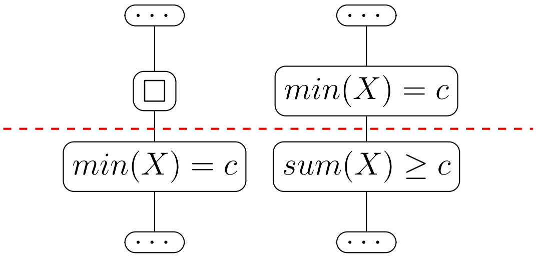
A12: maximum-less-than → average-less-than (attribute-constant)

Rule: $max(X) < c \sqsubseteq_K avg(X) < c$
Invariants: —
Preconditions: max is not supported
Postconditions: avg is supported



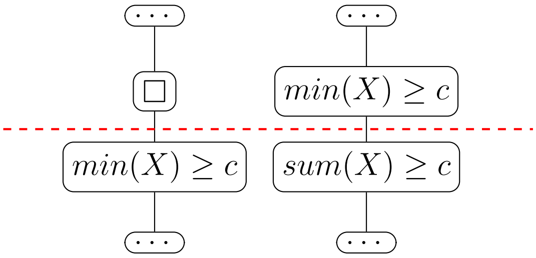
A13: minimum-equals → sum-greater-than-or-equals-to (attribute-constant)

Rule: $min(X) = c \sqsubseteq_K sum(X) \geq c$
Invariants: $c \geq 0$
Preconditions: min or $=$ are not supported
Postconditions: sum und $\geq$ is supported



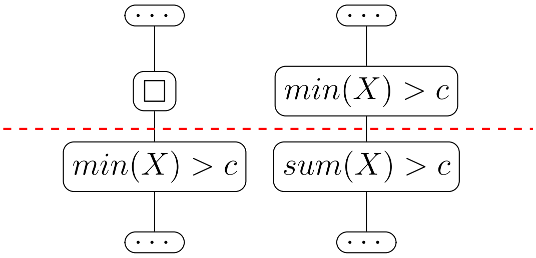
A14: minimum-greater-than-or-equals-to → sum-greater-than-or-equals-to (attribute-constant)

Rule: $min(X) \geq c \sqsubseteq_K sum(X) \geq c$
Invariants: $c \geq 0$
Preconditions: min is not supported
Postconditions: sum is supported



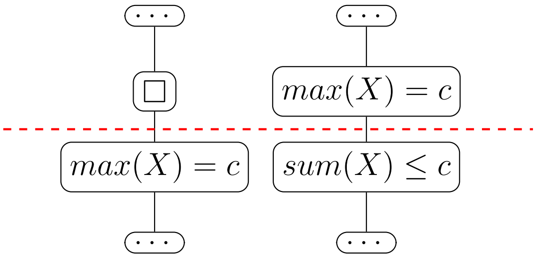
A15: minimum-less-than → sum-greater-than (attribute-constant)

Rule: $\mathit{min}(X) > c \sqsubseteq_K \mathit{sum}(X) > c$
Invariants: $c \geq 0$
Preconditions: *min* is not supported
Postconditions: *sum* is supported



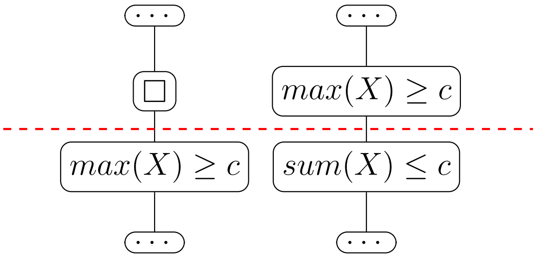
A16: maximum-equals → sum-less-than-or-equals-to (attribute-constant)

Rule: $\mathit{max}(X) = c \sqsubseteq_K \mathit{sum}(X) \leq c$
Invariants: $c \leq 0$
Preconditions: *max* or $=$ are not supported
Postconditions: *sum* und $\leq$ is supported



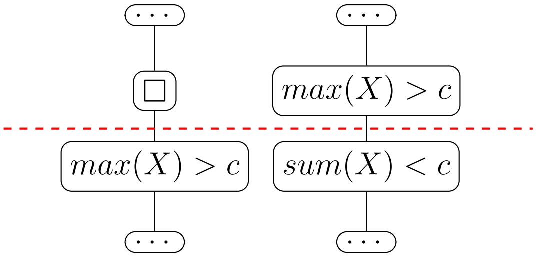
A17: maximum-greater-than-or-equals-to → sum-less-than-or-equals-to (attribute-constant)

Rule: $\mathit{max}(X) \geq c \sqsubseteq_K \mathit{sum}(X) \leq c$
Invariants: $c \leq 0$
Preconditions: *max* or $\geq$ are not supported
Postconditions: *sum* und $\leq$ is supported



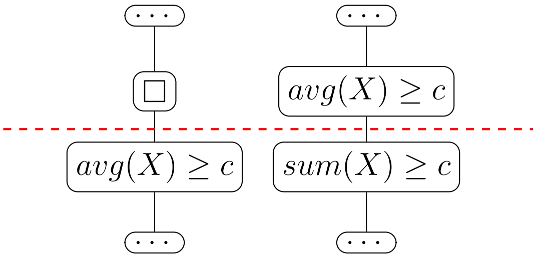
A18: maximum-greater-than → sum-less-than (attribute-constant)

Rule: $\mathit{max}(X) > c \sqsubseteq_K \mathit{sum}(X) < c$
Invariants: $c \leq 0$
Preconditions: *max* or $>$ are not supported
Postconditions: *sum* und $<$ is supported



A19: average-greater-than-or-equals-to → sum-greater-than-or-equals-to (attribute-constant)

Rule: $\mathit{avg}(X) \geq c \sqsubseteq_K \mathit{sum}(X) \geq c$
Invariants: $c \geq 0$
Preconditions: *avg* is not supported
Postconditions: *sum* is supported



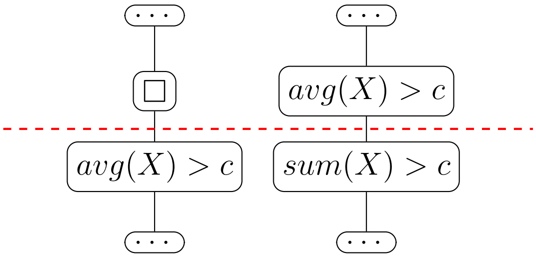
A20: average-less-than → sum-greater-than (attribute-constant)

Rule: $avg(X) > c \sqsubseteq_K sum(X) > c$

Invariants: $c \geq 0$

Preconditions: *avg* is not supported

Postconditions: *sum* is supported



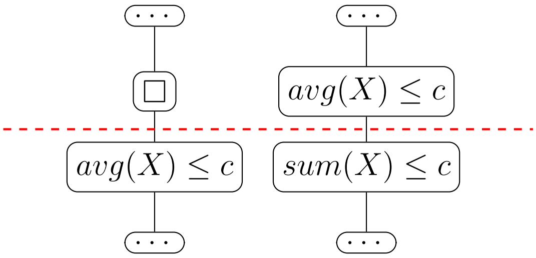
A21: average-less-than-or-equals-to → sum-less-than-or-equals-to (attribute-constant)

Rule: $avg(X) \leq c \sqsubseteq_K sum(X) \leq c$

Invariants: $c \leq 0$

Preconditions: *avg* is not supported

Postconditions: *sum* is supported



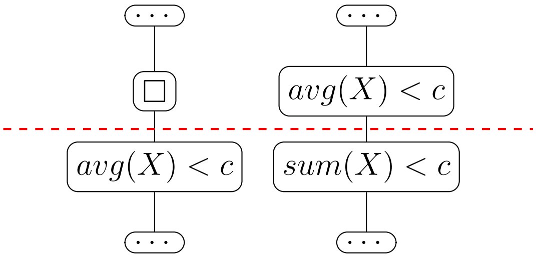
A22: average-less-than → sum-less-than (attribute-constant)

Rule: $avg(X) < c \sqsubseteq_K sum(X) < c$

Invariants: $c \leq 0$

Preconditions: *avg* is not supported

Postconditions: *sum* is supported



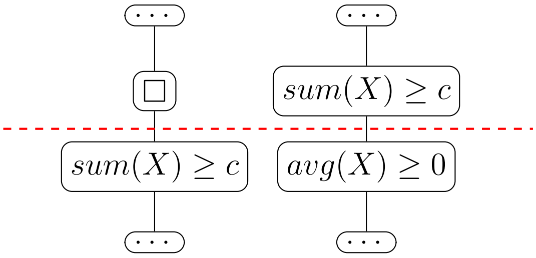
A23: sum-greater-than-or-equals-to → average-greater-than-or-equals-to-0 (attribute-constant)

Rule: $sum(X) \geq c \sqsubseteq_K avg(X) \geq 0$

Invariants: $c \geq 0$

Preconditions: *sum* is not supported

Postconditions: *avg* is supported



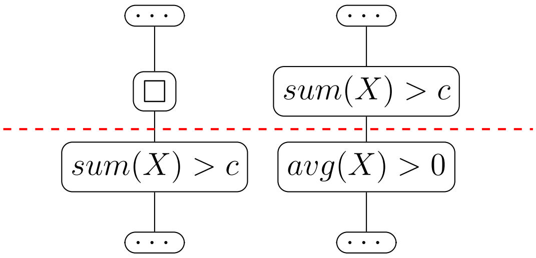
A24: sum-greater-than → average-less-than-0 (attribute-constant)

Rule: $sum(X) > c \sqsubseteq_K avg(X) > 0$

Invariants: $c \geq 0$

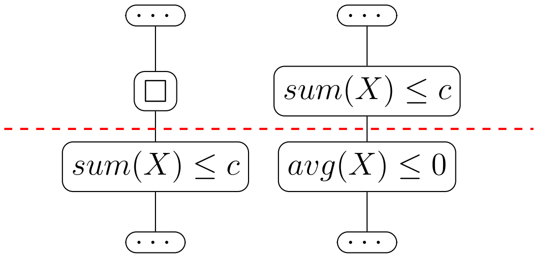
Preconditions: *sum* is not supported

Postconditions: *avg* is supported



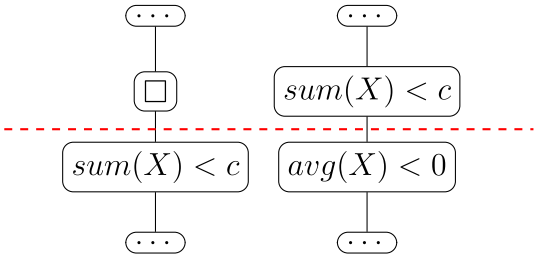
A25: sum-less-than-or-equals-to → average-less-than-or-equals-to-0 (attribute-constant)

Rule: $sum(X) \leq c \sqsubseteq_K avg(X) \leq 0$
Invariants: $c \leq 0$
Preconditions: sum is not supported
Postconditions: avg is supported



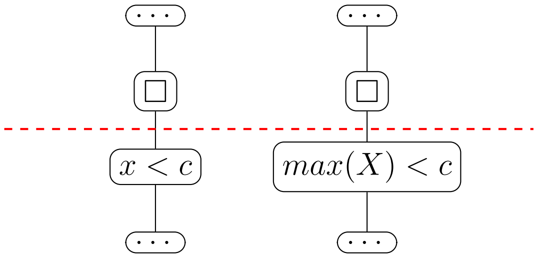
A26: sum-less-than → average-less-than-0 (attribute-constant)

Rule: $sum(X) < c \sqsubseteq_K avg(X) < 0$
Invariants: $c \leq 0$
Preconditions: sum is not supported
Postconditions: avg is supported



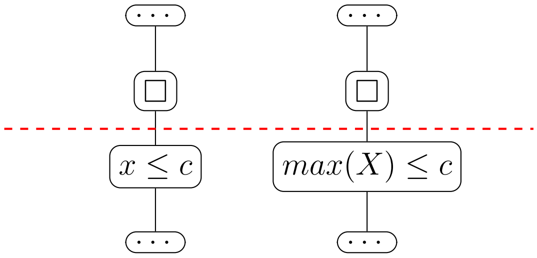
A27: less-than → maximum-less-than (attribute-constant)

Rule: $\forall x \in X : x < c \sqsubseteq_K max(X) < c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: max is supported



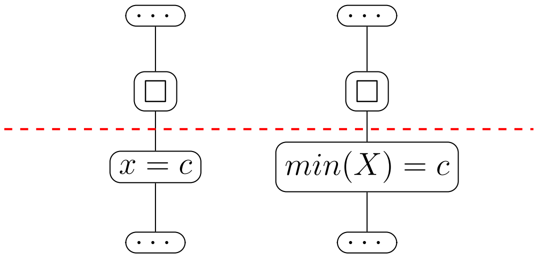
A28: less-than-or-equals-to → maximum-less-than-or-equals-to (attribute-constant)

Rule: $\forall x \in X : x \leq c \sqsubseteq_K max(X) \leq c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: max is supported



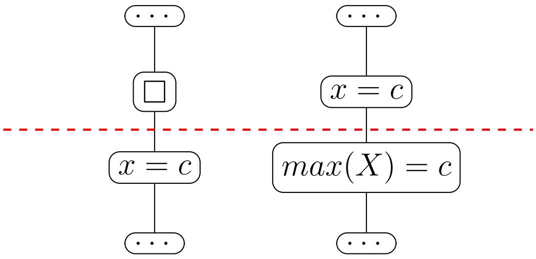
A29: equals → minimum-equals (attribute-constant)

Rule: $\forall x \in X : x = c \sqsubseteq_K min(X) = c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: min is supported



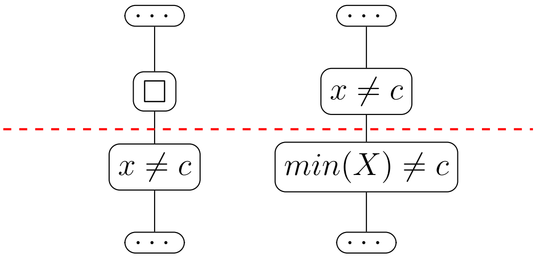
A30: equals → maximum-equals (attribute-constant)

Rule: $\forall x \in X : x = c \sqsubseteq_K max(X) = c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *max* is supported



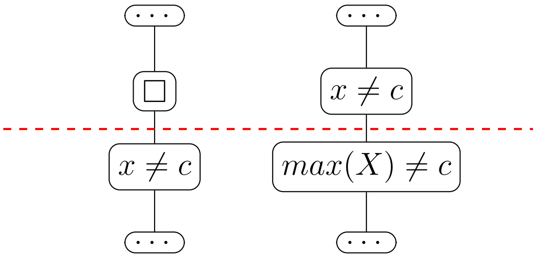
A31: unequals → Minimum-unequals (attribute-constant)

Rule: $\forall x \in X : x \neq c \sqsubseteq_K min(X) \neq c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *min* is supported



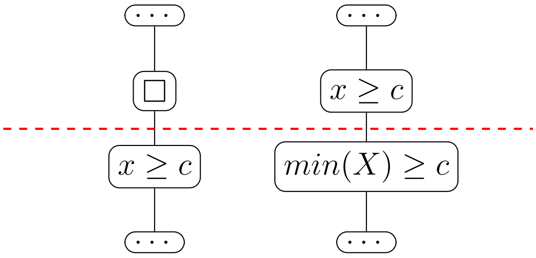
A32: unequals → Maximum-unequals (attribute-constant)

Rule: $\forall x \in X : x \neq c \sqsubseteq_K max(X) \neq c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *max* is supported



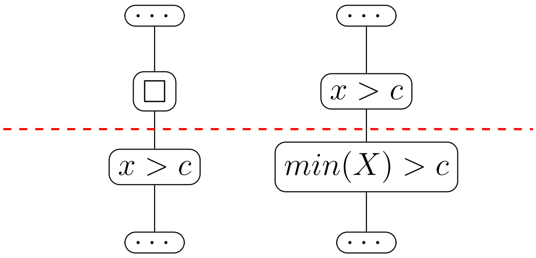
A33: greater-than-or-equals-to → minimum-greater-than-or-equals-to (attribute-constant)

Rule: $\forall x \in X : x \geq c \sqsubseteq_K min(X) \geq c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *min* is supported



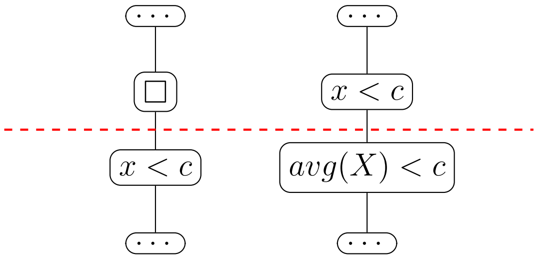
A34: greater-than → minimum-greater-than (attribute-constant)

Rule: $\forall x \in X : x > c \sqsubseteq_K min(X) > c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *min* is supported



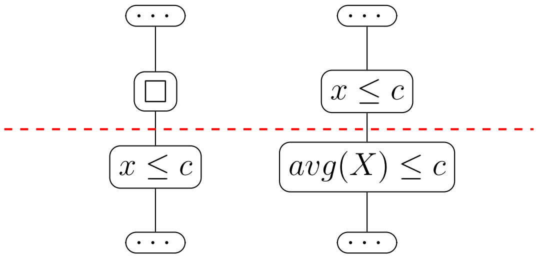
A35: less-than → average-less-than (attribute-constant)

Rule: $\forall x \in X : x < c \sqsubseteq_K avg(X) < c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *avg* is supported



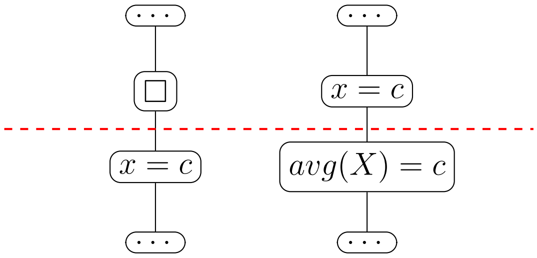
A36: less-than-or-equals-to → average-less-than-or-equals-to (attribute-constant)

Rule: $\forall x \in X : x \leq c \sqsubseteq_K avg(X) \leq c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *avg* is supported



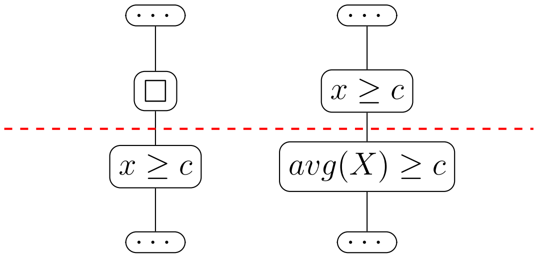
A37: equals → average-equals (attribute-constant)

Rule: $\forall x \in X : x = c \sqsubseteq_K avg(X) = c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *avg* is supported



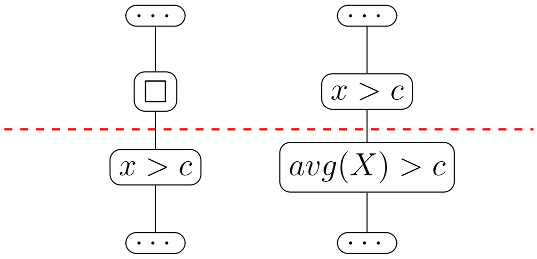
A38: greater-than-or-equals-to → average-greater-than-or-equals-to (attribute-constant)

Rule: $\forall x \in X : x \geq c \sqsubseteq_K avg(X) \geq c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *avg* is supported



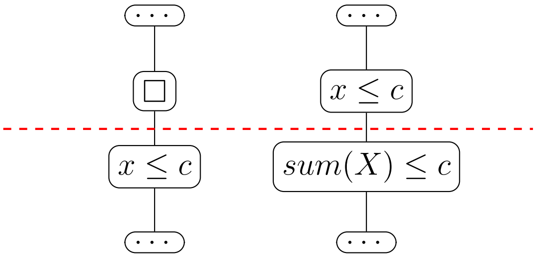
A39: greater-than → average-less-than (attribute-constant)

Rule: $\forall x \in X : x > c \sqsubseteq_K avg(X) > c$
Invariants: —
Preconditions: $\forall$ is not supported
Postconditions: *avg* is supported



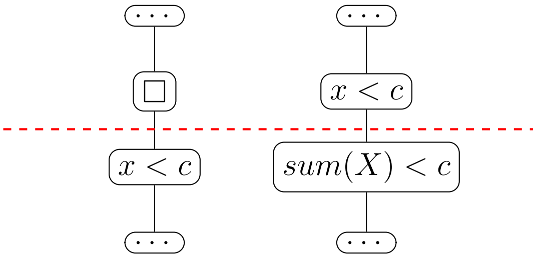
A40: less-than-or-equals-to → sum-less-than-or-equals-to (attribute-constant)

Rule: $\forall x \in X : x \leq c \sqsubseteq_K sum(X) \leq c$
Invariants: $c \leq 0$
Preconditions: $\forall$ is not supported
Postconditions: *sum* is supported



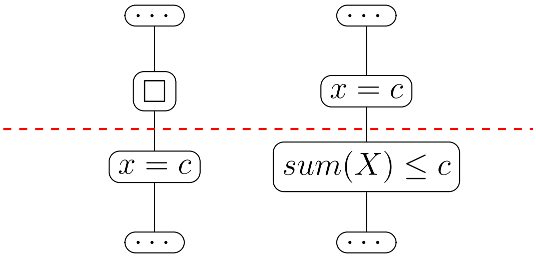
A41: less-than → sum-less-than (attribute-constant)

Rule: $\forall x \in X : x < c \sqsubseteq_K sum(X) < c$
Invariants: $c \leq 0$
Preconditions: $\forall$ is not supported
Postconditions: *sum* is supported



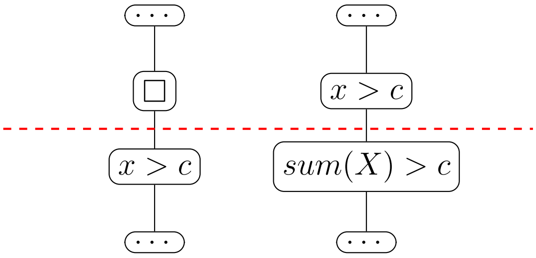
A42: equals → sum-less-than-or-equals-to (attribute-constant)

Rule: $\forall x \in X : x = c \sqsubseteq_K sum(X) \leq c$
Invariants: $c \leq 0$
Preconditions: $\forall$ or $=$ are not supported
Postconditions: *sum* und $\leq$ is supported



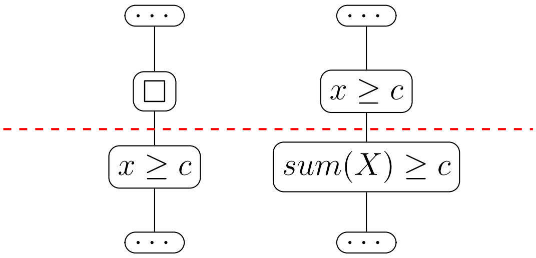
A43: greater-than → sum-greater-than (attribute-constant)

Rule: $\forall x \in X : x > c \sqsubseteq_K sum(X) > c$
Invariants: $c \geq 0$
Preconditions: $\forall$ or $=$ are not supported
Postconditions: *sum* und $\leq$ is supported



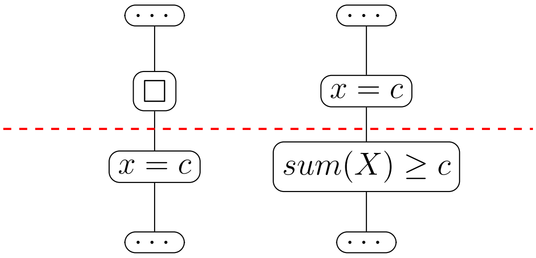
A44: greater-than-or-equals-to → sum-greater-than-or-equals-to (attribute-constant)

Rule: $\forall x \in X : x \geq c \sqsubseteq_K sum(X) \geq c$
Invariants: $c \geq 0$
Preconditions: $\forall$ is not supported
Postconditions: *sum* is supported



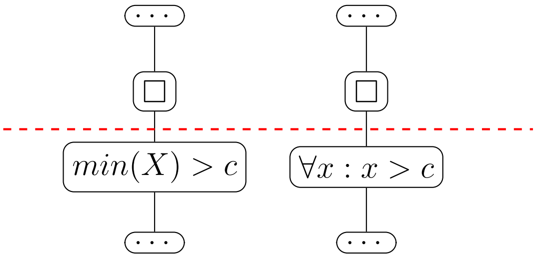
A45: equals → sum-greater-than-or-equals-to (attribute-constant)

Rule: $\forall x \in X : x = c \sqsubseteq_K \mathit{sum}(X) \geq c$
Invariants: $c \geq 0$
Preconditions: $\forall$ or $=$ are not supported
Postconditions: sum und $\geq$ is supported



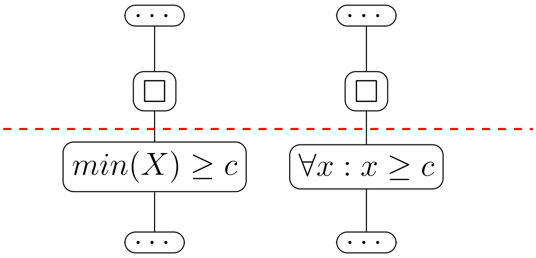
A46: minimum-greater-than → greater-than (attribute-constant)

Rule: $\mathit{min}(X) > c \sqsubseteq_K \forall x \in X : x > c$
Invariants: —
Preconditions: min is not supported
Postconditions: $\forall$ is supported



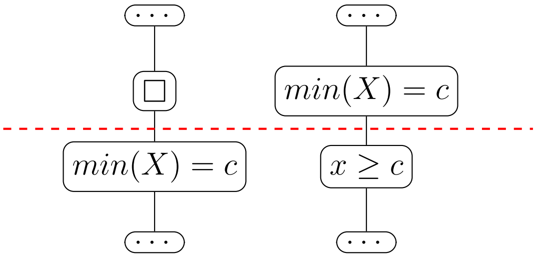
A47: minimum-greater-than-or-equals-to → greater-than-or-equals-to (attribute-constant)

Rule: $\mathit{min}(X) \geq c \sqsubseteq_K \forall x \in X : x \geq c$
Invariants: —
Preconditions: min is not supported
Postconditions: $\forall$ is supported



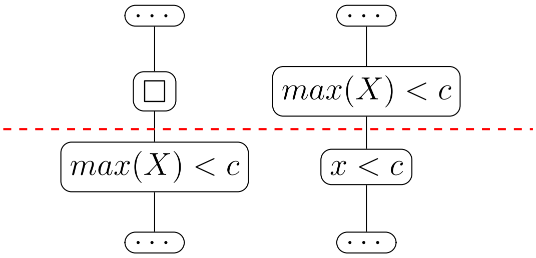
A48: minimum-equals → greater-than-or-equals-to (attribute-constant)

Rule: $\mathit{min}(X) = c \sqsubseteq_K \forall x \in X : x \geq c$
Invariants: —
Preconditions: min or $=$ are not supported
Postconditions: $\geq$ und $\forall$ is supported



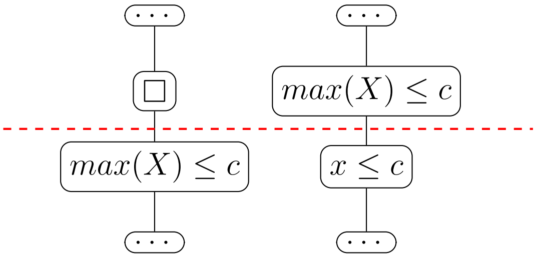
A49: maximum-less-than → less-than (attribute-constant)

Rule: $\mathit{max}(X) < c \sqsubseteq_K \forall x \in X : x < c$
Invariants: —
Preconditions: max is not supported
Postconditions: $\forall$ is supported



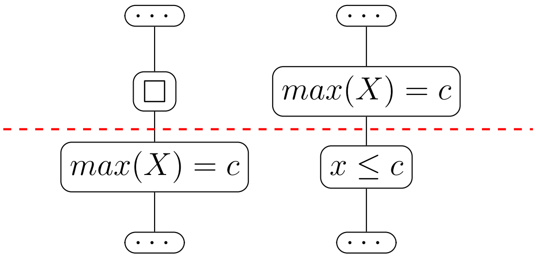
A50: maximum-less-than-or-equals-to → less-than-or-equals-to (attribute-constant)

Rule: $max(X) \leq c \sqsubseteq_K \forall x \in X : x \leq c$
Invariants: —
Preconditions: max is not supported
Postconditions: $\forall$ is supported



A51: maximum-equals → less-than-or-equals-to (attribute-constant)

Rule: $max(X) = c \sqsubseteq_K \forall x \in X : x \leq c$
Invariants: —
Preconditions: max or $=$ are not supported
Postconditions: $\leq$ and $\forall$ are supported



3 Measurements

3.1 Amarok dataset

Rule	A01		A02		A03		A04		A05		A06		A07		A08	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	41	23	12	37	11	23	11	30	12	27	12	27	12	26	11	25
#2	13	23	12	30	11	23	11	35	11	27	11	27	12	29	11	33
#3	12	22	12	26	11	30	11	29	11	27	11	30	11	27	11	26
#4	13	22	17	26	11	29	11	28	11	27	11	28	11	26	11	28
#5	12	37	14	23	11	28	11	28	14	27	11	27	11	25	11	26
#6	12	27	13	23	11	28	11	28	17	27	11	27	11	26	11	26
#7	12	23	12	23	11	38	11	34	14	27	11	27	11	30	11	24
#8	11	23	13	23	11	30	11	28	13	27	11	27	11	26	11	25
#9	11	22	12	22	11	40	11	37	12	27	11	31	11	29	11	24
#10	11	29	11	22	11	33	11	28	11	27	11	29	11	26	11	25
AVG	14,8	25,1	12,8	25,5	11,0	30,2	11,0	30,5	12,6	27,0	11,1	28,0	11,2	27,0	11,0	26,2
Overhead in %	69,59		99,22		174,55		177,27		114,29		152,25		141,07		138,18	

Table 1: Measurements for the rules A01 to A08 on the Amarok dataset.

Rule	A09		A10		A11		A12		A13		A14		A15		A16	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	11	25	12	22	12	22	12	31	12	28	12	28	12	31	14	34
#2	11	25	11	22	12	25	12	28	12	28	12	28	15	30	13	32
#3	11	25	11	22	12	23	12	23	12	28	12	28	15	31	13	33
#4	11	25	12	22	11	25	12	22	12	28	12	28	13	31	13	32
#5	11	25	12	22	12	24	12	22	11	29	12	27	12	30	12	33
#6	11	25	12	22	12	23	12	22	12	29	11	24	12	30	12	32
#7	11	25	15	22	12	23	12	22	12	29	12	24	11	36	12	33
#8	12	25	13	22	11	29	12	22	12	30	11	24	11	34	13	34
#9	12	25	12	22	12	24	12	22	11	29	12	29	12	25	13	32
#10	12	24	12	22	12	22	12	22	12	29	12	29	11	24	13	31
AVG	11,3	24,9	12,2	22,0	11,8	24,0	12,0	23,6	11,8	28,7	11,8	26,9	12,4	30,2	12,8	32,6
Overhead in %	120,35		80,33		103,39		96,67		143,22		127,97		143,55		154,69	

Table 2: Measurements for the rules A09 to A16 on the Amarok dataset.

Rule	A17		A18		A19		A20		A21		A22		A23		A24	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	13	32	13	27	19	25	14	25	15	29	16	29	12	25	13	25
#2	13	30	13	27	15	25	14	25	15	29	15	29	12	25	13	25
#3	13	27	13	27	14	25	14	25	15	29	15	29	12	25	13	25
#4	12	27	13	27	14	25	14	25	15	29	15	29	13	25	12	26
#5	12	27	13	27	14	25	14	25	15	29	22	29	14	25	12	25
#6	13	27	13	27	14	24	14	26	15	28	18	29	14	25	12	26
#7	13	43	13	27	14	25	14	26	15	28	16	29	13	25	12	26
#8	13	27	13	27	14	25	14	25	15	28	15	28	13	25	12	26
#9	14	27	13	27	14	25	14	25	15	29	15	29	13	25	12	25
#10	13	27	21	27	14	25	14	25	15	29	15	29	12	25	12	25
AVG	12,9	29,4	13,8	27,0	14,6	24,9	14,0	25,2	15,0	28,7	16,2	28,9	12,8	25,0	12,3	25,4
Overhead in %	127,91		95,65		70,55		80,00		91,33		78,40		95,31		106,50	

Table 3: Measurements for the rules A17 to A24 on the Amarok dataset.

Rule	**A25**		**A26**		**A27**		**A28**		**A29**		**A30**		**A31**		**A32**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	14	28	14	28	17	9	17	9	19	26	19	26	19	16	17	15
#2	13	28	13	29	25	9	17	9	19	26	20	26	18	15	17	19
#3	13	28	13	29	19	9	17	9	19	25	20	26	18	15	17	17
#4	14	29	13	29	17	9	17	9	20	26	19	26	22	16	17	17
#5	13	29	13	30	17	9	17	9	19	26	19	26	20	16	17	16
#6	13	29	13	30	17	9	17	9	19	26	19	26	18	16	17	16
#7	13	29	16	29	17	9	17	9	19	26	19	26	17	15	17	16
#8	13	29	17	29	17	9	17	9	20	26	19	26	17	15	17	16
#9	16	28	15	30	17	9	17	9	20	26	21	26	17	15	17	16
#10	14	28	13	29	17	9	17	9	19	26	21	26	17	15	17	16
AVG	13,6	28,5	14,0	29,2	18,0	9,0	17,0	9,0	19,3	25,9	19,6	26,0	18,3	15,4	17,0	16,4
Overhead in %	109,56		108,57		-50,00		-47,06		34,20		32,65		-15,85		-3,53	

Table 4: Measurements for the rules A25 to A32 on the Amarok dataset.

Rule	**A33**		**A34**		**A35**		**A36**		**A37**		**A38**		**A39**		**A40**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	17	9	16	10	17	27	16	27	19	28	17	27	18	27	19	28
#2	16	9	16	10	16	26	18	27	19	27	17	27	17	27	18	28
#3	16	9	16	10	16	26	17	27	19	27	17	26	17	30	19	27
#4	16	9	16	10	16	26	17	27	23	27	17	26	17	37	18	28
#5	16	10	16	10	16	26	17	27	23	27	17	26	18	29	18	28
#6	17	10	16	10	17	26	17	27	20	28	17	26	18	26	18	28
#7	16	10	17	10	16	26	19	26	20	45	17	27	17	26	18	28
#8	16	10	17	10	16	27	18	26	20	30	17	26	17	25	18	28
#9	16	9	16	10	16	27	16	26	20	28	17	26	17	26	18	28
#10	16	9	16	10	16	26	16	26	20	28	17	27	17	26	18	28
AVG	16,2	9,4	16,2	10,0	16,2	26,3	17,1	26,6	20,3	29,5	17,0	26,4	17,3	27,9	18,2	27,9
Overhead in %	-41,98		-38,27		62,35		55,56		45,32		55,29		61,27		53,30	

Table 5: Measurements for the rules A33 to A40 on the Amarok dataset.

Rule	**A41**		**A42**		**A43**		**A44**		**A45**		**A46**		**A47**		**A48**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	18	29	18	31	17	33	14	33	16	34	9	16	11	16	16	28
#2	17	28	18	31	16	33	14	31	16	34	9	16	11	16	14	30
#3	16	28	18	40	16	31	14	30	17	42	9	16	11	16	12	35
#4	15	28	18	36	16	32	14	30	16	33	9	16	11	16	11	28
#5	15	28	18	38	16	32	14	35	17	32	9	16	11	16	11	27
#6	15	28	18	39	17	28	14	39	16	32	9	16	11	16	10	27
#7	15	28	18	38	17	26	14	31	16	44	9	16	11	19	10	27
#8	15	28	18	38	16	26	14	30	16	36	9	16	11	17	9	27
#9	15	28	18	37	16	26	14	32	16	35	12	16	11	17	9	27
#10	15	28	18	37	16	26	14	31	16	32	11	16	11	17	9	27
AVG	15,6	28,1	18,0	36,5	16,3	29,3	14,0	32,2	16,2	35,4	9,5	16,0	11,0	16,6	11,1	28,3
Overhead in %	80,13		102,78		79,75		130,00		118,52		68,42		50,91		154,95	

Table 6: Measurements for the rules A43 to A48 on the Amarok dataset.

Rule	**A49**		**A50**		**A51**	
Measurement in ms	-	+	-	+	-	+
#1	9	17	9	17	9	27
#2	9	16	9	16	9	27
#3	9	16	9	16	9	28
#4	9	16	9	17	9	30
#5	9	16	9	16	9	35
#6	9	17	9	16	9	29
#7	9	16	9	16	9	37
#8	9	17	9	17	9	29
#9	9	16	9	16	9	35
#10	9	16	9	17	9	30
AVG	9,0	16,3	9,0	16,4	9,0	30,7
Overhead in %	81,11		82,22		241,11	

Table 7: Measurements for the rules A49 to A51 on the Amarok dataset.

3.2 TPCH dataset

Rule	**K01**		**K02**		**K03**		**K04**		**K05**		**K06**		**K07**		**K08**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	40	24	8	22	68	93	84	89	34	31	7	10	6	10	11	8
#2	38	23	7	21	59	85	83	82	32	24	6	10	7	10	11	8
#3	39	22	7	21	60	96	81	81	32	24	6	11	6	10	11	8
#4	39	20	7	21	59	97	84	81	32	25	6	11	6	10	11	8
#5	39	20	7	22	59	98	82	81	34	26	6	10	6	10	11	7
#6	38	21	7	21	60	95	81	83	33	27	6	11	6	10	11	8
#7	38	21	6	21	61	94	83	82	32	27	6	11	6	11	11	7
#8	40	21	6	21	59	97	84	78	33	24	6	10	7	11	11	7
#9	39	21	6	21	62	99	84	80	32	24	7	10	6	14	11	7
#10	39	20	7	22	59	94	83	78	32	24	6	10	6	13	11	7
AVG	38,9	21,3	6,8	21,3	60,6	94,8	82,9	81,5	32,6	25,6	6,2	10,4	6,2	10,9	11,0	7,5
Overhead in %	-45,24		213,24		56,44		-1,69		-21,47		67,74		75,81		-31,82	

Table 8: Measurements for the rules K01 to K08 on the Amarok dataset.

Rule	**K09**		**K10**		**K11**		**K12**		**K13**		**K14**		**K15**		**K16**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	24	21	8	16	22	18	7	16	5	3	3	3	2	1	1	1
#2	24	21	7	16	21	17	6	16	3	1	2	3	2	2	2	2
#3	22	22	7	16	21	17	6	16	2	2	2	3	1	1	1	2
#4	21	22	7	16	23	17	6	16	3	2	3	3	2	2	1	2
#5	21	22	6	16	21	17	6	16	3	1	2	4	2	1	2	2
#6	21	22	6	16	21	17	6	16	2	2	3	3	1	2	1	2
#7	21	22	6	16	21	17	6	16	3	1	2	3	2	1	1	2
#8	21	34	6	16	21	17	6	16	3	2	2	4	1	1	2	2
#9	25	27	6	16	21	17	6	16	3	2	3	3	2	2	1	1
#10	34	23	6	16	21	16	6	16	3	1	2	4	2	1	1	2
AVG	23,4	23,6	6,5	16,0	21,3	17,0	6,1	16,0	3,0	1,7	2,4	3,3	1,7	1,4	1,3	1,8
Overhead in %	0,85		146,15		-20,19		162,30		-43,33		37,50		-17,65		38,46	

Table 9: Measurements for the rules K09 to K16 on the Amarok dataset.

Rule	**K17**		**K18**		**K19**		**K20**		**K21**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+
#1	1	2	4	4	3	7	25	34	28	32
#2	2	2	5	3	4	5	24	34	27	31
#3	2	1	4	4	3	6	25	33	28	30
#4	3	2	5	3	4	5	24	33	28	30
#5	1	2	4	4	3	6	24	33	28	38
#6	2	1	5	3	3	5	24	33	28	33
#7	2	2	4	4	4	5	24	33	28	28
#8	2	2	5	5	3	4	24	33	28	28
#9	2	1	5	5	6	5	24	33	27	27
#10	1	2	4	5	3	5	24	33	28	27
AVG	1,8	1,7	4,5	4,0	3,6	5,3	24,2	33,2	27,8	30,4
Overhead in %	-5,56		-11,11		47,22		37,19		9,35	

Table 10: Measurements for the rules K17 to K21 on the Amarok dataset.

Rule	**L01**		**L02**		**L03**		**L04**		**L05**		**L06**		**L07**		**L08**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	15	20	7	11	6	18	15	19	10	12	6	9	8	15	23	29
#2	15	19	6	11	7	13	15	23	10	11	7	8	7	14	23	30
#3	14	18	7	11	6	12	15	26	11	13	7	9	6	14	22	30
#4	15	19	7	12	7	12	15	21	10	12	6	8	8	15	23	29
#5	15	19	6	10	6	11	14	20	10	12	7	8	7	14	22	30
#6	15	19	7	11	7	11	15	19	10	13	6	8	6	14	23	29
#7	16	20	6	12	6	11	14	19	10	12	7	9	7	14	23	31
#8	14	19	7	11	6	11	15	18	11	12	7	8	7	15	23	30
#9	15	19	6	10	7	11	14	19	12	13	6	8	6	14	23	30
#10	15	20	6	11	6	10	15	18	10	12	7	9	7	14	23	30
AVG	14,9	19,2	6,5	11,0	6,4	12,0	14,7	20,2	10,4	12,2	6,6	8,4	6,9	14,3	22,8	29,8
Overhead in %	28,86		69,23		87,50		37,41		17,31		27,27		107,25		30,70	

Table 11: Measurements for the rules L01 to L08 on the Amarok dataset.

3.3 MuSAMA dataset

Rule	**A01**		**A02**		**A03**		**A04**		**A05**		**A06**		**A07**		**A08**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	12664	35707	8898	36159	9684	35683	9256	34926	9406	35868	9748	35207	9563	27952	9563	27137
#2	9731	35011	9730	35928	9598	35497	9654	34188	9645	35917	9726	35853	9376	27463	9737	27435
#3	9168	35676	9597	36032	9166	35556	9816	35097	9520	36088	9641	35488	9765	27750	9845	27942
#4	9598	35966	9753	36233	9453	35764	9460	34101	9585	35834	9577	33113	9544	24630	9651	27555
#5	9635	35611	9657	35686	9761	36091	9359	35499	9880	35179	9690	34514	9552	25584	9642	27316
#6	9230	35742	9615	35977	9640	35628	9442	36158	9451	36542	9661	35749	9682	27787	9473	27775
#7	9621	35575	9671	35745	9714	35857	9457	35826	8796	36013	9602	35864	9578	27462	9459	27626
#8	9631	35623	9564	36106	9813	35303	9399	36412	9606	35664	9661	35146	9699	27293	9400	27351
#9	9591	35444	9795	36005	9913	35820	9662	35639	9661	35182	9771	35166	9822	27401	9719	27239
#10	9439	35790	9519	35062	9285	35569	9489	35495	9519	35867	9668	35865	9665	27352	9495	27334
AVG	9830,8	35614,5	9579,9	35893,3	9602,7	35676,8	9499,4	35334,1	9506,9	35815,4	9674,5	35196,5	9624,6	27067,4	9598,4	27471,0
Overhead in %	262,27		274,67		271,53		271,96		276,73		263,81		181,23		186,20	

Table 12: Measurements for the rules A01 to A08 on the TPC-H dataset.

Rule	**A09**		**A10**		**A11**		**A12**		**A13**		**A14**		**A15**		**A16**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	9691	27364	8773	28535	9903	29417	9606	28901	9544	36926	9677	36639	8890	37045	11806	41396
#2	9625	27452	8722	29373	9585	29120	9543	29019	9372	36983	9665	36590	9611	36142	11377	41521
#3	9771	27189	8737	29282	9387	29654	9786	28579	9183	37324	9637	37233	9568	37177	11581	41307
#4	9702	27476	8904	28721	9678	29126	9693	27783	9567	37286	9672	36641	9609	36868	11513	41559
#5	9690	27525	9196	29664	9381	29606	9559	28433	9460	37209	9652	36287	9276	36496	11759	41952
#6	9776	27717	9628	29484	9554	29607	9476	29480	9570	36867	9761	36101	10038	37031	11519	39489
#7	9676	27715	9643	29714	9680	28923	8857	29977	9638	36312	9738	36976	9353	37107	11430	41690
#8	9781	27583	9515	29582	9374	29691	9164	29328	9766	36434	9586	36961	9836	36809	11650	40969
#9	9001	26771	9603	28750	9447	29898	8920	28855	9534	36973	9753	36666	9702	35126	11089	40850
#10	8447	27463	9633	29393	9571	29331	9246	29243	9219	36586	9755	36140	9696	37074	11624	41240
AVG	9516,0	27425,5	9235,4	29249,8	9556,0	29437,3	9385,0	28959,8	9485,3	36890,0	9689,6	36623,4	9557,9	36687,5	11534,8	41197,3
Overhead in %	188,20		216,71		208,05		208,58		288,92		277,97		283,84		257,16	

Table 13: Measurements for the rules A09 to A16 on the TPC-H dataset.

Rule	**A17**		**A18**		**A19**		**A20**		**A21**		**A22**		**A23**		**A24**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	11589	41480	11588	41342	15373	43713	14536	41391	17067	46366	17020	46277	11647	42728	11853	42676
#2	11364	41115	11827	39590	15149	43180	14730	42535	16947	46131	16182	46685	11997	42057	10917	42626
#3	11013	40939	11358	38339	15233	42348	15139	42752	16861	45933	15797	46559	11925	42479	11726	43242
#4	11200	41460	11610	40998	15290	42512	15084	42618	16920	46969	16524	46338	11719	42545	11896	39898
#5	11350	41098	11547	40420	14199	42596	15506	43121	16840	45914	16874	45187	11551	42686	11460	43280
#6	11378	41099	11322	41370	14288	43014	15018	41952	16366	46701	16160	46256	11910	43025	11307	42926
#7	11261	41100	11600	40051	13900	42405	15053	42659	16929	46250	16837	46265	11475	43023	11759	42491
#8	11378	41450	11431	40917	14807	42923	15031	42137	16177	46135	16843	45705	12080	42821	11661	42860
#9	11540	41322	11013	41613	15270	41817	15313	42620	16941	46064	16874	45709	11989	42646	11576	43290
#10	11605	41390	11534	41449	15254	42705	15161	42759	16859	45795	16796	47085	11848	42949	11995	42954
AVG	11367,8	41245,3	11483,0	40608,9	14876,3	42721,3	15057,1	42454,4	16790,7	46225,8	16590,7	46206,6	11814,1	42695,9	11615,0	42624,3
Overhead in %	262,83		253,64		187,18		181,96		175,31		178,51		261,40		266,98	

Table 14: Measurements for the rules A17 to A24 on the TPC-H dataset.

Rule	**A25**		**A26**		**A27**		**A28**		**A29**		**A30**		**A31**		**A32**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	13873	46988	13652	46876	11852	9074	11048	9129	14486	22975	14150	22750	12384	35048	12123	36328
#2	13720	47576	14047	47360	12209	8866	12070	9010	14655	22660	14861	23330	12502	35197	12263	35944
#3	13861	47501	13915	46966	12229	9089	11804	9094	15042	22973	14792	23485	12678	35824	12223	36480
#4	13559	47045	12800	43895	12029	8894	12109	8986	14867	23114	14898	22247	12747	35398	12019	36203
#5	13774	47059	12717	43663	11914	9098	12125	9044	14031	23122	14865	22981	12571	35967	12079	35939
#6	13918	46550	13365	47058	11523	9021	11813	9084	15082	22968	14568	23156	12395	35821	12028	36114
#7	13784	46720	14122	46446	11993	9293	11834	8636	14889	23047	14944	22693	12416	35220	12146	35129
#8	13874	45915	13838	45491	11672	8952	11550	9274	14545	23143	14754	22006	12478	35588	12111	36118
#9	13874	46749	13787	46802	11049	9205	12040	9122	15003	22895	14718	22368	12365	35827	12112	36148
#10	13934	45593	12883	47172	12055	9099	11560	8863	14781	22912	14720	23259	12223	36629	11967	36401
AVG	13817,1	46769,6	13512,6	46172,9	11852,5	9059,1	11795,3	9024,2	14738,1	22980,9	14727,0	22827,5	12475,9	35651,9	12107,1	36080,4
Overhead in %	238,49		241,70		-23,57		-23,49		55,93		55,00		185,77		198,01	

Table 15: Measurements for the rules A25 to A32 on the TPC-H dataset.

Rule	**A33**		**A34**		**A35**		**A36**		**A37**		**A38**		**A39**		**A40**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	11939	8911	12212	8268	11949	25518	11973	25537	14797	26645	11614	23347	11391	24172	14021	26735
#2	12082	9226	11720	8783	11904	24643	11813	25642	14728	28179	12060	24748	11448	24395	13900	26343
#3	11948	9003	12072	8519	11996	25358	11140	25155	14781	27861	11556	25071	11147	23521	13374	26423
#4	11960	8924	11010	9020	11716	25084	11630	25769	14675	27107	11491	25188	10736	23327	13966	26961
#5	11792	8955	12035	9048	12090	25532	11701	25185	14499	26792	11868	24864	11346	24406	13753	26619
#6	11817	9030	12140	9042	11529	25519	11850	25815	13581	27204	11907	24197	11154	24982	13815	26911
#7	11961	9084	11474	8982	11892	24329	11890	25476	13551	25472	11935	24701	11318	25422	13171	26567
#8	12033	8913	12168	9056	11685	25406	11945	24913	13818	25855	11977	24597	12264	24689	12205	26543
#9	11299	8653	12186	9012	12039	25359	12097	25549	14506	26441	12029	24619	11511	24968	13859	26096
#10	11936	9179	12075	8597	11483	25685	11862	24578	14230	25621	11597	24786	12029	24658	13716	26786
AVG	11876,7	8987,8	11909,2	8832,7	11828,3	25243,3	11790,1	25361,9	14316,6	26717,7	11803,4	24611,8	11434,4	24454,0	13578,0	26598,4
Overhead in %	-24,32		-25,83		113,41		115,11		86,62		108,51		113,86		95,89	

Table 16: Measurements for the rules A33 to A40 on the TPC-H dataset.

Rule	**A41**		**A42**		**A43**		**A44**		**A45**		**A46**		**A47**		**A48**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	13821	27160	16589	26449	11999	22973	11931	23492	14727	25947	9375	11232	9513	10772	12181	20013
#2	13920	26941	16247	27954	11245	23544	11723	23543	14055	25054	9372	11142	9518	10829	8719	19849
#3	14066	26606	16705	27684	12046	23292	12048	23523	14411	25390	9436	10832	9450	10934	8822	19680
#4	14182	27089	16905	27973	11866	23211	11945	23428	14579	26418	9337	11161	9422	11057	8379	18981
#5	13862	25900	16694	27380	12037	22572	11977	23475	14565	25538	9313	10995	9469	11021	8315	19905
#6	14064	27317	16829	28099	11944	22773	11782	23283	14572	25309	9376	11113	9038	11071	7699	19820
#7	14056	26860	16776	27938	11794	23328	11989	23387	14477	26163	9294	11144	9513	11177	7630	19575
#8	13971	26887	16855	28175	11763	23317	11670	23205	14300	25718	9409	10978	9520	11023	8192	19836
#9	14137	25848	16692	28562	11803	22954	11984	22872	14629	25971	9331	11187	9371	10999	8217	19020
#10	14358	24989	16432	28108	12053	23461	11448	22909	14147	25328	8995	10958	9387	10875	7670	19356
AVG	14043,7	26559,7	16672,4	27832,2	11855,0	23142,5	11849,7	23311,7	14446,2	25683,6	9323,8	11074,2	9420,1	10975,8	8582,4	19603,5
Overhead in %	89,12		66,94		95,21		96,73		77,79		18,77		16,51		128,42	

Table 17: Measurements for the rules A41 to A48 on the TPC-H dataset.

Rule	**A49**		**A50**		**A51**	
Measurement in ms	-	+	-	+	-	+
#1	9302	10955	9506	10991	9566	19640
#2	9359	10909	9395	10765	9297	19699
#3	9297	11054	9383	10839	9282	19766
#4	9064	10848	9459	10923	9057	19519
#5	9162	10749	9313	10698	9388	19644
#6	9263	10946	9317	10775	9189	18373
#7	9301	10901	9437	10808	9404	18513
#8	9174	10995	9415	10764	9451	19268
#9	9537	11050	9518	10033	9359	19283
#10	9292	11166	9359	10667	9462	19270
AVG	9275,1	10957,3	9410,2	10726,3	9345,5	19297,5
Overhead in %	18,14		13,99		106,49	

Table 18: Measurements for the rules A49 to A51 on the TPC-H dataset.

Rule	**K01**		**K02**		**K03**		**K04**		**K05**		**K06**		**K07**		**K08**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	36240	10105	4239	10547	42836	41054	42685	51345	9922	12248	10074	16250	13812	9721	10369	11565
#2	34544	9734	3780	10499	41918	51024	44134	57881	10570	10854	10166	15606	13935	10642	9954	10742
#3	32245	9663	3889	10915	42832	47785	43302	56527	9859	11695	9779	16687	16771	10512	9975	11459
#4	32388	9082	3941	10625	45499	43121	43534	54647	10118	12487	10151	14487	13335	10082	9492	12273
#5	32195	9583	3901	10198	47512	42289	43376	54458	9336	10171	9559	15309	17162	10079	9101	11837
#6	33746	9706	3931	10641	44853	48198	43157	54430	10388	11266	9975	12993	14641	10172	9571	12592
#7	34311	9212	3839	10656	43827	41960	42906	54864	9971	16000	9797	12825	15332	10109	9872	11879
#8	34193	10270	3914	10592	41600	42062	43892	58959	9807	10425	9953	16080	14167	10388	10273	12175
#9	33615	8790	3945	10845	43427	45114	42883	53874	10022	16709	9051	18806	17766	10361	9877	11373
#10	34405	9437	3882	10687	43701	44967	42349	53178	10287	10934	9861	14063	13805	9449	10116	11905
AVG	33788,2	9558,2	3926,1	10620,5	43800,5	44757,4	43221,8	55016,3	10028,0	12278,9	9836,6	15310,6	15072,6	10151,5	9860,0	11780,0
Overhead in %	-71,71		170,51		2,18		27,29		22,45		55,65		-32,65		19,47	

Table 19: Measurements for the rules K01 to K08 on the TPC-H dataset.

Rule	**K09**		**K10**		**K11**		**K12**		**K13**		**K14**		**K15**		**K16**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	9114	9057	7545	9433	6695	6350	4130	7408	679	168	69	301	232	69	64	72
#2	9160	9166	7956	9389	7025	6275	4401	7331	259	60	59	279	221	70	72	74
#3	9185	9332	8368	9471	6967	7272	4151	7302	266	61	58	280	227	79	60	67
#4	8701	9020	8904	9475	6382	6252	4221	7363	264	61	61	288	263	69	69	74
#5	9104	7773	8813	9335	6785	6668	4217	7383	253	69	60	258	250	80	69	69
#6	9183	8669	8369	9318	6554	6193	4337	6976	258	90	69	258	239	68	60	70
#7	9089	9026	8098	8723	6867	6282	4397	7407	258	70	63	259	251	63	60	70
#8	9140	8911	8609	9499	6972	7084	4465	7368	241	64	64	273	234	63	60	68
#9	9426	8623	8761	8187	6715	6330	4035	7381	247	75	71	268	227	65	59	68
#10	8984	9250	9004	8900	7023	6503	4101	7264	257	66	70	271	235	70	79	68
AVG	9108,6	8882,7	8442,7	9173,0	6798,5	6520,9	4245,5	7318,3	298,2	78,4	64,4	273,5	237,9	69,6	65,2	70,0
Overhead in %	-2,48		8,65		-4,08		72,38		-73,71		324,69		-70,74		7,36	

Table 20: Measurements for the rules K09 to K16 on the TPC-H dataset.

Rule	**K17**		**K18**		**K19**		**K20**		**K21**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+
#1	75	234	852	272	317	653	32665	47399	42998	30451
#2	69	251	706	275	469	630	30328	43226	44123	30176
#3	70	238	675	273	351	609	28558	42505	42139	27297
#4	67	242	660	530	326	643	28829	42555	40885	27815
#5	66	246	624	344	325	618	28841	44766	41579	29937
#6	67	240	656	321	324	608	28866	43540	41957	30061
#7	68	235	675	328	337	571	28616	43873	42198	28154
#8	69	234	629	337	339	584	29292	43888	42005	28471
#9	70	239	637	361	321	566	28979	43281	42699	29966
#10	69	210	622	326	390	551	27723	43590	42171	29821
AVG	69,0	236,9	673,6	336,7	349,9	603,3	29269,7	43862,3	42275,4	29214,9
Overhead in %	243,33		-50,01		72,42		49,86		-30,89	

Table 21: Measurements for the rules K17 to K21 on the TPC-H dataset.

4 Performance Analysis

In this section, we show the effect of the number of nodes, the computing power and the selectivity towards the runtime for both unmodified and rewritten query. For the performance analysis, we took the an example query that takes data

Rule	L01		L02		L03		L04		L05		L06		L07		L08	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	7730	8934	2080	4338	2087	4130	6279	8437	7896	11359	3142	6520	3065	5535	5728	8185
#2	6558	9207	2101	4362	2090	3977	6155	7659	6774	10883	3031	6483	3084	5439	5515	8353
#3	6504	9097	2099	4201	2059	4196	6358	8145	6351	11097	3129	5967	3140	5609	5658	8196
#4	6937	9086	2108	4281	1974	4190	6455	6803	7592	10739	3074	6200	3110	5536	5780	8260
#5	6231	8558	2070	4311	2118	4282	5842	8228	6651	10153	2953	6388	3324	5454	5296	8307
#6	6391	8388	2074	4300	2065	4076	6573	8277	7223	10779	2914	6485	3110	5430	4897	8088
#7	6658	8299	2061	4306	2125	3729	5966	7463	6949	10562	3026	6408	3122	5419	5585	8073
#8	6035	8291	2117	4261	2197	3535	5939	8648	7433	10609	3066	6431	3129	5434	5786	8185
#9	6305	8335	2085	4187	2163	3608	5556	8575	7351	10417	3041	6463	3133	5589	5180	8079
#10	6342	8170	2076	4287	2112	4105	6431	7814	7509	10363	3051	6460	2981	5597	5508	8238
AVG	6569,1	8636,5	2087,1	4283,4	2099,0	3982,8	6155,4	8004,9	7172,9	10696,1	3042,7	6380,5	3119,8	5504,2	5493,3	8196,4
Overhead in %	31,47		105,23		89,75		30,05		49,12		109,70		76,43		49,21	

Table 22: Measurements for the rules L01 to L08 on the TPC-H dataset.

Rule	A01		A02		A03		A04		A05		A06		A07		A08	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	1663	2687	1126	2698	1127	1090	1104	2657	969	2703	1106	2526	1137	2686	1133	2711
#2	1153	2757	1119	2755	1114	1085	1144	2706	977	2685	1160	2578	1134	2726	1147	2784
#3	1135	2650	1121	2815	1110	1139	1126	2679	990	2697	1118	2500	1117	2727	1128	2772
#4	1132	2699	1126	2674	1120	1105	1171	2725	1132	2675	1120	2500	1164	2800	1119	2720
#5	1147	2669	1130	2579	1120	1105	1129	2742	1153	2679	1131	2435	1125	2789	1162	2738
#6	1123	2784	1128	2724	1145	1095	1134	2738	1115	2675	1115	2374	1109	2851	1152	2783
#7	1128	2653	1158	2733	1153	1102	1065	2694	1124	2721	1125	2466	1178	2753	1118	2702
#8	1149	2562	1173	2667	1140	1111	970	2567	1104	2648	1125	2556	1176	2703	1129	2769
#9	1118	2677	1101	2752	1152	1124	965	2474	1112	2701	1010	2569	1136	2705	1107	2751
#10	1146	2680	1166	2678	1145	1106	965	2725	1126	2750	958	2498	1185	2741	1112	2731
AVG	1189,4	2681,8	1134,8	2707,5	1132,6	1106,2	1077,3	2670,7	1080,2	2693,4	1096,8	2500,2	1146,1	2748,1	1130,7	2746,1
Overhead in %	125,48		138,59		-2,33		147,91		149,34		127,95		139,78		142,87	

Table 23: Measurements for the rules A01 to A08 on the MuSAMA dataset.

Rule	A09		A10		A11		A12		A13		A14		A15		A16	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	1136	2664	1118	1156	1127	1164	1155	963	1166	2793	1167,00	2689	1106	2707	1160	1062
#2	1168	2778	1150	1134	1153	1170	1121	1009	1120	2716	1151,00	2446	1121	2687	1190	950
#3	1162	2748	1124	1116	1149	1140	1147	1134	1173	2681	1113,00	2461	1129	2723	1161	1097
#4	1122	2698	1168	1110	1125	1143	1131	1129	1142	2746	1127,00	2656	1121	2645	1124	1083
#5	1127	2695	1200	1115	1133	1143	1166	1130	1161	2756	1141,00	2710	1132	2517	1139	1099
#6	1111	2642	1131	1105	1147	1164	1156	1142	1142	2727	1159,00	2666	1136	2710	1122	1119
#7	1137	2774	1132	1117	1126	1169	1166	1181	1125	2673	1137,00	2666	1130	2380	1117	1154
#8	1118	2720	1129	1192	1136	1028	1141	1188	1138	2428	1114,00	2719	1099	2591	1122	1109
#9	1112	2669	1120	1129	1138	986	1163	1153	1141	2376	1119,00	2701	1174	2678	1113	1117
#10	1115	2452	1142	1208	1136	970	1170	1159	1149	2544	1112,00	2694	1125	2771	1110	1007
AVG	1130,8	2684,0	1141,4	1138,2	1137,0	1107,7	1151,6	1118,8	1145,7	2644,0	1134,0	2640,8	1127,3	2640,9	1135,8	1079,7
Overhead in %	137,35		-0,28		-2,58		-2,85		130,78		132,87		134,27		-4,94	

Table 24: Measurements for the rules A09 to A16 on the MuSAMA dataset.

Rule	A17		A18		A19		A20		A21		A22		A23		A24	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	1120	1087	1110	1107	1150	2821	1144	2765	1179	940	1203	1105	1104	2655,00	1096	2279
#2	1123	1121	1154	1089	1153	2674	1150	2789	1181	1029	1148	1105	1137	2653,00	1120	2438
#3	1122	1100	1122	985	1162	2512	1153	2762	1205	1105	1014	990	1119	2709,00	1096	2683
#4	1166	1065	1132	1091	1180	2580	1104	2726	1157	1087	979	1000	1133	2669,00	1126	2721
#5	1152	951	1164	1083	1155	2581	1004	2735	1158	1124	1205	951	1101	2668,00	1128	2661
#6	1117	947	1103	1109	1173	2686	999	2734	1136	1131	1182	935	1104	2682,00	1179	2769
#7	1119	957	1156	1090	1172	2700	991	2710	1155	1125	1163	1055	1120	2639,00	1119	2821
#8	1126	1068	1152	1088	1198	2371	1059	2704	1182	1097	1146	1088	1146	2635,00	1115	2648
#9	1015	1091	1148	1100	1210	2766	1184	2641	1212	1063	1148	1086	1173	2723,00	1110	2755
#10	1119	1114	1139	1110	1149	2702	1154	2392	1144	1010	1152	1083	1115	2399,00	1138	2747
AVG	1117,9	1050,1	1138,0	1085,2	1170,2	2639,3	1094,2	2695,8	1170,9	1071,1	1134,0	1039,8	1125,2	2643,2	1122,7	2652,2
Overhead in %	-6,06		-4,64		125,54		146,37		-8,52		-8,31		134,91		136,23	

Table 25: Measurements for the rules A17 to A24 on the MuSAMA dataset.

Rule	A25		A26		A27		A28		A29		A30		A31		A32	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	1106	1116	1128	1119	2008	1097	1561	1109	1549,00	1508	1543	1432	1876	2291	1924	2207
#2	1132	1118	1140	1173	2050	1077	1517	1088	1521,00	1511	1539	1456	1854	2200	1914	2202
#3	1131	1130	1156	1149	1985	1098	1565	998	1545,00	1323	1543	1481	1676	2164	1921	2195
#4	1136	1134	1121	1162	1979	1108	1537	1079	1524,00	1498	1522	1511	1814	2015	1926	2187
#5	1106	1125	1102	1121	1984	1166	1543	1098	1514,00	1486	1538	1501	1900	2153	1914	2208
#6	1099	1153	1131	1129	2032	1141	1559	1124	1557,00	1535	1592	1525	1928	2190	1932	2275
#7	1112	1111	1132	1161	1980	1093	1531	1090	1542,00	1516	1563	1490	1910	2133	1902	2236
#8	1156	1171	1107	1066	1969	1088	1552	953	1546,00	1479	1554	1489	1867	2225	1938	2209
#9	1124	1158	1131	1094	1968	1063	1541	1076	1539,00	1457	1583	1482	1875	2205	1885	2012
#10	1157	1143	1105	1102	1992	1144	1581	1164	1557,00	1497	1546	1510	1892	2279	1849	1917
AVG	1125,9	1135,9	1125,3	1127,6	1994,7	1107,5	1548,7	1077,9	1539,4	1481,0	1552,3	1487,7	1859,2	2185,5	1910,5	2164,8
Overhead in %	0,89		0,20		-44,48		-30,40		-3,79		-4,16		17,55		13,31	

Table 26: Measurements for the rules A25 to A32 on the MuSAMA dataset.

from multiple sensor nodes, executes an selection and finally, a self-join is executed. Each sensor node “produces” one million tuples, which can be processed in parallel on each sensor. The remaining tuples are send to the upper node

Rule	A33		A34		A35		A36		A37		A38		A39		A40	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	1325	1101	1884	1135	2015	1869	1525	1728	1516	1799	1444	2631,00	1866	2775	1529	1589
#2	1362	1106	1721	1186	2052	1839	1513	1855	1502	1840	1420	2427,00	1852	3032	1538	1650
#3	1478	1136	1824	1141	2022	1853	1533	1880	1558	1782	1505	2621,00	1874	2703	1541	1806
#4	1469	1128	1714	1121	1945	1794	1524	1875	1525	1800	1478	2685,00	1838	2862	1538	1817
#5	1463	1127	1637	1166	1992	1837	1502	1743	1500	1757	1435	2622,00	1815	2798	1518	1767
#6	1392	1127	1782	1121	1994	1743	1535	1855	1336	1836	1455	2640,00	1813	2944	1524	1767
#7	1475	1097	1816	1101	1821	1856	1384	1840	1573	1810	1427	2608,00	1839	2997	1535	1780
#8	1497	1096	1831	1092	1743	1905	1555	1798	1587	1709	1452	2579,00	1800	3012	1548	1809
#9	1498	1142	1837	1115	1975	1867	1550	1821	1523	1740	1486	2570,00	1874	3015	1516	1805
#10	1468	1013	1840	1095	1783	1935	1516	1822	1535	1917	1495	2568,00	1844	3035	1509	1811
AVG	1442,7	1107,3	1788,6	1127,3	1934,2	1849,8	1513,7	1821,7	1515,5	1799,0	1459,7	2595,1	1841,5	2917,3	1529,6	1760,1
Overhead in %	-23,25		-36,97		-4,36		20,35		18,71		77,78		58,42		15,07	

Table 27: Measurements for the rules A33 to A40 on the MuSAMA dataset.

Rule	A41		A42		A43		A44		A45		A46		A47		A48	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	1937	1813	1527	1816	1817	3012	1425,00	2442	1475	2591	1091	1850	1115	1347	1128	2517
#2	1943	1789	1529	1772	1837	2956	1465,00	2215	1471	2618	1099	1661	1110	1274	1094	2453
#3	1990	1817	1517	1815	1829	2891	1466,00	2479	1494	2537	1089	1661	1113	1433	1142	2638
#4	2027	1774	1533	1828	1851	2984	1426,00	2534	1504	2587	1138	1859	1108	1414	1119	2555
#5	1847	1557	1479	1819	1823	2970	1433,00	2585	1537	2595	1120	1630	1113	1468	1121	2519
#6	1824	1596	1481	1838	1843	2962	1452,00	2586	1577	2664	1134	1927	1089	1497	1036	2614
#7	1989	1749	1476	1799	1826	2781	1326,00	2370	1545	2458	1143	1848	1099	1476	996	2711
#8	1985	1645	1464	1807	1845	2905	1262,00	2564	1479	2680	1116	1879	947	1437	967	2600
#9	1944	1542	1561	1865	1887	2969	1434,00	2585	1496	2590	1119	1784	944	1432	985	2431
#10	2009	1746	1485	1806	1836	2964	1411,00	2529	1498	2635	1150	1696	1039	1378	1121	2634
AVG	1949,5	1702,8	1505,2	1816,5	1839,4	2939,4	1410,0	2488,9	1507,6	2595,5	1119,9	1779,5	1067,7	1415,6	1070,9	2567,2
Overhead in %	-12,65		20,68		59,80		76,52		72,16		58,90		32,58		139,72	

Table 28: Measurements for the rules A41 to A48 on the MuSAMA dataset.

Rule	A49		A50		A51	
Measurement in ms	-	+	-	+	-	+
#1	1102	2025	1109	1497	1110	1422
#2	1026	1995	1094	1472	1095	1299
#3	1091	1949	1093	1501	1108	1295
#4	1135	1898	1097	1507	1110	1305
#5	1117	1761	1122	1483	1102	1346
#6	1148	1829	1123	1485	1106	1356
#7	1140	1720	1096	1487	1098	1331
#8	1081	1717	1091	1562	1113	1330
#9	1094	1736	1095	1488	1124	1301
#10	1089	1695	1115	1471	1096	1362
AVG	1102,3	1832,5	1103,5	1495,3	1106,2	1334,7
Overhead in %	66,24		35,51		20,66	

Table 29: Measurements for the rules A49 to A51 on the MuSAMA dataset.

Rule	K01		K02		K03		K04		K05		K06		K07		K08	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	8110	2029	53898	270013	270113	278144	336884	322876	4604	5855	2161	1595	1600	1792	849	849
#2	6620	2430	55008	270584	270517	276902	337133	324559	5025	5510	1893	1792	1862	1777	844	834
#3	6070	2219	54742	268952	262173	276755	333543	318498	5147	5686	1602	1807	1878	1750	820	834
#4	6428	2383	56947	272324	266850	277963	332977	317667	5759	5842	1705	1587	1743	1773	832	812
#5	6392	2328	57627	272538	269757	275467	331844	319708	6025	5742	1729	1761	1607	1709	764	808
#6	6566	2015	58333	273119	269580	276259	338804	316608	5883	5761	1864	1764	1633	1611	739	861
#7	5950	1876	53998	262822	267027	275306	337965	319595	5572	5927	1839	1812	1601	1552	737	765
#8	6409	2104	55342	273356	265951	275061	334529	319157	5862	5618	1845	1704	1631	1730	741	739
#9	6592	2097	57860	273023	270913	274232	333452	321550	5767	5798	1822	1688	1868	1762	715	708
#10	6523	2440	55460	269662	267351	269583	332480	315122	5248	5474	1698	1798	1813	1766	755	793
AVG	6566,0	2192,1	55921,5	270639,3	268023,2	275567,2	334961,1	319534,0	5489,2	5721,3	1815,8	1730,8	1723,6	1722,2	779,6	800,3
Overhead in %	-66,61		383,96		2,81		-4,61		4,23		-4,68		-0,08		2,66	

Table 30: Measurements for the rules K01 to K08 on the MuSAMA dataset.

Rule	K09		K10		K11		K12		K13		K14		K15		K16	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	4435	4184	3690	3953	2463	2710	2177	2896	2601	1822	1882	2566	2787	2088	2073	1832
#2	3398	3791	3588	3444	2381	2667	2403	2685	2621	1954	2037	2465	2526	2037	1862	1844
#3	3477	4263	3354	3706	2429	2408	2315	2818	2588	1818	1831	2539	2814	1973	1974	1949
#4	3503	3996	3468	3639	2490	2711	2373	2576	2872	1857	1773	2568	2683	2069	1861	2086
#5	3714	3932	3615	3856	2706	2491	2228	2609	2677	1865	1779	2573	2896	2170	2112	1977
#6	3516	4116	3341	3364	2468	2479	2328	2691	2546	1980	1993	2573	2798	1999	2031	1858
#7	3382	3897	3119	3783	2439	2541	2467	2830	2351	1810	1771	2811	2823	1873	1982	2131
#8	3239	4223	3052	3700	2588	2603	2503	2658	2596	1723	1697	2724	2617	1729	2159	1940
#9	3436	4059	3092	3567	2664	2395	2335	2745	2788	1920	1744	2481	2631	1877	1948	2126
#10	2791	4121	3657	3281	2699	2713	2384	2861	2694	1915	2013	2749	2739	1875	1994	2068
AVG	3489,1	4058,2	3397,6	3629,3	2532,7	2571,8	2351,3	2736,9	2633,4	1866,4	1852,0	2604,9	2731,4	1969,0	1999,6	1981,1
Overhead in %	16,31		6,82		1,54		16,40		-29,13		40,65		-27,91		-0,93	

Table 31: Measurements for the rules K09 to K16 on the MuSAMA dataset.

and processed by the rest of the query. For the rewritten query, an additional selection is executed on the sensors. This leads to the following complexities:

Rule	**K17**		**K18**		**K19**		**K20**		**K21**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+
#1	3640	2830	3418	3254	5759	6499	185279	187837	104671	113672
#2	3030	2704	3353	3666	5381	5990	188360	186640	102895	111024
#3	2815	2905	3451	3352	5473	6267	188561	183748	103883	113225
#4	2825	2905	3262	3213	5449	5936	185043	187687	104195	113076
#5	2852	2879	3272	3285	6011	6343	189850	183474	101074	109286
#6	3009	2886	3342	3328	5557	6136	184977	188526	109027	110781
#7	2878	2785	3312	3198	5293	5958	184221	191196	112076	111005
#8	2971	2815	3139	3392	6039	6169	176550	190825	109143	111235
#9	2924	2736	3080	3241	6484	6654	183064	191182	111375	110515
#10	2892	2887	2935	3477	6129	6646	183342	187238	110777	104594
AVG	2983,6	2833,2	3256,4	3340,6	5757,5	6259,8	184924,7	187835,3	106911,6	110841,3
Overhead in %	-5,04		2,59		8,72		1,57		3,68	

Table 32: Measurements for the rules K17 to K21 on the MuSAMA dataset.

Rule	**L01**		**L02**		**L03**		**L04**		**L05**		**L06**		**L07**		**L08**	
Measurement in ms	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
#1	1169	1270	3747	4688	3754	5210	1980	2917	991	1900	4260	5071	4002	4609	1718	2695
#2	675	1266	4152	4816	3619	5160	2000	2952	940	1823	4014	5181	3984	5182	1732	2636
#3	677	1280	4107	4697	3976	5168	2192	2894	933	1596	4063	5028	3752	5243	1710	2650
#4	662	1252	3964	4946	4112	5463	2028	2752	909	1636	4099	5072	4110	5245	1722	2527
#5	669	1495	4233	4648	4136	5126	2040	2633	1011	1772	4176	4982	3845	5213	1686	2788
#6	660	1468	4084	4931	3862	5094	2132	2905	1035	1895	4154	5035	3991	5118	1751	2494
#7	667	1415	4338	4828	4145	4824	2094	2890	1043	1910	3840	5254	4256	5147	1862	2761
#8	620	1418	4138	4947	4150	4821	2044	2504	1056	1868	3875	5156	4247	5256	1693	2703
#9	686	1411	4158	4901	4154	4349	1998	2564	1030	1952	4025	5040	4155	5209	1559	2828
#10	651	1493	3861	4898	4168	4419	1717	2721	1021	1881	3910	5136	4183	5245	1594	2683
AVG	713,6	1376,8	4078,2	4830,0	4007,6	4963,4	2022,5	2773,2	996,9	1823,3	4041,6	5095,5	4052,5	5146,7	1702,7	2676,5
Overhead in %	92,94		18,43		23,85		37,12		82,90		26,08		27,00		57,19	

Table 33: Measurements for the rules L01 to L08 on the MuSAMA dataset.

- Unmodified query: self-join + selection: $O(n^2) + O(n)$,
- Rewritten query: self-join + selection: $O(n^2) + O(n)$ + selection on lower nodes: $O(n)$,

where n is the number of tuples.

For the sake of simplicity, we only take the complexity and not the system-dependent implementation of the operators for the calculation of the runtime. The calculation sums up the cost for each operator, which consists of the amount of tuple, multiplied with the selectivity and divided by the computing power of the corresponding device there the operator is executed. For each performance test, we give an table with the parameters and the runtime result, as well as a diagram to visualize the runtime of both queries.

4.1 Number of nodes

With an increasing number of nodes, the runtime for both unmodified and rewritten query increase with the higher amount of data. Due to the parallelism of the lower nodes, the runtime of the rewritten query increases slower than the the runtime of the unmodified query. As a results, the rewritten query is faster than the unmodified query when a certain amount of nodes on the lower layer work in parallel. The results are shown in Table 34 and Figure 2 for lower nodes with 200 MHz.

number of nodes	Hz lower layer	Hz upper layer	selectivity lower layer	selectivity upper layer	runtime unmodified query	runtime rewritten query
1	200	2000	0,1	0,01	0,55	5,10
2	200	2000	0,1	0,01	1,20	5,30
3	200	2000	0,1	0,01	1,95	5,60
4	200	2000	0,1	0,01	2,80	6,00
5	200	2000	0,1	0,01	3,75	6,50
6	200	2000	0,1	0,01	4,80	7,10
7	200	2000	0,1	0,01	5,95	7,80
8	200	2000	0,1	0,01	7,20	8,60
9	200	2000	0,1	0,01	8,55	9,50
10	200	2000	0,1	0,01	10,00	10,50
11	200	2000	0,1	0,01	11,55	11,60
12	200	2000	0,1	0,01	13,20	12,80
13	200	2000	0,1	0,01	14,95	14,10
14	200	2000	0,1	0,01	16,80	15,50
15	200	2000	0,1	0,01	18,75	17,00
16	200	2000	0,1	0,01	20,80	18,60
17	200	2000	0,1	0,01	22,95	20,30
18	200	2000	0,1	0,01	25,20	22,10
19	200	2000	0,1	0,01	27,55	24,00
20	200	2000	0,1	0,01	30,00	26,00

Table 34

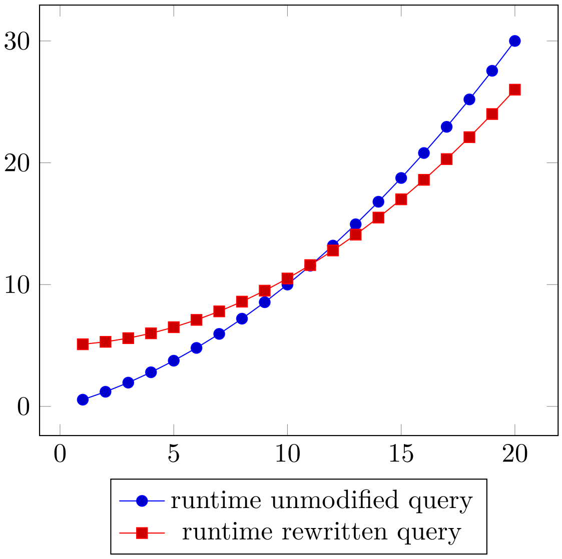


Figure 2: Runtime of the unmodified query and the rewritten query with a different number of nodes on the lower layer. Every lower node is equipped with a 200 MHz processor while the upper layer is equipped with a 2000 MHz processor. The selectivity on the lower nodes is 10%, while the selectivity on the upper node is 1%.

With an increase of the computing power of all lower nodes, the runtime decreases with a constant value for the rewritten query. The results are shown in Table 35 and Figure 3 for lower nodes with 1000 MHz. With additional 400% computing power, the runtime is increased by 80%, which is equal to 4 s in our example.

number of nodes	Hz lower layer	Hz upper layer	selectivity lower layer	selectivity upper layer	runtime unmodified query	runtime rewritten query
1	1000	2000	0,1	0,01	0,55	1,10
2	1000	2000	0,1	0,01	1,20	1,30
3	1000	2000	0,1	0,01	1,95	1,60
4	1000	2000	0,1	0,01	2,80	2,00
5	1000	2000	0,1	0,01	3,75	2,50
6	1000	2000	0,1	0,01	4,80	3,10
7	1000	2000	0,1	0,01	5,95	3,80
8	1000	2000	0,1	0,01	7,20	4,60
9	1000	2000	0,1	0,01	8,55	5,50
10	1000	2000	0,1	0,01	10,00	6,50
11	1000	2000	0,1	0,01	11,55	7,60
12	1000	2000	0,1	0,01	13,20	8,80
13	1000	2000	0,1	0,01	14,95	10,10
14	1000	2000	0,1	0,01	16,80	11,50
15	1000	2000	0,1	0,01	18,75	12,00
16	1000	2000	0,1	0,01	20,80	14,60
17	1000	2000	0,1	0,01	22,95	16,30
18	1000	2000	0,1	0,01	25,20	18,10
19	1000	2000	0,1	0,01	27,55	20,00
20	1000	2000	0,1	0,01	30,00	22,00

Table 35

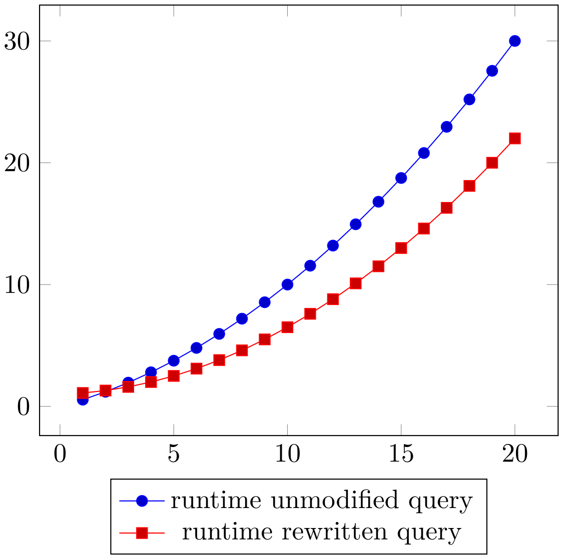


Figure 3: Runtime of the unmodified query and the rewritten query with a different number of nodes on the lower layer. Every lower node is equipped with a 1000 MHz processor while the upper layer is equipped with a 2000 MHz processor. The selectivity on the lower nodes is 10%, while the selectivity on the upper node is 1%.

4.2 Computing power

In the following, we will briefly discuss the effect on the computing power of the lower and upper nodes while the number of nodes stays the same. Table 36 and Figure 4 show, that with a linear increase in the computing power the runtime decreases logarithmically for the rewritten query. The unmodified query stays constant, because the upper layer, which performs all operations, has the same computing power all the time.

number of nodes	Hz lower layer	Hz upper layer	selectivity lower layer	selectivity upper layer	runtime unmodified query	runtime rewritten query
5	100	2000	0,1	0,01	3,75	11,50
5	200	2000	0,1	0,01	3,75	6,50
5	300	2000	0,1	0,01	3,75	4,83
5	400	2000	0,1	0,01	3,75	4,00
5	500	2000	0,1	0,01	3,75	3,50
5	600	2000	0,1	0,01	3,75	3,17
5	700	2000	0,1	0,01	3,75	2,93
5	800	2000	0,1	0,01	3,75	2,75
5	900	2000	0,1	0,01	3,75	2,61
5	1000	2000	0,1	0,01	3,75	2,50
5	1100	2000	0,1	0,01	3,75	2,41
5	1200	2000	0,1	0,01	3,75	2,33
5	1300	2000	0,1	0,01	3,75	2,27
5	1400	2000	0,1	0,01	3,75	2,21
5	1500	2000	0,1	0,01	3,75	2,17
5	1600	2000	0,1	0,01	3,75	2,13
5	1700	2000	0,1	0,01	3,75	2,09
5	1800	2000	0,1	0,01	3,75	2,06
5	1900	2000	0,1	0,01	3,75	2,03
5	2000	2000	0,1	0,01	3,75	2,00

Table 36

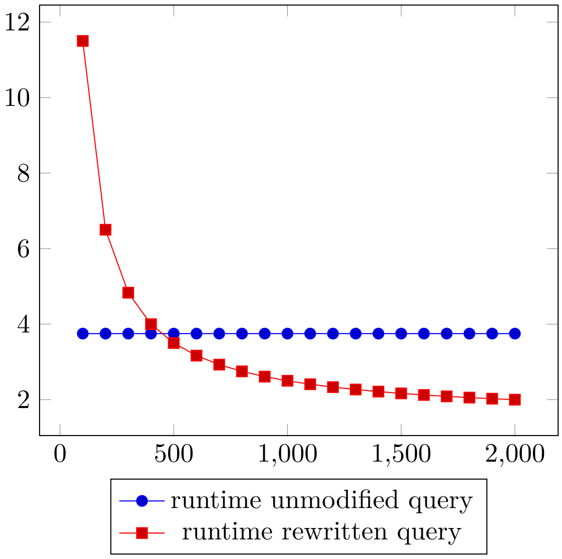


Figure 4: Runtime of the unmodified query and the rewritten query with a different computing power of the five nodes on the lower layer. Every lower node is equipped with a 100 – 2000 MHz processor while the upper layer is equipped with a 2000 MHz processor. The selectivity on the lower nodes is 10%, while the selectivity on the upper node is 1%.

In Table 37 and Figure 5 the effect of the computing power of the upper node towards the runtime is examined. With an increase of the computing power, the runtime for both rewritten and unmodified query decrease. This effect is slightly higher for the unmodified query, so that with more computing power of the upper node the unmodified query becomes faster than the rewritten query.

number of nodes	Hz lower layer	Hz upper layer	selectivity lower layer	selectivity upper layer	runtime unmodified query	runtime rewritten query
5	500	3000	0,1	0,01	2,50	3,00
5	500	2900	0,1	0,01	2,59	3,03
5	500	2800	0,1	0,01	2,68	3,07
5	500	2700	0,1	0,01	2,78	3,11
5	500	2600	0,1	0,01	2,88	3,15
5	500	2500	0,1	0,01	3,00	3,20
5	500	2400	0,1	0,01	3,13	3,25
5	500	2300	0,1	0,01	3,26	3,30
5	500	2200	0,1	0,01	3,41	3,36
5	500	2100	0,1	0,01	3,57	3,43
5	500	2000	0,1	0,01	3,75	3,50
5	500	1900	0,1	0,01	3,95	3,58
5	500	1800	0,1	0,01	4,17	3,67
5	500	1700	0,1	0,01	4,41	3,76
5	500	1600	0,1	0,01	4,69	3,88
5	500	1500	0,1	0,01	5,00	4,00
5	500	1400	0,1	0,01	5,36	4,14
5	500	1300	0,1	0,01	5,77	4,31
5	500	1200	0,1	0,01	6,25	4,50
5	500	1100	0,1	0,01	6,82	4,73

Table 37

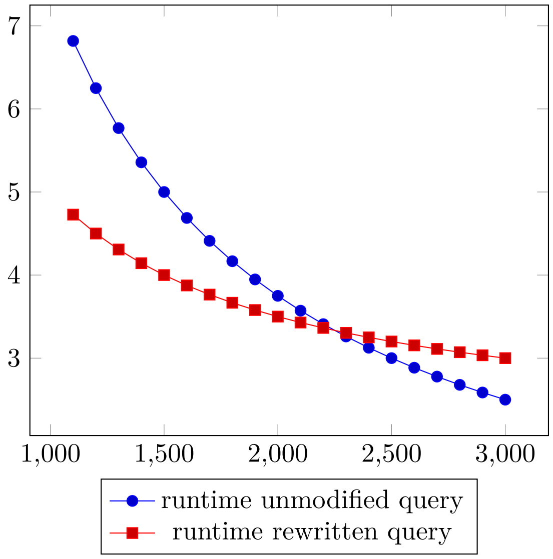


Figure 5: Runtime of the unmodified query and the rewritten query with a different computing power of the node on the upper layer. Every lower node is equipped with a 500 MHz processor while the upper layer is equipped with a processor in a 1100 to 2000 MHz range. The selectivity on the lower nodes is 10%, while the selectivity on the upper node is 1%.

4.3 Selectivity

As a last aspect, we examined the effect of the selectivity on the lower layers towards the runtime. For the evaluation, the number of nodes on the lower layer, the computing power of all nodes and the selectivity on the upper layer are fixed to constant values. Table 38 and Figure 6 show, that a higher selectivity (lower value) decreases the runtime for the rewritten query, because less data has to be processed on the upper node. The runtime of the unmodified query stays constant, because the change of the selectivity is not valid for the operators of the unmodified query.

number of nodes	Hz lower layer	Hz upper layer	selectivity lower layer	selectivity upper layer	runtime unmodified query	runtime rewritten query
5	500	2000	0,95	0,01	3,75	5,63
5	500	2000	0,90	0,01	3,75	5,50
5	500	2000	0,85	0,01	3,75	5,38
5	500	2000	0,80	0,01	3,75	5,25
5	500	2000	0,75	0,01	3,75	5,13
5	500	2000	0,70	0,01	3,75	5,00
5	500	2000	0,65	0,01	3,75	4,88
5	500	2000	0,60	0,01	3,75	4,75
5	500	2000	0,55	0,01	3,75	4,63
5	500	2000	0,50	0,01	3,75	4,50
5	500	2000	0,45	0,01	3,75	4,38
5	500	2000	0,40	0,01	3,75	4,25
5	500	2000	0,35	0,01	3,75	4,13
5	500	2000	0,30	0,01	3,75	4,00
5	500	2000	0,25	0,01	3,75	3,88
5	500	2000	0,20	0,01	3,75	3,75
5	500	2000	0,15	0,01	3,75	3,63
5	500	2000	0,10	0,01	3,75	3,50
5	500	2000	0,05	0,01	3,75	3,38
5	500	2000	0,01	0,01	3,75	3,28

Table 38

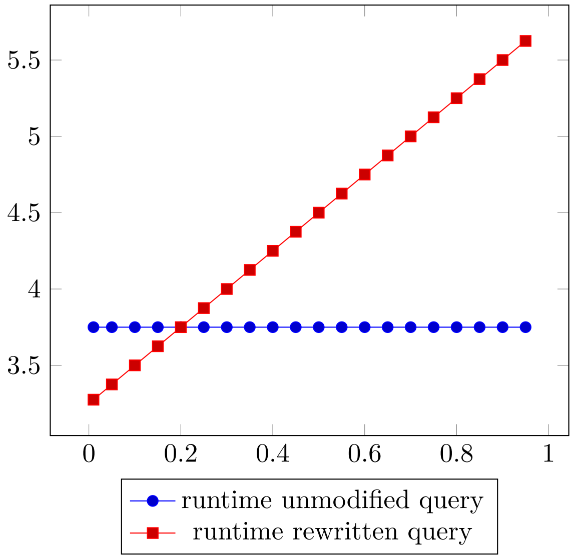


Figure 6: Runtime of the unmodified query and the rewritten query with a different selectivity of the five nodes on the lower layer. Every lower node is equipped with a 500 MHz processor while the upper layer is equipped with a 2000 MHz processor. The selectivity on the lower nodes ranges from 1% to 95%, while the selectivity on the upper node is 1%.

5 Example Query

In the motivation, we had the following SQL query:

```
SELECT Composers.name AS Composer, Albums.name AS Album
FROM Composers JOIN Tracks
ON(Composers.id=Tracks.composer)
JOIN Albums ON(Albums.id=Tracks.album)
WHERE Composers.name !=  ''
GROUP BY Albums.id, Composers.name, Albums.name
HAVING min(length) < 15*1000
AND max(length) >= 10*60*1000
```

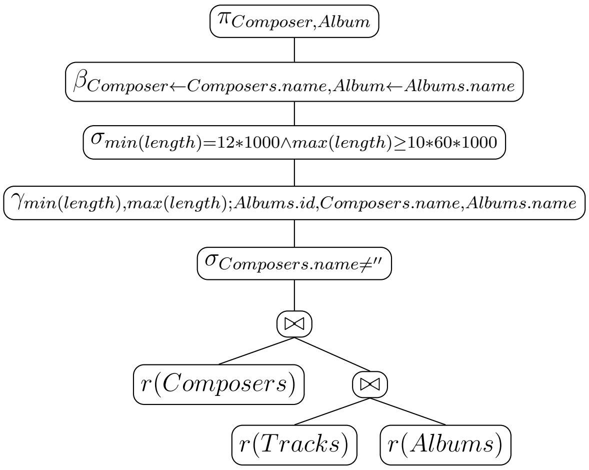

$\pi_{Composer,Album}$
$\beta_{Composer \leftarrow Composers.name, Album \leftarrow Albums.name}$
$\sigma_{min(length)=12*1000 \wedge max(length) \geq 10*60*1000}$
$\gamma_{min(length),max(length);Albums.id,Composers.name,Albums.name}$
$\sigma_{Composers.name \neq ''}$
$\bowtie$
$r(Composers)$
$\bowtie$
$r(Tracks)$
$r(Albums)$

Figure 7: Algebra tree before rewriting

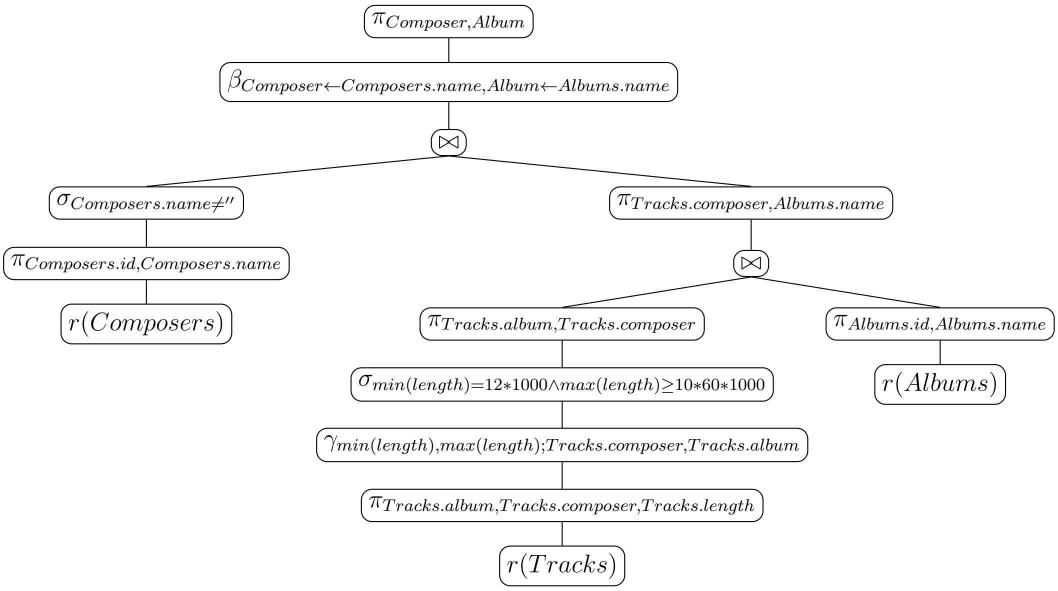

$\pi_{Composer,Album}$
$\beta_{Composer \leftarrow Composers.name, Album \leftarrow Albums.name}$
$\bowtie$
$\sigma_{Composers.name \neq ''}$
$\pi_{Tracks.composer,Albums.name}$
$\pi_{Composers.id,Composers.name}$
$\bowtie$
$r(Composers)$
$\pi_{Tracks.album,Tracks.composer}$
$\pi_{Albums.id,Albums.name}$
$\sigma_{min(length)=12*1000 \wedge max(length) \geq 10*60*1000}$
$r(Albums)$
$\gamma_{min(length),max(length);Tracks.composer,Tracks.album}$
$\pi_{Tracks.album,Tracks.composer,Tracks.length}$
$r(Tracks)$

Figure 8: Algebra tree after rewriting

The corresponding relational algbra term is:

```
$\pi_{Composer,Album}$ (
  $\beta_{Composer \leftarrow Composers.name, Album \leftarrow Albums.name}$ (
    $\sigma_{min(length)=12*1000 \wedge max(length) \geq 10*60*1000}$ (
      $\gamma_{min(length),max(length);Albums.id,Composers.name,Albums.name}$ (
        $\sigma_{Composers.name \neq ''}$ (
          $r(Composers) \bowtie r(Tracks) \bowtie r(Albums)$
        )
      )
    )
  )
)
```

As a relational algebra tree, it looks like follows:

The algebra term is a 1:1 mapping from the unoptimized SQL query. After applying optimization rules, we get the following algebra tree:

The corresponding optimized SQL query is the following:

```
SELECT DISTINCT X.name AS Composer, Z.name AS Album
FROM (
  SELECT Composers.name, Composers.id
```

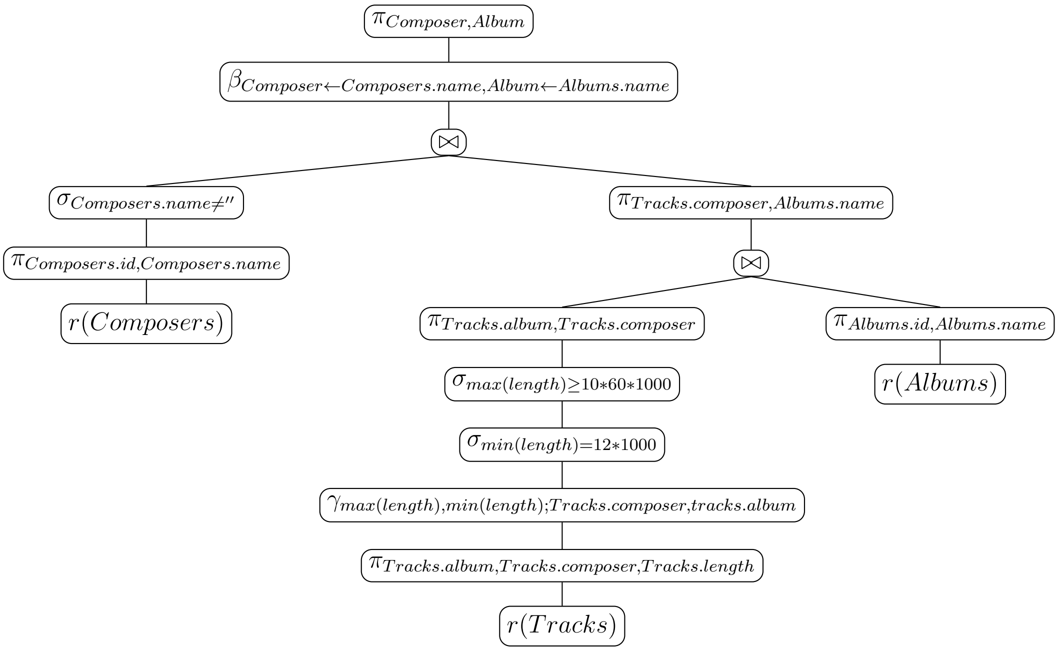

$\pi_{Composer,Album}$
$\beta_{Composer \leftarrow Composers.name, Album \leftarrow Albums.name}$
$\bowtie$
$\sigma_{Composers.name \neq ''}$
$\pi_{Tracks.composer,Albums.name}$
$\pi_{Composers.id,Composers.name}$
$\bowtie$
$r(Composers)$
$\pi_{Tracks.album,Tracks.composer}$
$\pi_{Albums.id,Albums.name}$
$\sigma_{max(length) \geq 10*60*1000}$
$r(Albums)$
$\sigma_{min(length)=12*1000}$
$\gamma_{max(length),min(length);Tracks.composer,tracks.album}$
$\pi_{Tracks.album,Tracks.composer,Tracks.length}$
$r(Tracks)$

Figure 9: Algebra tree after splitting the selection predicate into two predicates.

```
  FROM Composers
  WHERE Composers.name != ''
) AS X JOIN(
  SELECT W.composer, Y.name
  FROM (
    SELECT Tracks.album, Tracks.composer
    FROM Tracks
    GROUP BY Tracks.album, Tracks.composer
    HAVING min(length) = 12*1000
    AND max(length) >= 600*1000
  ) AS W JOIN(
    SELECT Albums.id, Albums.name
    FROM Albums
  ) AS Y ON (Y.id=W.album)
) AS Z ON (X.id=Z.composer)
```

Starting at this point, we can apply our rule set. First we split up the group selection predicate into two selection predicates:

For splitting up the tree on two layers, we have to ensure that the length is also available for the upper selection predicate. To achieve this, we have to add an additional join operator:

This results in the following SQL query with CTEs:

```
WITH A AS(
  SELECT tracks.album, tracks.composer
  FROM tracks
  GROUP BY tracks.album, tracks.composer
  HAVING min(length) = 12*1000
), B AS(
  SELECT tracks.album, tracks.composer, tracks.length
  FROM tracks
), C AS(
  SELECT A.album AS aid, A.composer as cid
  FROM A NATURAL JOIN B
  GROUP BY album, composer
  HAVING max(length) >= 10*60*1000
), D AS(
  SELECT albums.id AS aid, albums.name AS aname
  FROM albums
```

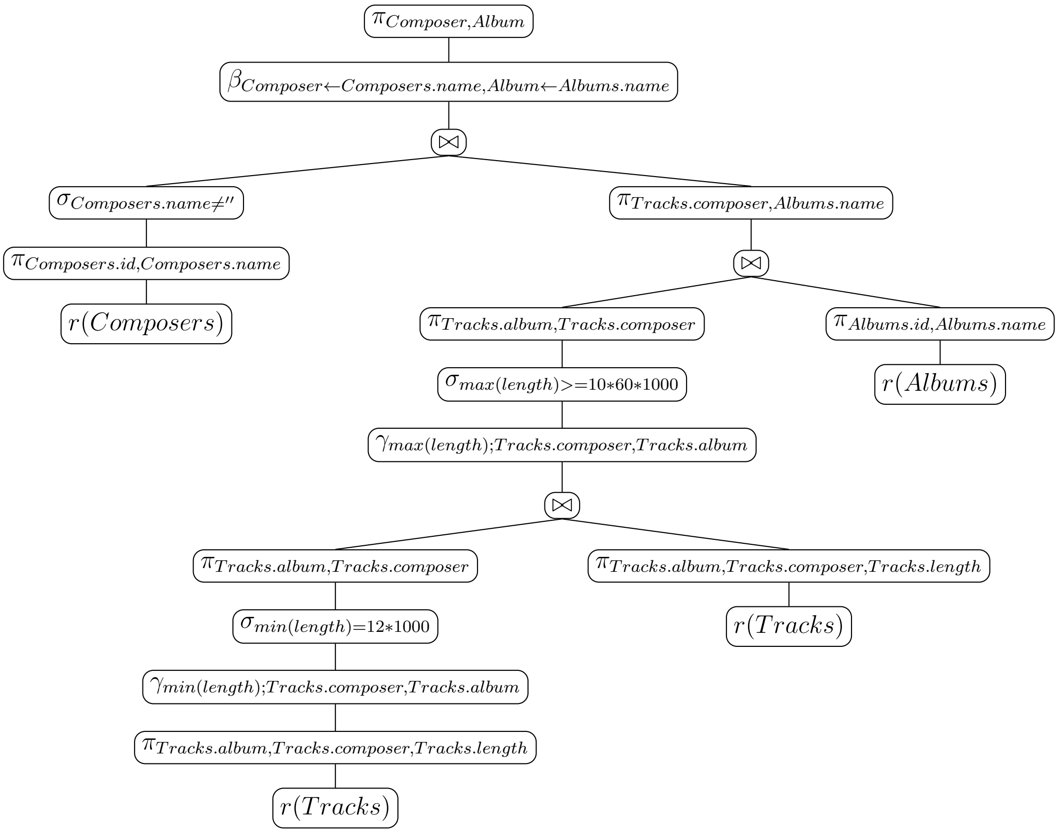

$\pi_{Composer,Album}$
$\beta_{Composer \leftarrow Composers.name, Album \leftarrow Albums.name}$
$\bowtie$
$\sigma_{Composers.name \neq ''}$
$\pi_{Tracks.composer,Albums.name}$
$\pi_{Composers.id,Composers.name}$
$\bowtie$
$r(Composers)$
$\pi_{Tracks.album,Tracks.composer}$
$\pi_{Albums.id,Albums.name}$
$\sigma_{max(length)>=10*60*1000}$
$r(Albums)$
$\gamma_{max(length);Tracks.composer,Tracks.album}$
$\bowtie$
$\pi_{Tracks.album,Tracks.composer}$
$\pi_{Tracks.album,Tracks.composer,Tracks.length}$
$\sigma_{min(length)=12*1000}$
$r(Tracks)$
$\gamma_{min(length);Tracks.composer,Tracks.album}$
$\pi_{Tracks.album,Tracks.composer,Tracks.length}$
$r(Tracks)$

Figure 10: Algebra tree after applying Contract-based rules.

```
),E AS(
  SELECT composers.id AS cid, composers.name AS cname
  FROM composers
  WHERE composers.name != ''
)
SELECT cname AS composer, aname AS album
FROM C NATURAL JOIN D NATURAL JOIN E
```

Further rules can be applied, like some of the linear arithmetic constraints. For readability, we leave the example untouched.

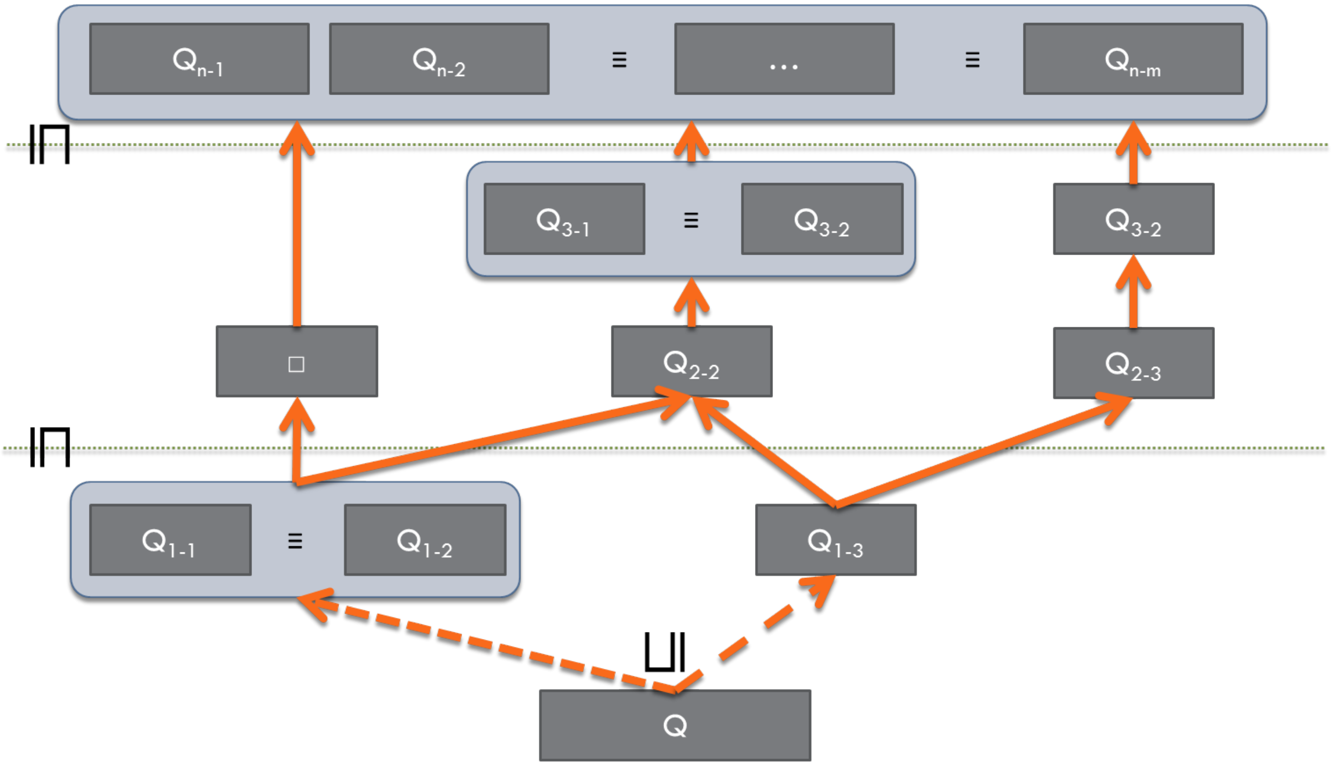


Figure 11: fig:alg

6 Algorithm for applying the rules

In general, we first apply all "classical" optimization rules to push down simple operators like projection and selection to get a good starting point for applying further rules. The relational algebra tree is traversed from the leaf nodes to the root node. When an unsupported operator is detected, we check all rules that contain this operator as a precondition (see Algorithm **??**). If multiple rules can be triggered (see Figure **??**), we choose the rule that results in the smallest overhead. This will be the smallest chain of rules that fulfills every postcondition.

Algorithm 1: Algorithm for applying multiple rules to get the rewritten query

```
Data: The original query $Q$, distributed on layers 1 . . . n
Result: The rewritten query $Q'$ on layers 1 . . . n
Determine unsupported operators as preconditions $PRE$. while $PRE \neq \emptyset$ do
    Find rule $r$ with $pre(r) \subseteq PRE$;
    $PRE = PRE/pre(r)$;
    $INV = INV \cup inv(r)$ $POST = POST \cup post(r)$ for $p \in PRE$ do
        if $unfulfilled(p)$ then
            $PRE = PRE \cup \{p\}$
        else
        end
        ..
    end
end
```

Imprint

University of Rostock
Institute of Computer Science
Database Research Group
Albert-Einstein-Straße 22
18059 Rostock

Represented by: Prof. Dr. rer. nat. habil. Andreas Heuer