An Algebraic Approach to Rule Analysis in Expert Database Systems

Elena Baralis†  Jennifer Widom

Department of Computer Science
Stanford University
Stanford, CA 94305-2140
{baralis,widom}@cs.stanford.edu

Abstract. Expert database systems extend the functionality of conventional database systems by providing a facility for creating and automatically executing Condition-Action rules. While Condition-Action rules in database systems are very powerful, they also can be very difficult to program, due to the unstructured and unpredictable nature of rule processing. We provide methods for static analysis of Condition-Action rules; our methods determine whether a given rule set is guaranteed to terminate, and whether rule execution is confluent (has a guaranteed unique final state). Our methods are based on previous methods for analyzing rules in active database systems. We improve considerably on the previous methods by providing analysis criteria that are much less conservative: our methods often determine that a rule set will terminate or is confluent when previous methods could not. Our improved analysis is based on a “propagation” algorithm, which uses a formal approach based on an extended relational algebra to accurately determine when the action of one rule can affect the condition of another. Our algebraic approach yields methods that are applicable to a broad class of expert database rule languages.

1 Introduction

In the past decade there has been a surge of interest in adding rule processing to database systems. Deductive database systems use logic rules to provide expressive query facility [9, 24]. Active database systems use Event-Condition-Action rules to provide reactive behavior [15]. In this paper we focus on what we refer to as expert database systems. An expert database system is a conventional database system extended with a facility for creating and automatically executing Condition-Action rules. Expert database systems originated by coupling a rule processor for a production rule language such as OPS5 [7] to a conventional DBMS; this approach is taken in, e.g., [23]. More recently the prevalent approach has been to build rule processing directly into the database system. Examples of recent or ongoing projects in expert database systems are [6, 11, 12, 13, 21]. Note that some systems described as active database systems actually use the Condition-Action rule paradigm, and hence fall into the class of expert database systems as we use the term here; examples of such systems are [14, 22]. Since expert database systems evolved from production rule systems such as OPS5 and are closely related to active and deductive database systems, the techniques presented in this paper certainly can be adapted for other database rule paradigms.

While expert database systems are very powerful, developing even small applications can be a difficult task, due to the unstructured and unpredictable nature of rule processing. During rule processing, rules can activate and deactivate each other, and the intermediate and final states of the database can depend on which rules are activated and executed in which order. It is highly beneficial if the rule programmer can predict in advance some aspects of rule behavior. This can be achieved by providing a facility that statically analyzes a set of rules before installing the rules in the database [1]. Static rule analysis can form the basis of a design methodology and programming environment for expert database systems.

As has been observed in the past [1, 17, 25], two important and desirable properties of rule behavior are termination and confluence. A rule set is guaranteed to terminate if, for any database state and set of modifications, rule processing cannot continue forever (i.e. rules cannot activate each other indefinitely). A rule set is confluent if, for any database state and set of modifications, the final database state after rule processing is independent of the order in which activated rules are executed.

In this paper we propose a generally applicable algorithm for determining when the action of one rule can affect the condition of another rule. The algorithm uses an extension of relational algebra to model rule conditions and actions. Essentially, the algorithm “propagates” one rule’s action through another rule’s condition to determine how the action may affect the condition; hence,
we call it the Propagation Algorithm. The Propagation Algorithm is useful for analysing termination since it can determine when one rule may activate another rule. The Propagation Algorithm also is useful for analysing confluence since it can determine when the execution order of two rules is significant. The Propagation Algorithm determines these properties much more accurately than previous methods, e.g., [1,16]. In addition, since we take a general approach based on relational algebra, our method is applicable to most expert database systems that use the relational model.

1.1 Previous Related Work

In traditional expert systems, i.e. production rule systems such as OPS5 [7], predicting properties such as termination and confluence is of less importance than in the database environment. Consequently, to our knowledge there has been little work on rule analysis in traditional expert systems.

In the database context, [16,26] give methods for analysing Condition-Action rules that are similar to the rules we consider. However, the goal of their work is to impose restrictions on rule sets so that confluence (a "unique fixed point" in their model) is guaranteed; we instead provide techniques for analysing the behavior of arbitrary rule sets. In addition, the methods in [16,26] have been shown to be weaker than the methods in [1], which in turn are weaker than the methods we present here. The methods in [1] are developed in the context of the Starburst Rule System, which uses an Event-Condition-Action (active database) rule model. Their technique for analysing rule interaction relies on a shallow comparison of the actions performed by one rule and the events and conditions of another rule. We improve on this approach significantly by using a formal algebraic model that allows us to accurately analyse the interaction between rules using the semantics of rule conditions and actions. In an initial report we applied our approach to termination only [3]; here we refine the techniques in [3] and propose a general framework for analysis of both termination and confluence.

In other related work, [25] analyses rule behavior in the context of object-oriented active database systems. Their work focuses on differences between instance-oriented and set-oriented rules (we consider only set-oriented rules in this paper) and on decidability properties for rule analysis. Their rule model is rather restricted, in that rule actions (methods) can only modify data selected by the corresponding rule condition, and deletions and insertions seem to be disallowed. The properties of confluence and of termination within some fixed number of steps are shown to be decidable using an approach based on "typical databases"; a typical database contains all possible data instances that could affect the outcome of rule processing. The rule set is "run" over the typical database and the outcome is checked for the desired properties. This approach is clearly infeasible in practical applications, so lower complexity algorithms are proposed, but the details and applicability of these algorithms are not clarified.

A rather different approach to rule analysis is taken in a recent paper [17], where Event-Condition-Action rules are reduced to term rewriting systems, and known analysis techniques for termination and confluence of term rewriting systems are applied. The rule model they use is quite different from ours, and it is unclear whether a general relational rule model such as ours can be expressed as a term rewriting system. However, in the future we plan to explore the relationship between these different approaches.

Our Propagation Algorithm is closely related to the problem of independence of queries and updates, addressed in, e.g., [18]. [18] gives an algorithm for detecting if the outcome of a query, expressed as a Datalog program, can be affected by a given insertion or deletion. For analysing expert database rules, we need a somewhat stronger technique: when a query and update are not independent, we need to know whether the update adds to, removes from, or modifies the result of the query. Furthermore, while the algorithm presented in [18] applies to more general queries than we consider here (e.g. recursive queries), their model for database updates is considerably simpler than ours.

Finally, our Propagation Algorithm is somewhat related to incremental evaluation, as in [4,19,20]: both problems address the effect of a database modification on a relational expression. However, incremental evaluation techniques are designed for run time, when the actual modifications are known, while our techniques apply at compile time, when the modifications are expressed as database operations.

1.2 Outline of the Paper

In Section 2 we present our algebraic Condition-Action rule language and provide several examples that are used throughout the paper. Section 3 contains the Propagation Algorithm, examples of its application, and a correctness Theorem. In Sections 4 and 5 we apply the algorithm to the analysis of termination and confluence, respectively; again, several examples are included. In Section 6 we draw conclusions and outline future work.

2 Algebraic Rule Language

A rule in our language has a condition and an action. Rule conditions are expressed as queries over the database; rule actions are database modifications. We use a language in which conditions and actions are both represented by relational algebra expressions. In this section we describe the extensions to relational algebra that are required to represent general rule conditions and actions. Then we specify the syntax of our rule language using this algebra, and we describe the semantics of rule processing in our model. Finally, we give several examples of how Condition-Action rules may be represented in our algebraic language.

2.1 Algebraic Operators

Based on [8], we define an extension to relational algebra that allows us to represent any queries that are express-
able in SQL (or Quel), with the exception of the handling of duplicates and ordering conditions. We also introduce an extension that allows us to represent the SQL data modification operations insert, delete, and update.

Our extended relational algebra includes the basic relational algebra operators select (σ), project (π), cross-product (×), natural join (⋈), union (∪), and difference (−), which we do not elaborate on here; see [24]. The first two lines of Table 1 present useful operators derived from the basic operators, while the next three lines present additional operators that we use. In the table, X and A denote attributes, B, A₁, and A₂ denote attribute lists, a is an aggregate function, and expr is an expression (explained below). In line 1, E₁ ×ₚ E₂ can be defined as: 

\[
E_1 \times_{p} E_2 = E_1 - (E_1 \times_{p} E_2)
\]

2.1.2 Aggregate Functions and Expression Evaluation

The attribute extension operators allow us to extend a relational expression E with a new attribute; this approach is used for aggregate functions and for modification operations. We have:

- The E operator, which computes expressions applied to each tuple of E
- The A operator, which computes aggregate functions (e.g. max, min, avg, sum, cnt) over partitions of E

E is a unary operator applied to a relational expression E producing a result with schema schema(E) ∪ \{X\}. Recall from Table 1 that the E operator is expressed as:

\[
E[X = expr]E
\]

expr is an expression evaluated over each tuple t of E (a conventional expression involving attributes of t and constants) yielding one value for each tuple; this value is entered into the new attribute X for each tuple of E. For details of similar operators see [8]; examples are given in later sections.

A is also a unary operator applied to a relational expression E producing a result with schema schema(E) ∪ \{X\}. Recall from Table 1 that the A operator is expressed as:

\[
A[X = a(A); B]E
\]

B defines a set of attributes on which the result of E is partitioned; each group in the partition contains all the tuples with the same B value. a is an aggregate function that is applied to the (multiset of) values contained in the projection of each partition on attribute A, yielding one value for each partition; this value is entered into the new attribute X for each tuple of the partition. The attributes B are optional: when B is omitted, no grouping is performed, and the aggregate function a is applied to the entire result of E, yielding one value; that value is entered into the new attribute X for each tuple of E. For details see [8].

2.1.3 Modification Operations

We represent data modification operations in relational algebra by characterising the operations in terms of the database state they produce. Table 2 presents inserts, deletes, and updates by indicating the algebraic expressions that are used to denote the operations, and the way in which these expressions are applied to a relation R to produce a new value for R. In the table, Aᵣ denotes the attributes of R that are updated, Aᵣ' denotes primed versions of these attributes (explained below), and Aᵣ = schema(R) − Aᵣ.

Insert operation. An insert operation is denoted by a relational expression E_in. E_in produces the tuples to be inserted (either a set of constant tuples or the result of an algebraic expression). The schema of E_in must coincide with the schema of R.

Delete operation. A delete operation is denoted by a relational expression E_del. E_del produces the tuples to be deleted. The schema of E_del must coincide with the schema of R.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E[x = expr]E</td>
<td>attribute extension and expression evaluation</td>
</tr>
<tr>
<td>A[X = a(A); B]E</td>
<td>attribute extension and aggregate function evaluation</td>
</tr>
<tr>
<td>E ×ₚ E</td>
<td>not-exists semijoin with predicate p</td>
</tr>
<tr>
<td>E ×ₚ E</td>
<td>semijoin with predicate p</td>
</tr>
<tr>
<td>A₁; A₂</td>
<td>attribute rename</td>
</tr>
<tr>
<td>(\mathcal{E}[X = \text{expr}])</td>
<td>attribute extension and expression evaluation</td>
</tr>
</tbody>
</table>

Table 1: Additional algebraic operators
Update operation. An update operation is denoted by a relational expression $E_{upd}$. $E_{upd}$ has schema $\text{schema}(R) \cup A'_\alpha$, where attributes $A'_\alpha$ contain the new values for the updated attributes $A_\alpha$. As convention, the new values for the updated attributes are always assigned the corresponding "primed" attribute names. That is, if attribute $A \in A_\alpha$ is updated, then the new value for $A$ is assigned to attribute $A'$. A typical way to express $E_{upd}$ is:

$$E_{upd} = E[A_{\alpha_1} = \text{expr}_1] E[A_{\alpha_2} = \text{expr}_2] \ldots E[A_{\alpha_n} = \text{expr}_n] E_c$$

where $E_c$ is an expression producing the tuples to be updated (i.e. the "selection condition" of the update operation). The schema of $E_c$ must coincide with the schema of $R$. $E[A_{\alpha_i} = \text{expr}_i]$ evaluates expression $\text{expr}_i$ on each tuple of $E_c$ and assigns the result to the new attribute $A'_{\alpha_i}$. Although this is a useful form, in its generality $E_{upd}$ can be any relational expression with schema $\text{schema}(R) \cup A'_\alpha$.

As specified in Table 2, the new state of $R$ after the update operation is the union of two terms:

1. The first term $R \times \exists E_{upd}$ includes in the result all tuples in $R$ that are not modified by the update operation.
2. The second term $\alpha_A';A_\alpha (\pi_{A_\alpha} A_\alpha E_{upd})$ includes in the result the original values for the non-updated attributes of the modified tuples and the new values for the modified attributes, with the primed attribute names replaced by the original attribute names.

Given a relational expression $E$ (say) with schema $\text{schema}(R) \cup A'_\alpha$, we often need the corresponding expression that is compatible in schema with $R$ and contains either the pre-updated (old) or the updated (new) values for the modified attributes. For convenience we will use the abbreviations $E_\text{old}(E) = \pi_{\text{schema}(E)-A'_\alpha} E$ and $E_\text{new}(E) = \alpha_{A'_\alpha};A_\alpha (\pi_{\text{schema}(E)-A'_\alpha} E)$.

### 2.2 Rule Syntax and Semantics

A Condition-Action rule in our language is defined as:

$$E_{\text{cond}} \rightarrow E_{\text{act}}$$

where:

- $E_{\text{cond}}$ states the rule's condition as an expression in our extended relational algebra.
- $E_{\text{act}}$ states the rule's action as a data modification operation expressed using $E_{\text{ins}}, E_{\text{del}},$ or $E_{\text{upd}}$ as given in Table 2.\(^1\)

\(^1\)For simplicity, we consider rules with a single action here, although many expert database systems allow rules with a sequence of actions. Our methods easily extend to multiple actions, usually simply by applying the method once for each action [2].
Example 2.1: Rule bad-account states that if an account has a balance less than 500 and an interest rate greater than 0%, then that account's interest rate is set to 0%. In our language the rule is expressed as $E_{\text{cond}} \rightarrow E_{\text{upd}}$ with:

- $E_{\text{cond}} = \pi_{\text{bal}, \text{rate}}(\sigma_{\text{bal} < 500 \land \text{rate} > 0}(\text{ACCT}))$
- $E_{\text{upd}} = E[rate' = 0]E_c$
- $E_c = \sigma_{\text{bal} < 500 \land \text{rate} > 0}(\text{ACCT})$

Example 2.2: Rule raise-rate states that if an account has an interest rate greater than 1% but less than 2%, then the interest rate is raised to 2%. In our language the rule is expressed as $E_{\text{cond}} \rightarrow E_{\text{upd}}$ with:

- $E_{\text{cond}} = \pi_{\text{rate}}(\sigma_{\text{rate} > 1 \land \text{rate} < 2}(\text{ACCT}))$
- $E_{\text{upd}} = E[rate' = 2]E_c$
- $E_c = \sigma_{\text{rate} > 1 \land \text{rate} < 2}(\text{ACCT})$

Example 2.3: Rule SF-bonus states that when the number of customers living in San Francisco exceeds 1000, then the interest rate of all San Francisco customers' accounts with a balance greater than 5000 and an interest rate less than 3% is increased by 1%. In our language the rule is expressed as $E_{\text{cond}} \rightarrow E_{\text{upd}}$ with:

- $E_{\text{cond}} = \pi_{\text{name}}(\sigma_{\text{num} > 1000}(A[C = \text{cnt}(\text{name})](\text{city} = \text{SF}\land \text{CUST})))$
- $E_{\text{upd}} = E[rate' = \text{rate} + 1]E_c$
- $E_c = (\nu_{\text{city}})(\nu_{\text{bal}}(\nu_{\text{name}}(\nu_{\text{rate}} \nu_{\text{time}}))(\nu_{\text{rate}}))$

where name abbreviates ACCT.name = CUST.name.

Example 2.4: Rule add-bad states that if an account has a balance less than 500 and is not yet recorded in the LOW-ACC relation, then the information on that account is inserted into the LOW-ACC relation, "time-stamped" with the current date. In our language the rule is expressed as $E_{\text{cond}} \rightarrow E_{\text{ins}}$ with:

- $E_{\text{cond}} = \pi_{\text{num}, \text{bal}}(\sigma_{\text{bal} < 500}(\text{ACCT})) \land \text{name}(\text{LOW-ACC})$
- $E_{\text{ins}} = E[\text{date} = \text{today}()]$

where num abbreviates ACCT.num = LOW-ACC.num and today() is a system-defined function returning the current date.

Example 2.5: Rule del-bad states that if an account in the LOW-ACC relation has a balance of at least 500 in the ACT relation, then the account is deleted from the LOW-ACC relation. In our language the rule is expressed as $E_{\text{cond}} \rightarrow E_{\text{del}}$ with:

- $E_{\text{cond}} = \pi_{\text{num}}(\text{LOW-ACC} \land \text{name}(\sigma_{\text{bal} > 500}(\text{ACCT})))$
- $E_{\text{del}} = \text{LOW-ACC} \land \text{name}(\sigma_{\text{bal} > 500}(\text{ACCT}))$

where num abbreviates LOW-ACC.num = ACCT.num.

3 The Propagation Algorithm

We describe a general algorithm that uses syntactic analysis to predict how a database query (i.e. a rule condition) can be affected by the execution of a data modification operation (i.e. a rule action). The outcome of our algorithm is zero or more of the operations insert, delete, and update, characterizing how the result of the query may change due to the execution of the modification. If the algorithm produces an insert operation, then the query may contain more data after the modification; if the algorithm produces a delete operation, then the query may contain less data after the modification; if the algorithm produces an update operation, then the query may contain updated data after the modification; if no operations are produced, then the result of the query cannot change due to the modification. The operations produced by our algorithm are represented as relational expressions in the same way that we algebraically represent data modification operations in rule actions, except here the modifications apply to arbitrary relational expressions instead of only to single relations.

The algorithm takes as input a rule condition $C$ and a rule action $A$, both expressed in extended relational algebra as defined in Section 2. As an initial filter, if the condition $C$ does not reference the relation modified by $A$, then clearly $A$ cannot affect the result of $C$. Otherwise, $A$ is "propagated" through a tree representation of $C$'s query. The leaves of the tree are relations, and one of these leaves corresponds to the relation $R$ that is modified by $A$. (For simplicity here we assume there is only one reference to $R$ in condition $C$; our method can easily be extended to handle multiple references [2].) Action $A$ is propagated from the affected relation up the query tree, and it may be transformed into one or more different actions (modification operations) during the propagation process. To describe the propagation, we give formal rules specifying how arbitrary actions are propagated through arbitrary nodes of the tree. After each propagation through a node in the tree, the actions obtained are checked for "consistency" (explained next). Inconsistent actions are discarded, while consistent actions are further propagated. The propagation process continues until the root of the query tree is reached or all actions have been discarded as inconsistent. At each point during the propagation process, the actions associated with a node $N$ in the tree indicate the actions that may occur to $N$'s subtree as a result of performing the original action $A$. Hence, the consistent actions that reach the root of the tree describe how the original action $A$ may affect condition $C$.

An action produced by the propagation process is consistent when the algebraic expression describing the action does not contain contradictions, i.e. it is satisfiable. Satisfiability of relational expressions is undecidable in the general case, so we can give sufficient but not necessary conditions for satisfiability of the expressions representing the propagated actions. However, for many expressions that arise in practice we can see trivially whether the expression is satisfiable (as in examples below), and for some classes of expressions we can verify satisfiability using the tableau method in [24].

2Note that a "conservative" test for satisfiability is not really a limitation here, since our entire approach is based on syntactic analysis and hence is conservative: when an expres-
The rules for propagation are given in tables based on the kind of incoming action: insert, delete, and update actions in Tables 3, 4, and 5 respectively. Each row in the tables contains the propagated action(s), $E_{out}$, as a function of the incoming action, $E_{in}$, and the relational operator in the query tree. The column labeled "Applicability condition" specifies when different propagation rules are used for different cases. In the tables, $A_1$, $A_2$, and $B$ are attribute lists, $A_{in} = \text{schema}(E_1) \cap \text{schema}(E_2)$, $A_{out} = \text{schema}(E_2)$, $A_u$ are the updated attributes, $A_p$ and $A_e$ are the attributes involved in predicate $p$ and expression expr respectively, $p = \alpha_{A_1:A_p} \land$ and $\text{expr}' = \alpha_{A_1:A_e} \land \text{expr}$, $p(B)$ equates all attributes in list $B$, and $p'(A_{out}B)$ equates all attributes in $A_{out}B$ with the corresponding $B$ attributes. Since the natural join, cartesian product, and union operators are symmetric, without loss of generality we assume that the first operand is modified; analogous rules apply for modifications to the second operand. Observe that aggregate functions require, in addition to the incoming action, the entire relational expression $E$ to which the aggregate function is applied.

The formulas given in Table 5 don't take into account the internal structure of selection predicates and update expressions. In the case of simple predicates (comparisons between an attribute and a constant\(^3\)) and simple arithmetic update expressions (addition or subtraction of constants from an attribute), in many cases it is possible to eliminate some of the propagated actions. For example, consider the propagation of the update $A = A + 1$ through the operation $\sigma_{A > 5}$. Intuitively, this update will never cause tuples to be deleted from the expression rooted in $\sigma_{A > 5}$. Thus, the propagated delete operation can be eliminated. Table 6 shows the actions that can be eliminated in the different cases. In the table, "other" indicates an arbitrary arithmetic expression, which in the case of an equality or non-equality predicate still allows an update action to be eliminated.

### Examples

We give two examples of the Propagation Algorithm applied to rules from Section 2.3. In each example, we analyze the effect of one rule’s action on another rule’s condition by fully describing the propagation process and the satisifability test.

#### Example 3.1

Consider condition $E_{cond}$ in rule badaccount (Example 2.1) and the update action in rule SF-bonus (Example 2.3). The input to the algorithm is:

$$C = \pi_{\text{bal, rate}}(\sigma_{\text{bal} < 500 \land \text{rate} > 0}(\text{ACCT}))$$

$$A = E_{upd} = E_{ins}$$

where name abbreviates $\text{ACCT.name} = \text{CUST.name}$ and $\text{SF-cust} abbreviates $\sigma_{\text{city} = \text{SF-CUST}}$. Using Table 5, the propagation of $E_{ins}$ through the selection operation in $C$ yields insert and update actions (the delete action is eliminated, see Table 6). We have:

### Table 6: Eliminated actions

<table>
<thead>
<tr>
<th>Arithmetic update expression</th>
<th>Addition</th>
<th>Subtraction</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_u = h, A_u \neq h$</td>
<td>$E_{ins}$</td>
<td>$E_{del}$</td>
<td>$E_{ins}$</td>
</tr>
<tr>
<td>$A_u &gt; h, A_u \geq h$</td>
<td>$E_{ins}$</td>
<td>$E_{ins}$</td>
<td>$E_{ins}$</td>
</tr>
<tr>
<td>$A_u &lt; h, A_u \leq h$</td>
<td>$E_{ins}$</td>
<td>$E_{del}$</td>
<td>$E_{ins}$</td>
</tr>
</tbody>
</table>
### Table 4: Delete action propagation

<table>
<thead>
<tr>
<th>Node</th>
<th>Applicability condition</th>
<th>Resulting expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_{A}E$</td>
<td>$E_{del}^{a} = \pi_{A}E_{del}^{a}$</td>
<td>$E_{del}^{a}$</td>
</tr>
<tr>
<td>$E_{1} \bowtie E_{2}$</td>
<td>$E_{del}^{a} = E_{del}^{a} \bowtie E_{del}^{a}$</td>
<td>$E_{del}^{a}$</td>
</tr>
<tr>
<td>$E_{1} \times E_{2}$</td>
<td>$E_{del}^{a} = E_{del}^{a} \times E_{del}^{a}$</td>
<td>$E_{del}^{a}$</td>
</tr>
<tr>
<td>$E_{1} \cup E_{2}$</td>
<td>$E_{del}^{a} = E_{del}^{a} \cup E_{del}^{a}$</td>
<td>$E_{del}^{a}$</td>
</tr>
<tr>
<td>$E_{1} \rhd E_{2}$</td>
<td>$E_{del}^{a} = E_{del}^{a} \rhd E_{del}^{a}$</td>
<td>$E_{del}^{a}$</td>
</tr>
<tr>
<td>$E_{1} \bowtie_{E} E_{2}$</td>
<td>$E_{del}^{a} = E_{del}^{a} \bowtie_{E} E_{del}^{a}$</td>
<td>$E_{del}^{a}$</td>
</tr>
</tbody>
</table>

In both cases, predicates bal < 500 and bal > 5000 (the latter from $E_{upd}$) are contradictory, so both expressions $E_{ins}^{\prime}$ and $E_{upd}^{\prime}$ are unsatisfiable. Intuitively, action $A$ operates on data not read by condition $C$. We conclude that action $A$ cannot affect condition $C$.

**Example 3.2**: Consider condition $E_{cond}$ in rule badaccount (Example 2.1) and the update action in rule raise-rate (Example 2.2). The input to the algorithm is:

\[
C = \pi_{bal,rate}(\sigma_{bal<500 \land rate>0}ACCT) \\
A = E_{upd} = \xi[rate = 2](\sigma_{rate>1}rate<2ACCT)
\]

The propagation of $E_{upd}$ through the selection operation in $C$ yields insert, delete, and update actions:

\[
E_{ins}^{\prime} = \rho_{new}(\sigma_{bal<500 \land rate>0}E_{upd}) \bowtie_{E} \\
E_{ins} = (\sigma_{bal<500 \land rate>0}E_{upd}) \bowtie_{E} \\
E_{del}^{\prime} = \rho_{del}(\sigma_{bal<500 \land rate>0}E_{upd}) \bowtie_{E} \\
E_{del} = (\sigma_{bal<500 \land rate>0}E_{upd}) \bowtie_{E} \\
E_{upd} = (\sigma_{bal<500 \land rate>0}E_{upd}) \bowtie_{E}
\]

These expressions do not contain contradictory predicates, thus they may be satisfiable and the propagation continues. The propagation of $E_{ins}^{\prime}$, $E_{del}^{\prime}$, and $E_{upd}^{\prime}$ through the projection operation in $C$ yields:

\[
E_{ins}^{\prime} = \pi_{bal,rate}E_{ins}^{\prime} \\
E_{del}^{\prime} = \pi_{bal,rate}E_{del}^{\prime} \\
E_{upd}^{\prime} = \pi_{bal,rate}E_{upd}^{\prime}
\]

and $E_{upd}^{\prime}$ describe the actions that can be performed on $C$ as a result of the execution of $A$.

### 3.2 Correctness of the Algorithm

The following Theorem states the correctness of the Propagation Algorithm. Due to space constraints, the proof of the Theorem is omitted here. A proof sketch appears in [5] and a complete proof appears in [2].

**Theorem 3.1**: Let $Q$ be the query tree corresponding to a relational expression $C$ and let $A$ be an action performed on a relation in $Q$. Let $E_{out}$ be the actions produced at the root of $Q$ by application of the propagation rules in Tables 3-5. $E_{out}$ describes a superset of all actions that can be performed on expression $C$ as a result of executing the original action $A$.

### 4 Termination Analysis

Recall the rule processing loop from Section 2.2. Termination for a rule set is guaranteed if rule processing always reaches a state in which no rule has a true condition. Notice that, according to the semantics in Section 2.2, after the first execution of each rule $r$, $r$'s condition is true again if and only if new data satisfies the condition. Hence, informally, rule processing does not terminate if and only if rules provide new data to each other indefinitely.

We say that a rule $r_1$ may activate a rule $r_2$ if executing $r_1$'s action may cause new data to satisfy $r_2$'s condition. We analyze termination by building an Activation Graph. In the graph, nodes represent rules, and directed edges indicate that one rule may activate the other. If there are no cycles in the graph, then rule processing is guaranteed to terminate [1,3]. Hence, the core of termination analysis is determining when an edge should be included in the graph, i.e. when one rule may activate
### Table 5: Update action propagation

<table>
<thead>
<tr>
<th>( A \cap B = A \cap B )</th>
<th>( A \cap B = 0 )</th>
<th>( A = B )</th>
<th>( A = 0 )</th>
<th>( B = 0 )</th>
<th>( A \cap B = 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_{A \cup B} = (E_{A} \cup E_{B}) )</td>
<td>( E_{A \cap B} = (E_{A} \cap E_{B}) )</td>
<td>( E_{A} = E_{A} )</td>
<td>( E_{B} = E_{B} )</td>
<td>( E_{A} = E_{A} )</td>
<td>( E_{B} = E_{B} )</td>
</tr>
</tbody>
</table>

Note: Propagated actions for each condition are listed in the rightmost column.
another rule. The more accurately we can make this decision, the more accurately we can analyze termination.

We use our Propagation Algorithm to decide when an edge \( r_1 \rightarrow r_2 \) belongs in the Activation Graph. Note that rules may activate themselves, so \( r_1 = r_2 \) is included in the analysis. To detect if \( r_1 \) may activate \( r_2 \), the Propagation Algorithm is applied to \( r_2 \)'s condition \( C \) and \( r_1 \)'s action \( A \). If the algorithm yields an insert or update operation, then \( r_1 \) may provide new data satisfying \( r_2 \)'s condition. Thus, \( r_1 \) may activate \( r_2 \), and the edge \( r_1 \rightarrow r_2 \) belongs in the graph. If only a delete operation or no operation is produced by the algorithm, then \( r_1 \) cannot provide new data to \( r_2 \)'s condition, and the edge is not included in the graph.

Our use of the Activation Graph is similar to, e.g., [1, 10], but our approach is far less conservative since we exploit the algebraic structure of conditions and actions to accurately determine when edges belong in the graph.

4.1 Examples

Consider the rules from Section 2.3. We present two examples where we apply our analysis techniques to determine that a pair of rules does not produce a cycle in the Activation Graph, i.e. the rules cannot activate each other indefinitely. In both of these examples, the technique in [1] is unable to determine that these rules terminate.

Example 4.1: Consider rule bad-account (Example 2.1) and rule raise-rate (Example 2.2) that here will be called \( r_1 \) and \( r_2 \) respectively. Both rule conditions reference attribute rate, and both rule actions update rate. Hence, intuitively (and according to the method in [1]), the two rules might activate each other indefinitely. We have shown in Example 3.2 that \( r_2 \)'s action may provide data to \( r_1 \)'s condition (since insert and update operations are produced by the Propagation Algorithm), thus the edge \( r_2 \rightarrow r_1 \) belongs in the Activation Graph. Now we use the Propagation Algorithm to determine if \( r_1 \) may activate \( r_2 \). The input to the algorithm is:

\[
C = \pi_{\text{rate}}(\sigma_{\text{rate}>1} \land \text{rate}<2ACCT)
\]

\[
A = E_{\text{upd}} = E[\text{rate}' = 0](\sigma_{\text{bal}<500} \land \text{rate}>0ACCT)
\]

The propagation of \( E_{\text{upd}} \) through the selection operation in \( C \) yields:

\[
E'_{\text{ins}} = \rho_{\text{new}}((\sigma_{\text{rate}>1} \land \text{rate}<2E_{\text{upd}}) \times y)
\]

\[
E'_{\text{upd}} = (\sigma_{\text{rate}>1} \land \text{rate}<2E_{\text{upd}}) \times M
\]

\[
E'_{\text{del}} = \rho_{\text{del}}((\sigma_{\text{rate}>1} \land \text{rate}<2E_{\text{upd}}) \times y)
\]

Since predicates rate' > 1 and rate' = 0 (the latter from \( E_{\text{upd}} \)) are contradictory, expressions \( E'_{\text{ins}} \) and \( E'_{\text{upd}} \) are not satisfiable and hence are discarded. The propagation of \( E'_{\text{del}} \) through the projection operation in \( C \) yields:

\[
E'_{\text{del}} = \pi_{\text{rate}}E'_{\text{del}}
\]

which is satisfiable. Thus, \( r_1 \)'s action may result in a deletion of tuples from \( r_2 \)'s condition. However, since neither an insert nor an update action is produced, \( r_1 \) cannot activate \( r_2 \), the edge \( r_1 \rightarrow r_2 \) is not included in the Activation Graph, and we conclude that rules \( r_1 \) and \( r_2 \) will always terminate.

Example 4.2: Consider rule add-bad (Example 2.4) and rule del-bad (Example 2.5) that here will be called \( r_1 \) and \( r_2 \) respectively. Here again, intuitively (and according to the method in [1]), the two rules might activate each other indefinitely. We use the Propagation Algorithm to determine if \( r_1 \) may activate \( r_2 \). The input to the algorithm is:

\[
C = \pi_{\text{num}}(\text{LOW-ACCT}) \land \pi_{\text{num}}(\sigma_{\text{bal}<500}\text{ACCT})
\]

\[
A = E_{\text{ins}} = E[\text{date} = \text{today}]((\pi_{\text{num}}.,(\pi_{\text{num}}.,(\pi_{\text{num}}.,(\pi_{\text{num}}.,\text{low-bal})<\text{num},\text{low-bal}))<\text{num},\text{high-bal})
\]

where num abbreviates low-bal < 500ACCT and high-bal abbreviates low-bal < 500ACCT. This expression is not satisfiable, since it requires a tuple with a given num value to satisfy both predicates bal < 500 and bal ≥ 500. Hence, \( r_1 \) cannot activate \( r_2 \), edge \( r_1 \rightarrow r_2 \) is not included in the Activation Graph, and rules \( r_1 \) and \( r_2 \) are guaranteed to terminate.

5 Confluence Analysis

Recall again the rule processing loop from Section 2.2. In each iteration, there may be multiple rules eligible for execution, since more than one rule may have a true condition. A rule set is confluent if the final state of the database does not depend on which eligible rule is chosen for execution at any iteration.

To formally describe confluence and confluence analysis, we introduce the notion of a rule execution state and a rule execution sequence. Let \( R \) be the set of rules under consideration.

Definition 5.1: A rule execution state \( S \) is a pair \((db, RA)\), where \( db \) is a state of the database and \( RA \subseteq R \) is a set of activated rules. □

Definition 5.2: A rule execution sequence is a sequence \( \sigma \) consisting of a series of rule execution states linked by (executed) rules. A rule execution sequence is complete if the last state is \((db, \emptyset)\), i.e. the last state has no activated rules. A rule execution sequence is valid if it represents a correct execution sequence: only activated rules are executed, and pairs of adjacent states properly represent the effect of executing the corresponding rule; for details see [1,2]. □

We now define confluence in terms of execution sequences.
Definition 5.3: A rule set is confluent if, for every initial rule execution state $S$ (corresponding to an initial database and set of modifications), every valid and complete rule execution sequence beginning with $S$ has the same final state.

Clearly we cannot use this definition directly to analyze confluence, since it requires the exhaustive verification of all possible execution sequences for all possible initial states. We give sufficient conditions for confluence based on the commutativity of rule pairs. Two rules $r_i$ and $r_j$ commute if, starting with any execution state $S$, executing $r_i$ followed by $r_j$ produces the same execution state as executing $r_j$ followed by $r_i$. The conditions for commutativity require that the two rules cannot activate or deactivate each other, and their actions can be executed in either order. These conditions are stated in the following Lemma, whose proof is obvious [1].

Lemma 5.1: Distinct rules $r_i$ and $r_j$ commute if: (1) $r_i$'s action cannot affect the outcome of $r_j$'s condition (i.e., $r_i$ can neither activate nor deactivate $r_j$); (2) executing $r_i$'s action cannot change the effect of executing $r_j$'s action; (3) conditions (1) and (2) with $i$ and $j$ reversed. □

Note that even though conditions (1)–(3) are not necessarily satisfied when $r_i = r_j$, it is the case that a rule always commutes with itself.

We state two Lemmas, followed by the main Theorem on confluence. The proofs are omitted due to space constraints; all proofs appear in [5]. The first Lemma states, under the assumption of commutative rules, that two execution sequences with the same initial state and executed rules have the same final state; the second Lemma states, again under the assumption of commutativity, that two sequences with the same initial state must have the same executed rules.

Lemma 5.2: Let all pairs of rules in $R$ commute. Let $\sigma_1$ and $\sigma_2$ be two valid and complete rule execution sequences with the same initial state, such that the same rules are executed in $\sigma_1$ and $\sigma_2$ although not necessarily in the same order. Then $\sigma_1$ and $\sigma_2$ have the same final state. □

Lemma 5.3: Let all pairs of rules in $R$ commute. Let $\sigma_1$ and $\sigma_2$ be two valid and complete rule execution sequences with the same initial state. Then the same rules are executed in $\sigma_1$ and $\sigma_2$. □

Based on these Lemmas, the following Theorem presents a sufficient condition to guarantee confluence of a rule set.

Theorem 5.1: A rule set $R$ is confluent if all pairs of rules in $R$ commute. □

The requirement for confluence in Theorem 5.1 may seem rather strong, but there is no way to weaken this requirement in a rule model without a more sophisticated conflict resolution policy or priorities among rules. We believe this argues for the importance of rule priorities, which we plan to investigate in this context as future work. Notice also that, in the case where no rule can activate itself, the confluence requirement as stated in Theorem 5.1 trivially implies termination, since the pairwise commutativity of all rules includes the requirement that no rule activates another rule. However, if one or more rules can activate themselves, then confluence does not imply termination.

Commutativity of rule pairs forms the basis of most methods for analyzing confluence of database rules, e.g. [1,25]. The remainder of this section describes our technique for determining commutativity of rule pairs. Since commutativity itself is a "subroutine" to proving confluence, our commutativity analysis technique also can be applied in other contexts, e.g. [1]. Needless to say, we use our Propagation Algorithm to analyze commutativity, exploiting the algebraic description of rule conditions and actions to yield a much more accurate analysis technique than, e.g., [1].

To guarantee commutativity of two rules $r_i$ and $r_j$, we must verify conditions (1), (2), and (3) in Lemma 5.1. For (1), we determine that $r_i$ cannot activate $r_j$ exactly as we have done for termination; recall Section 4. To show that $r_i$ cannot deactivate $r_j$, we must show that $r_i$'s action $A$ cannot "take away" data from $r_j$'s condition $C$. It is easy to see that action $A$ can take away data from condition $C$ only if the Propagation Algorithm applied to $A$ and $C$ produces a delete operation. Hence, one application of the Propagation Algorithm is sufficient for verifying (1).

For (2), we must determine if $r_i$'s action $A_i$ can change the effect of $r_j$'s action $A_j$. We do this by transforming action $A_j$ into a condition $C_j$ such that if the result of condition $C_j$ cannot be affected by the execution of $A_i$, then $A_i$ cannot change the effect of action $A_j$. We then apply the Propagation Algorithm to analyze $A_i$ and $C_j$; if the algorithm produces $\emptyset$, then $A_i$ cannot change the effect of $A_j$; if the algorithm produces one or more of insert, delete, or update, then $A_i$ may change the effect of $A_j$.

Consider how condition $C_j$ is derived from action $A_j$. If $A_j$ is an insert operation, then $A_j = E_{ins}$ is a condition describing the inserted data, hence we let $C_j = E_{ins}$. Similarly, if $A_j$ is a delete operation, then $A_j = E_{del}$ is a condition describing the deleted data, and we let $C_j = E_{del}$. Suppose $A_j$ is an update operation on an attribute $A_j$ defined by $E_{upd} = A[A' - expr]E_C$. We start with the "selection condition" $E_C$. $C_j$ is the projection of $E_C$ onto all attributes referenced within $E_C$ together with all attributes referenced in the $E$ operation (both $A$ and the attributes referenced in $expr$). If any of these attributes can be affected by the execution of $A_i$, then $A_i$ may change the effect of $A_j$'s update; if not, then

---

4Rule $r_i$ deactivates rule $r_j$ if $r_i$'s action deletes all new data satisfying $r_j$'s condition.

5The extension to multiple updated attributes is obvious. Note that here we require updates to be specified with explicit use of $E_C$; this does not limit expressiveness.
conclude that rules \( r_i \) and \( rz \)'s action cannot affect \( r_i \)'s condition. Hence, we analyze \( r_i \)'s action, and we have already shown in Example 3.1 that these two rules are confluent. In both of these examples, the technique in [1] is unable to determine that these rules commute.

**Example 5.1:** Consider rule bad-account (Example 2.1) and rule SF-bonus (Example 2.3), that here will be called \( r_1 \) and \( r_2 \) respectively. Both rules reference attribute \textit{rate} and both update this attribute. Hence, intuitively (and according to the method in [1]), the two rules may not commute. We first analyze the effect of \( r_1 \)'s action on \( r_2 \). Since \( r_2 \)'s condition does not reference the relation updated by \( r_1 \), \( r_1 \)'s action trivially cannot affect \( r_2 \)'s condition. We use the Propagation Algorithm to analyze the effect of \( r_1 \)'s action on the condition corresponding to \( r_2 \)'s action: \( \pi_{\text{bal}, \text{rate}, \text{name}, \text{city} \rightarrow \text{EC}} \). The input to the algorithm is:

\[
C = \pi_{\text{bal} \geq 5000 \land \text{rate} < 3} (\text{\textit{ACCT} \leftarrow \text{\textit{name}(\text{\textit{city} = \text{\textit{SF}_CUST}})))
\]

\[
A = E_{upd} = [\text{\textit{rate'}} = 0] (\pi_{\text{bal} < 5000 \land \text{rate} > 0} \text{\textit{ACCT}})
\]

The propagation of \( E_{upd} \) through the semijoin operation in \( C \) yields:

\[
E'_{upd} = E_{upd} \bowtie_{\text{\textit{name}}} (\pi_{\text{\textit{city} = \text{\textit{SF}_CUST}}})
\]

The propagation of \( E'_{upd} \) through the selection operation in \( C \) yields:

\[
E''_{upd} = \rho_{\text{\textit{new}}}(\pi_{\text{bal} > 5000 \land \text{rate'} < 3} E'_{upd}) \bowtie_{\text{\textit{day}}}
\]

\[
E''_{del} = \rho_{\text{\textit{old}}}(\pi_{\text{bal} = 5000 \land \text{rate'} < 3} E'_{upd}) \bowtie_{\text{\textit{day}}}
\]

\[
E''_{upd} = (\pi_{\text{bal} = 5000 \land \text{rate'} < 3} E'_{upd}) \bowtie_{\text{\textit{day}}}
\]

In all three expressions, predicates \( \text{bal} > 5000 \) and \( \text{bal} < 5000 \) (the latter from \( E_{upd} \)) are contradictory, so the expressions are unsatisfiable. Hence, the Propagation Algorithm produces no actions and we conclude that executing \( r_1 \)'s action cannot change the effect of \( r_2 \)'s action.

A similar analysis reveals that \( r_2 \)'s action cannot affect \( r_1 \)'s action, and we have already shown in example 3.1 that \( r_2 \)'s action cannot affect \( r_1 \)'s condition. Hence, we conclude that rules \( r_1 \) and \( r_2 \) commute.

**Example 5.2:** Consider rule add-bad (Example 2.4) and rule del-bad (Example 2.5) that here will be called \( r_1 \) and \( r_2 \) respectively. We have already shown in Example 4.2 that \( r_1 \)'s action cannot affect \( r_2 \)'s condition. An analogous analysis shows that \( r_1 \)'s action cannot affect the condition corresponding to \( r_2 \)'s action, i.e. rule \( \text{del-bad} \). Consider the effect of rule \( r_2 \) on rule \( r_1 \). We first apply the Propagation Algorithm to \( r_2 \)'s action and \( r_1 \)'s condition. The input to the algorithm is:

\[
C = \pi_{\text{num,bal}}(\pi_{\text{bal} < 5000} \leftarrow \text{\textit{low-bal}} \text{\textit{ACC}}) \leftarrow \text{\textit{num}} \text{\textit{LOW-ACC}})
\]

\[
A = E_{del} = \text{\textit{LOW-ACC}} \leftarrow \text{\textit{num}} (\pi_{\text{bal} > 5000} \text{\textit{ACCT}})
\]

where \( \text{\textit{num}} \) abbreviates \( \text{\textit{LOW-ACC}} \leftarrow \text{\textit{num}} \text{\textit{high-bal}} \). The propagation of \( E_{del} \) through the \( \bowtie_{\text{\textit{day}}} \) operation in \( C \) yields:

\[
E''_{ins} = \text{\textit{low-bad}} \leftarrow \text{\textit{num}} (\text{\textit{LOW-ACC}} \leftarrow \text{\textit{num}} \text{\textit{high-bal}})
\]

where \( \text{\textit{low-bal}} \) abbreviates \( \sigma_{\text{\textit{bal} < 5000}} \text{\textit{ACCT}} \) and \( \text{\textit{high-bal}} \) abbreviates \( \sigma_{\text{\textit{bal} > 5000}} \text{\textit{ACCT}} \). This expression is not satisfiable, as it requires a tuple with a given num value to satisfy both predicates \( \text{bal} < 5000 \) and \( \text{bal} > 5000 \). Hence, \( r_2 \)'s action cannot affect \( r_1 \)'s condition. An analogous analysis shows that \( r_2 \)'s action cannot affect the condition corresponding to \( r_1 \)'s action, i.e. rule \( \text{del-bad} \). Hence, we conclude that rules \( r_1 \) and \( r_2 \) commute.

**6 Conclusions and Future Work**

We have defined a representation of Condition-Action expert database rules based on an extended relational algebra, and we have described a generally applicable algorithm for analyzing the interactions between one rule's condition (a query) and another rule's action (a modification). We have shown how this algorithm is applied to check termination and confluence for sets of rules. Our technique improves considerably upon previous methods, because our formal approach allows us to exploit the semantics of conditions and actions to analyze the interaction between rules. Note that the methods we describe also are applicable to rule languages that “pass data” from the condition to the action (e.g. [13,14]), since our algorithm detects the actual modifications to rule conditions (inserts, deletes, and updates), not simply the transition between true and false. As in [1], our analysis techniques identify the responsible rules when termination or confluence is not guaranteed; hence, our techniques can be used as the kernel of an interactive development tool that helps rule definers develop sets of rules that are guaranteed to have the desired properties.

We plan to extend our rule model and analysis techniques to incorporate additional features of expert database rules:

- **Rule priorities and conflict resolution.** Priorities restrict the possible execution sequences of rules, making analysis more complex but perhaps more precise. Coupling our accurate analysis of rule interactions with the priority-based methods in [1] should immediately produce a quite powerful analysis method for prioritized rules.

- **Different semantics for rule condition evaluation.** In some database rule languages, rule conditions may be evaluated over the entire database, as opposed to considering only “new” data as we
have done here. This interpretation yields a different notion of rule activation, since a rule condition remains true unless execution of some rule action renders it false.

- Events. We can handle Event-Condition-Action rules that have a semantics similar to our Condition-Action rules, e.g. the event-based rules of Ariel [14], with minor modifications to our techniques. (In fact, it is our feeling that event-based rules often are programmed this way in practice, e.g. [10].) However, general Event-Condition-Action rules, especially those in which the condition is evaluated over the entire database, will require a redefinition of rule activation (as discussed in the previous point), along with corresponding modifications to our method.

We also hope to use our algebraic rule model and Propagation Algorithm as the basis for compile-time and run-time optimisations to rule processing.

Acknowledgements

Thanks to the members of the Stanford Database Group, especially Ashish Gupta and Jeff Ullman, for lively and useful discussions, and to Stefano Ceri for providing the technical impetus and enabling the collaboration.

References


