Noname manuscript No.
(will be inserted by the editor)

Jan Chomicki

Database Querying under Changing Preferences

Received: date / Accepted: date

Abstract We present here a formal foundation for an iterative and incremental approach to constructing and evaluating preference queries. Our main focus is on *query modification*: a query transformation approach which works by revising the preference relation in the query. We provide a detailed analysis of the cases where the order-theoretic properties of the preference relation are preserved by the revision. We consider a number of different revision operators: union, prioritized and Pareto composition. We also formulate algebraic laws that enable incremental evaluation of preference queries. Finally, we consider two variations of the basic framework: finite restrictions of preference relations and weak-order extensions of strict partial order preference relations.

Keywords preference queries · preference revision · query evaluation · strict partial orders · weak orders

1 Introduction

The notion of *preference* is common in various contexts involving decision or choice. Classical utility theory (Fis70) views preferences as *binary relations*. This view has recently been adopted in database research (Cho02; Cho03; Kie02; KK02), where preference relations are used in formulating *preference queries*. In AI, various approaches to compact specification of preferences have been explored (BBD$^+$04). The semantics underlying such approaches typically relies on preference relations between worlds.

Research supported by NSF grant IIS-0307434. This paper is an expanded version of (Cho06).

Jan Chomicki
Dept. of Computer Science and Engineering, University at Buffalo, Buffalo, NY 14260-2000
Tel.: +716-645-3180 x103
Fax: +716-645-3464
E-mail: chomicki@cse.buffalo.edu

Preferences can be embedded into database query languages in several different ways. First, (Cho02; Cho03; Kie02; KK02) propose to introduce a special operator *"find all the most preferred tuples according to a given preference relation."* This operator is called *winnow* in (Cho02; Cho03). A special case of winnow is called *skyline* (BKS01) and has been recently extensively studied (PTFS03; BGZ04). Second, (AW00; HP04) assume that preference relations are defined using numeric utility functions and queries return tuples ordered by the values of a supplied utility function. It is well-known that numeric utility functions cannot represent all strict partial orders (Fis70), not even those that occur in database applications in a natural way (Cho03). For example, utility functions cannot capture skylines. Also, ordered relations go beyond the classical relational model of data. The evaluation and optimization of queries over such relations requires significant changes to relational query processors and optimizers (ISWGA04). On the other hand, winnow can be seamlessly combined with any relational operators.

We adopt here the first approach, based on winnow, within the preference query framework of (Cho03) (a similar model was described in (Kie02)). In this framework, preference relations between tuples are defined by first-order logical formulas.

Example 1 Consider the relation $Car(Make, Year)$ and the following preference relation $\succ_{C_1}$ between *Car* tuples:

within each make, prefer a more recent car,

which can be defined as follows:

$$(m,y) \succ_{C_1} (m',y') \equiv m = m' \wedge y > y'.$$

The winnow operator ω_{C_1} returns for every make the most recent car available. Consider the instance r_1 of *Car* in Figure 1a. The set of tuples $\omega_{C_1}(r_1)$ is shown in Figure 1b.

	Make	*Year*
t_1	VW	2002
t_2	VW	1997
t_3	Kia	1997

(a)

	Make	*Year*
t_1	VW	2002
t_3	Kia	1997

(b)

Fig. 1 (a) The Car relation; (b) Winnow result

In this paper, we focus on preference queries of the form $\omega_{\succ}(R)$, consisting of a single occurrence of winnow. Here $\succ$ is a preference relation (typically defined by a formula), and R is a database relation. The relation R represents the space of possible choices. We also briefly discuss how our results can be applied to more general preference queries.

Past work on preference queries has made the assumption that preferences are *static*. However, this assumption is often not satisfied. User preferences change, sometimes as a direct consequence of evaluating a preference query. Therefore, we view preference querying as a *dynamic, iterative process*. The user submits a query and inspects the result. The result may be satisfactory, in which case the querying process terminates. Or, the result may be too large or too small, contain unexpected answers, or fail to contain expected answers. By inspecting the query answer, the user may realize some previously unnoticed aspects of her preferences. It is also possible that not all the relevant data was included in the database over which the preference query is evaluated.

So if the user is not satisfied with the preference query result, she has several further options:

Modify and resubmit the query. This is appropriate if the user decides to refine or change her preferences. For example, the user may have started with a partial or vague concept of her preferences (PFT03). We consider here query modification consisting of *revising* the preference relation $\succ$, although, of course, more general transformations may also be envisioned.

Update the database. This is appropriate if the user discovers that there are more (or fewer) possible choices than originally envisioned. For example, in comparison shopping the user may have discovered a new source of relevant data.

In this context we pursue the following research challenges:

Defining a repertoire of suitable preference relation revisions. In this work, we consider revisions obtained by *composing* the original preference relation with a new preference relation, and *transitively closing* the result (to guarantee transitivity). We study different composition operators: union, and prioritized and Pareto composition. Those operators represent several basic ways of combining preferences and have already been incorporated into preference query languages (Cho03; Kie02). The operators reflect different user attitudes towards *preference conflicts*. (A conflict is, intuitively, a situation in which two preference relations order the same pair of tuples differently.) Union ignores conflicts (and thus such conflicts need to be prevented if we want to obtain a preference relation which is a strict partial order). Prioritized composition resolves preference conflicts by consistently giving priority to one of the preference relations. Pareto composition resolves conflicts in a symmetric way. We emphasize that revision is done using composition because we want the revised preference relation to be uniquely defined in the same first-order language as the original preference relation. Clearly, the revision repertoire that we study in this paper does not exhaust all meaningful scenarios. One can also imagine approaches where axiomatic properties of preference revisions are studied, as in belief revision (GR95).

Identifying essential properties of preference revisions. We claim that revisions should preserve the order-theoretic properties of the original preference relations. For example, if we start with a preference relation which is a strict partial order, the revised relation should also have those properties. This motivates, among others, transitively closing preference relations to guarantee transitivity. Preserving order-theoretic properties of preference relations is particularly important in view of the iterative construction of preference queries where the output of a revision can serve as the input to another one. We study both necessary and sufficient conditions on the original and revising preference relations that yield the preservation of

their order-theoretic properties. Necessary conditions are connected with the absence of preference conflicts. However, such conditions are typically not sufficient and stronger assumptions about the preference relations need to be made. Somewhat surprisingly, a special class of strict partial orders, *interval orders*, plays an important role in this context. The conditional preservation results we establish in this paper supplement those in (Cho03; Kie02) and may be used in other contexts where preference relations are composed, for example in the implementation of preference query languages. Another desirable property of revisions is minimality in some well-defined sense. We define minimality in terms of symmetric difference of preference relations but there are clearly other possibilities.

Incremental evaluation of preference queries. At each point of the interaction with the user, the results of evaluating *previous* versions of the given preference query are available. Therefore, they can be used to make the evaluation of the *current* query more efficient. For both the preference revision and database update scenarios, we formulate algebraic laws that validate new query evaluation plans that use materialized results of past query evaluations. The laws use order-theoretic properties of preference relations in an essential way.

Example 2 Consider Example 1. Seeing the result of the query $\omega_{C_1}(r_1)$, a user may realize that the preference relation $\succ_{C_1}$ is not quite what she had in mind. The result of the query may contain some unexpected or unwanted tuples, for example t_3. Thus the preference relation needs to be modified, for example by revising it with the following preference relation $\succ_{C_2}$:

$$(m,y) \succ_{C_2} (m',y') \equiv m = \text{''VW''} \wedge m' \neq \text{''VW''} \wedge y = y'.$$

As there are no conflicts between $\succ_{C_1}$ and $\succ_{C_2}$, the user chooses union as the composition operator. However, to guarantee transitivity of the resulting preference relation, $\succ_{C_1} \cup \succ_{C_2}$ has to be transitively closed. So the revised relation is $\succ_{C*} \equiv TC(\succ_{C_1} \cup \succ_{C_2})$. (The explicit definition of $\succ_{C*}$ is given in Example 6.) The tuple t_3 is now dominated by t_2 (i.e., $t_2 \succ_{C*} t_3$) and will not be returned to the user.

The plan of the paper is as follows. In Section 2, we define the basic notions. In Section 3, we introduce preference revision. In Section 4, we discuss query modification and the preservation by revisions of order-theoretic properties of preference relations. In Section 5, we discuss incremental evaluation of preference queries in the context of query modification and database updates. Subsequently, we consider two variations of our basic framework: (finite) restrictions of preference relations (Section 6) and weak-order extensions of strict partial order preference relations (Section 7). We briefly discuss related work in Section 8 and conclude in Section 9.

2 Basic notions

We are working in the context of the relational model of data. Relation schemas consist of finite sets of attributes. For concreteness, we consider two infinite, countable domains: $\mathscr{D}$ (uninterpreted constants, for readability shown as strings) and $\mathscr{Q}$ (rational numbers), but our results, except where explicitly indicated, hold also for finite domains. We assume that database instances are finite sets of tuples. Additionally, we have the standard built-in predicates.

2.1 Preference relations

We adopt here the framework of (Cho03).

Definition 1 Given a relation schema $R(A_1 \cdots A_k)$ such that U_i, $1 \leq i \leq k$, is the domain (either $\mathscr{D}$ or $\mathscr{Q}$) of the attribute A_i, a relation $\succ$ is a *preference relation over R* if it is a subset of $(U_1 \times \cdots \times U_k) \times (U_1 \times \cdots \times U_k)$.

Although we assume that database instances are finite, in the presence of infinite domains preference relations can be infinite.

Typical properties of a preference relation $\succ$ include (Fis70):

- *irreflexivity*: $\forall x.\ x \not\succ x$;
- *transitivity*: $\forall x,y,z.\ (x \succ y \wedge y \succ z) \Rightarrow x \succ z$;
- *negative transitivity*: $\forall x,y,z.\ (x \not\succ y \wedge y \not\succ z) \Rightarrow x \not\succ z$;
- *connectivity*: $\forall x,y.\ x \succ y \vee y \succ x \vee x = y$;
- *strict partial order* (SPO) if $\succ$ is irreflexive and transitive;
- *interval order* (IO) (Fis85) if $\succ$ is an SPO and satisfies the condition

$$\forall x,y,z,w.\ (x \succ y \wedge z \succ w) \Rightarrow (x \succ w \vee z \succ y);$$

- *weak order* (WO) if $\succ$ is a negatively transitive SPO;
- *total order* if $\succ$ is a connected SPO.

Every total order is a WO; every WO is an IO.

Definition 2 A *preference formula (pf)* $C(t_1,t_2)$ is a first-order formula defining a preference relation $\succ_C$ in the standard sense, namely

$$t_1 \succ_C t_2 \text{ iff } C(t_1,t_2).$$

An *intrinsic preference formula (ipf)* is a preference formula that uses only built-in predicates.

By using the notation $\succ_C$ for a preference relation, we assume that there is an underlying pf C. Occasionally, we will limit our attention to ipfs consisting of the following two kinds of atomic formulas (assuming we have two kinds of variables: $\mathscr{D}$-variables and $\mathscr{Q}$-variables):

- *equality constraints*: $x = y$, $x \neq y$, $x = c$, or $x \neq c$, where x and y are $\mathscr{D}$-variables, and c is an uninterpreted constant;
- *rational-order constraints*: $x \lambda y$ or $x \lambda c$, where $\lambda \in \{=, \neq, <, >, \leq, \geq\}$, x and y are $\mathscr{Q}$-variables, and c is a rational number.

An ipf all of whose atomic formulas are equality (resp. rational-order) constraints will be called an *equality* (resp. *rational-order*) ipf. If both equality and rational-order constraints are allowed in a formula, the formula will be called *ERO*. Clearly, ipfs are a special case of general constraints (KLP00; KKR95), and define *fixed*, although possibly infinite, relations.

Proposition 1 *Satisfiability of quantifier-free ERO formulas is in NP.*

Proof Satisfiability of conjunctions of atomic ERO constraints can be checked in linear time (GSW96). In an arbitrary quantifier-free ERO formula negation can be eliminated. Then in every disjunction one needs to nondeterministically select one disjunct, ultimately obtaining a conjunction of atomic constraints. □

Proposition 1 implies that all the properties that can be polynomially reduced to validity of ERO formulas, for example all the order-theoretic properties listed above, can be decided in co-NP.

Every preference relation $\succ$ generates an indifference relation $\sim$: two tuples t_1 and t_2 are *indifferent* ($t_1 \sim t_2$) if neither is preferred to the other one, i.e., $t_1 \not\succ t_2$ and $t_2 \not\succ t_1$. We denote by $\sim_C$ the indifference relation generated by $\succ_C$.

Composite preference relations are defined from simpler ones using logical connectives. We focus on the following basic ways of composing preference relations over the same schema:

- *union:* $t_1\ (\succ_1 \cup \succ_2)\ t_2$ iff $t_1 \succ_1 t_2 \vee t_1 \succ_2 t_2$;
- *prioritized composition:* $t_1\ (\succ_1 \rhd \succ_2)\ t_2$ iff $t_1 \succ_1 t_2 \vee (t_2 \not\succ_1 t_1 \wedge t_1 \succ_2 t_2)$;
- *Pareto composition:*

$$t_1\ (\succ_1 \otimes \succ_2)\ t_2 \text{ iff } (t_1 \succ_1 t_2 \wedge t_2 \not\succ_2 t_1) \vee (t_1 \succ_2 t_2 \wedge t_2 \not\succ_1 t_1).$$

We will use the above composition operators to construct revisions of given preference relations. We also consider transitive closure:

Definition 3 The *transitive closure* of a preference relation $\succ$ over a relation schema R is a preference relation $TC(\succ)$ over R defined as:

$$(t_1, t_2) \in TC(\succ) \text{ iff } t_1 \succ^n t_2 \text{ for some } n > 0,$$

where:

$$t_1 \succ^1 t_2 \equiv t_1 \succ t_2$$
$$t_1 \succ^{n+1} t_2 \equiv \exists t_3.\ t_1 \succ t_3 \wedge t_3 \succ^n t_2.$$

Clearly, in general Definition 3 leads to infinite formulas. However, in the cases that we consider in this paper the preference relation $\succ_{TC(\succ)}$ will in fact be defined by a finite formula.

Proposition 2 *Transitive closure of every preference relation defined by an ERO ipf is definable using an ERO ipf of at most exponential size, which can be computed in exponential time.*

Proof This is because transitive closure can be expressed in Datalog and the evaluation of Datalog programs over equality and rational-order constraints terminates in exponential time (combined complexity) (KKR95). □

In the cases mentioned above, the transitive closure of a given preference relation is a relation definable in the signature of the preference formula. But clearly transitive closure, unlike union and prioritized or Pareto composition, is itself not a first-order definable operator.

2.2 Winnow

We define now an algebraic operator that picks from a given relation the set of the *most preferred tuples*, according to a given preference relation.

Definition 4 (Cho03) If R is a relation schema and $\succ$ a preference relation over R, then the *winnow operator* is written as $\omega_{\succ}(R)$, and for every instance r of R:

$$\omega_{\succ}(r) = \{t \in r \mid \neg\exists t' \in r.\, t' \succ t\}.$$

If a preference relation is defined using a pf C, we write simply ω_C instead of $\omega_{\succ_C}$. A *preference query* is a relational algebra query containing at least one occurrence of the winnow operator.

3 Preference revisions

The basic setting is as follows: We have an *original* preference relation $\succ$ and revise it with a *revising* preference relation $\succ_0$ to obtain a *revised* preference relation $\succ'$. We also call $\succ'$ a *revision* of $\succ$. We assume that $\succ$, $\succ_0$, and $\succ'$ are preference relations over the same schema, and that all of them satisfy at least the properties of SPOs.

In our setting, a revision is obtained by composing $\succ$ with $\succ_0$ using union, prioritized or Pareto composition, and transitively closing the result (if necessary to obtain transitivity). However, we formulate some properties, like minimality or compatibility, in more general terms.

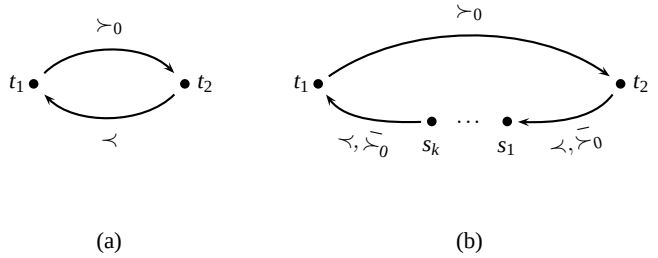
To define minimality, we order revisions using the symmetric difference ($\triangle$).

Definition 5 Assume $\succ_1$ and $\succ_2$ are two revisions of a preference relation $\succ$ with a preference relation $\succ_0$. We say that $\succ_1$ is *closer* than $\succ_2$ to $\succ$ if $\succ_1 \triangle \succ \subset \succ_2 \triangle \succ$.

To further describe the behavior of revisions, we define several kinds of *preference conflicts*. The intuition here is to characterize those conflicts that, when eliminated by prioritized or Pareto composition, reappear if the resulting preference relation is closed by transitivity.

Definition 6 A 0-*conflict* between a preference relation $\succ$ and a preference relation $\succ_0$ is a pair (t_1, t_2) such that $t_1 \succ_0 t_2$ and $t_2 \succ t_1$. A 1-*conflict* between $\succ$ and $\succ_0$ is a pair (t_1, t_2) such that $t_1 \succ_0 t_2$ and there exist $s_1, \ldots s_k$, $k \geq 1$, such that $t_2 \succ s_1 \succ \cdots \succ s_k \succ t_1$ and $t_1 \not\succ_0 s_k \not\succ_0 \cdots \not\succ_0 s_1 \not\succ_0 t_2$. A 2-*conflict* between $\succ$ and $\succ_0$ is a pair (t_1, t_2) such that there exist $s_1, \ldots, s_k$, $k \geq 1$ and $w_1, \ldots, w_m$, $m \geq 1$, such that $t_2 \succ s_1 \succ \cdots \succ s_k \succ t_1$, $t_1 \not\succ_0 s_k \not\succ_0 \cdots \not\succ_0 s_1 \not\succ_0 t_2$, $t_1 \succ_0 w_1 \succ_0 \cdots \succ_0 w_m \succ t_2$, and $t_2 \not\succ w_m \not\succ \cdots \not\succ w_1 \not\succ t_1$

A 1-conflict is a 0-conflict if $\succ$ is an SPO, but not necessarily vice versa. A 2-conflict is a 1-conflict if $\succ_0$ is an SPO. The different kinds of conflicts are pictured in Figures 2 and 3 ($\overline{\succ}$ denotes the complement of $\succ$).



(a) (b)

Fig. 2 (a) 0-conflict; (b) 1-conflict

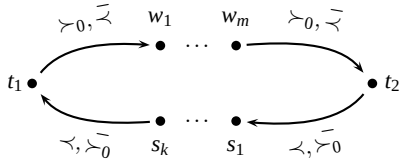


Fig. 3 2-conflict

Example 3 If $\succ_0 = \{(a,b)\}$ and $\succ = \{(b,a)\}$, then (a,b) is a 0-conflict which is not a 1-conflict. If we add (b,c) and (c,a) to $\succ$, then the conflict becomes a 1-conflict ($s_1 = c$). If we further add (c,b) or (a,c) to $\succ_0$, then the conflict is not a 1-conflict anymore. On the other hand, if we add (a,d) and (d,b) to $\succ_0$ instead, then we obtain a 2-conflict.

We assume here that the preference relations $\succ$ and $\succ_0$ are SPOs. If $\succ' = TC(\succ \cup \succ_0)$, then for every 0-conflict between $\succ$ and $\succ_0$, we still obviously have $t_1 \succ' t_2$ and $t_2 \succ' t_1$. Therefore, we say that the union does not resolve any conflicts. On the other hand, if $\succ' = TC(\succ_0 \rhd \succ)$, then for each 0-conflict (t_1, t_2), $t_1 \succ_0 \rhd \succ t_2$ and $\neg(t_2 \succ_0 \rhd \succ t_1)$. In the case of 1-conflicts, we get again $t_1 \succ' t_2$ and $t_2 \succ' t_1$. But in the case where a 0-conflict is not a 1-conflict, we get only $t_1 \succ' t_2$. Thus we say that prioritized composition resolves those 0-conflicts that are not 1-conflicts. Finally, if $\succ' = TC(\succ \otimes \succ_0)$, then for each 1-conflict (t_1, t_2), $\neg(t_1 \succ \otimes \succ_0 t_2)$ and $\neg(t_2 \succ \otimes \succ_0 t_1)$. We get $t_1 \succ' t_2$ and $t_2 \succ' t_1$ if the conflict is a 2-conflict, but if it is not, we obtain only $t_2 \succ' t_1$. Thus we say that Pareto composition resolves those 1-conflicts that are not 2-conflicts. (Pareto composition resolves also conflicts that are symmetric versions of 1-conflicts, with $\succ_0$ and $\succ$ interchanged, which are not 2-conflicts.)

We now characterize those combinations of $\succ$ and $\succ_0$ that avoid different kinds of conflicts.

Definition 7 A preference relation $\succ$ is *i-compatible*($i = 0, 1, 2$) with a preference relation $\succ_0$ if there are no *i*-conflicts between $\succ$ and $\succ_0$.

0- and 2-compatibility are symmetric. 1-compatibility is not necessarily symmetric. For SPOs, 0-compatibility implies 1-compatibility and 1-compatibility implies 2-compatibility. Examples 1 and 2 show a pair of 0-compatible relations. 0-compatibility of $\succ$ and $\succ_0$ *does not require* the acyclicity of $\succ \cup \succ_0$ or that one of the following hold: $\succ \subseteq \succ_0$, $\succ_0 \subseteq \succ$, or $\succ \cap \succ_0 = \emptyset$.

Propositions 1 and 2 imply that all the variants of compatibility defined above are decidable for ERO ipfs. For example, 1-compatibility is expressed by the condition $\succ_0^{-1} \cap TC(\succ - \succ_0^{-1}) = \emptyset$ where $\succ_0^{-1}$ is the inverse of the preference relation $\succ_0$.

0-compatibility of $\succ$ and $\succ_0$ is a *necessary* condition for $TC(\succ \cup \succ_0)$ to be irreflexive, and thus an SPO. Similar considerations apply to $TC(\succ_0 \rhd \succ)$ and 1-compatibility, and $TC(\succ \otimes \succ_0)$ and 2-compatibility. In the next section, we will see that those conditions are not *sufficient*: further restrictions on the preference relations will be introduced.

We conclude by noting the relationships between the three notions of preference composition introduced above.

Lemma 1 *For every preference relations* $\succ$ *and* $\succ_0$

$$\succ_0 \otimes \succ \subseteq \succ_0 \rhd \succ \subseteq \succ_0 \cup \succ,$$

and if $\succ_0$ *and* $\succ$ *are* 0*-compatible*

$$\succ_0 \otimes \succ = \succ_0 \rhd \succ = \succ_0 \cup \succ.$$

4 Query modification

In this section, we study preference query modification [1]. A given preference query $\omega_{\succ}(R)$ is transformed to the query $\omega_{\succ'}(R)$ where $\succ'$ is obtained by composing the original preference relation $\succ$ with the revising preference relation $\succ_0$, and transitively closing the result. (The last step is clearly unnecessary if the obtained preference relation is already transitive.) We want $\succ'$ to satisfy the same order-theoretic properties as $\succ$ and $\succ_0$, and to be minimally different from $\succ$. To achieve those goals, we impose additional conditions on $\succ$ and $\succ_0$.

For every $\theta \in \{\cup, \rhd, \otimes\}$, we consider the order-theoretic properties of the preference relation $\succ' = \succ_0 \theta \succ$, or $\succ' = TC(\succ_0 \theta \succ)$ if $\succ_0 \theta \succ$ is not guaranteed to be transitive. To ensure that this preference relation is an SPO, only irreflexivity has to be guaranteed; for weak orders one has also to establish negative transitivity.

[1] The term *query modification* was used in early relational systems like INGRES to denote a technique that produced a changed version of a query submitted by a user. The changes were meant to incorporate view definitions, integrity constraints and security specifications. We feel that it is justified to use the same term in the context of composition of a preference relation in a query with some other preference relation, to produce a new query.

4.1 Strict partial orders

SPOs have several important properties from the user's point of view, and thus their preservation is desirable. For instance, all the preference relations defined in (Kie02) and in the language Preference SQL (KK02) are SPOs. Moreover, if $\succ$ is an SPO, then the winnow $\omega_{\succ}(r)$ is nonempty if (a finite) r is nonempty. The fundamental algorithms for computing winnow require that the preference relation be an SPO (Cho03). Also, in that case incremental evaluation of preference queries becomes possible (Proposition 5 and Theorem 7).

Theorem 1 *For every* 0*-compatible preference relations* $\succ$ *and* $\succ_0$ *such that one is an interval order (IO) and the other an SPO, the preference relation* $TC(\succ_0 \ \theta \ \succ)$*, where* $\theta \in \{\cup, \rhd, \otimes\}$*, is an SPO. If the IO is a WO, then* $TC(\succ_0 \ \theta \ \succ) = \succ_0 \ \theta \ \succ$.

Proof By Lemma 1, 0-compatibility implies that $\succ_0 \cup \succ = \succ_0 \rhd \succ = \succ_0 \otimes \succ$. Thus, WLOG we consider only union. Assume $\succ_0$ is an IO. If $TC(\succ \cup \succ_0)$ is not irreflexive, then $\succ \cup \succ_0$ has a cycle. Consider such cycle of minimum length. It consists of edges that are alternately labeled $\succ_0$ (only) and $\succ$ (only). (Otherwise the cycle can be shortened). If there is more than one non-consecutive $\succ_0$-edge in the cycle, then $\succ_0$ being an IO implies that the cycle can be shortened. So the cycle consists of two edges: $t_1 \succ_0 t_2$ and $t_2 \succ t_1$. But this is a 0-conflict violating 0-compatibility of $\succ$ and $\succ_0$. □

It is easy to see that there is no preference relation which is an SPO, contains $\succ \cup \succ_0$, and is closer (in the sense of Definition 5) to $\succ$ than $TC(\succ \cup \succ_0)$.

As can be seen from the above proof, the fact that one of the preference relations is an interval order makes it possible to eliminate those paths (and thus also cycles) in $TC(\succ \cup \succ_0)$ that interleave $\succ$ and $\succ_0$ more than once. In this way acyclicity reduces to the lack of 0-conflicts.

It seems that the interval order (IO) requirement in Theorem 1 cannot be weakened without needing to strengthen the remaining assumptions. If neither of $\succ$ and $\succ_0$ is an IO, then we can find such elements $x_1, y_1, z_1, w_1, x_2, y_2, z_2, w_2$ that

$$x_1 \succ y_1, z_1 \succ w_1, x_1 \not\succ w_1, z_1 \not\succ y_1, x_2 \succ_0 y_2, z_2 \succ_0 w_2, x_2 \not\succ_0 w_2,$$

and $z_2 \not\succ_0 y_2$. If we choose $y_1 = x_2$, $z_1 = y_2$, $w_1 = z_2$, and $x_1 = w_2$, then we get a cycle in $\succ \cup \succ_0$. Note that in this case $\succ$ and $\succ_0$ are still 0-compatible. Also, there is no SPO preference relation which contains $\succ \cup \succ_0$ because each such relation has to contain $TC(\succ \cup \succ_0)$. This situation is pictured in Figure 4.

Example 4 Consider again the preference relation $\succ_{C_1}$:

$$(m,y) \succ_{C_1} (m',y') \equiv m = m' \wedge y > y'.$$

Suppose that the new preference information is captured as $\succ_{C_3}$ which is an IO but not a WO:

$$(m,y) \succ_{C_3} (m',y') \equiv m = ''\mathrm{VW}'' \wedge y = 1999 \wedge m' = ''\mathrm{Kia}'' \wedge y' = 1999.$$

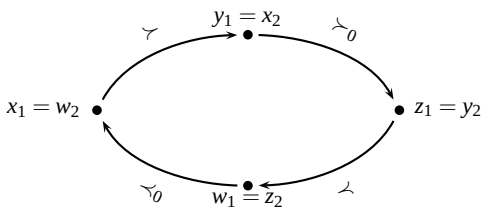


Fig. 4 A cycle for 0-compatible relations that are not IOs.

Then $TC(\succ_{C_1} \cup \succ_{C_3})$, which properly contains $\succ_{C_1} \cup \succ_{C_3}$, is defined as the SPO $\succ_{C_4}$:

$$(m,y) \succ_{C_4} (m',y') \equiv m = m' \wedge y > y' \vee \\ m = \text{''VW''} \wedge y \geq 1999 \wedge m' = \text{''Kia''} \wedge y' \leq 1999.$$

Theorem 1 implies that if $\succ$ and $\succ_0$ are 0-compatible and one of them contains only one pair, then $TC(\succ \cup \succ_0)$ is an SPO. So what will happen if we break up the preference relation $\succ_0$ from Figure 4 into two one-element relations $\succ_1$ and $\succ_2$ and attempt to apply Theorem 1 twice? Unfortunately, such a "strategy" does not work. The second step is not possible because the preference relation $\succ_2$ is not 0-compatible with the revision of $\succ$ with $\succ_1$.

For dealing with *prioritized composition*, 0-compatibility can be replaced by a less restrictive condition, 1-compatibility, because prioritized composition already provides a way of resolving some conflicts.

Theorem 2 *For every preference relations $\succ$ and $\succ_0$ such that $\succ_0$ is an IO, $\succ$ is an SPO and $\succ$ is 1-compatible with $\succ_0$, the preference relation $TC(\succ_0 \rhd \succ)$ is an SPO.*

Proof We assume that $TC(\succ_0 \rhd \succ)$ is not irreflexive and consider a cycle of minimum length in $\succ_0 \rhd \succ$. If the cycle has two non-consecutive edges labeled (not necessarily exclusively) by $\succ_0$, then it can be shortened, because $\succ_0$ is an IO. The cycle has to consist of an edge $t_1 \succ_0 t_2$ and a sequence of edges (labeled only by $\succ$): $t_2 \succ t_3, \ldots, t_{n-1} \succ t_n, t_n \succ t_1$ such that $n > 2$. and $t_1 \not\succ_0 t_n \not\succ_0 \ldots \not\succ_0 t_3 \not\succ_0 t_2$. (We cannot shorten sequences of consecutive $\succ$-edges because $\succ$ is not necessarily preserved in $\succ_0 \rhd \succ$.) Thus (t_1, t_2) is a 1-conflict violating 1-compatibility of $\succ$ with $\succ_0$. □

Clearly, there is no SPO preference relation which contains $\succ_0 \rhd \succ$, and is closer to $\succ$ than $TC(\succ_0 \rhd \succ)$. Violating any of the conditions of Theorem 2 may lead to a situation in which no SPO preference relation which contains $\succ_0 \rhd \succ$ exists.

If $\succ_0$ is a WO, the requirement of 1-compatibility and the computation of transitive closure are unnecessary. We first recall some basic properties of weak orders.

Proposition 3 *Let $\succ$ be a WO preference relation over a schema R and $\sim$ the indifference relation generated by $\succ$. If $x \succ y$, $y \sim z$ and $z \succ w$, then also $x \succ z$ and $y \succ w$.*

Theorem 3 *For every preference relations $\succ_0$ and $\succ$ such that $\succ_0$ is a WO and $\succ$ an SPO, the preference relation $\succ_0 \rhd \succ$ is an SPO.*

Proof Clearly, $\succ' = \succ_0 \rhd \succ$, as a subset of $\succ_0 \cup \succ$, is irreflexive. To show transitivity, consider $t_1 \succ' t_2$ and $t_2 \succ' t_3$. There are four possibilities: (1) If $t_1 \succ_0 t_2$ and $t_2 \succ_0 t_3$, then $t_1 \succ_0 t_3$ and $t_1 \succ' t_3$. (2) If $t_1 \succ_0 t_2$, $t_3 \not\succ_0 t_2$ and $t_2 \succ t_3$, then also $t_2 \succ_0 t_3$ or $t_2 \sim_0 t_3$ (where $\sim_0$ is the indifference relation generated by $\succ_0$). In either case, $t_1 \succ_0 t_3$ and $t_1 \succ' t_3$ (the second case requires using Proposition 3). (3) $t_2 \not\succ_0 t_1$, $t_1 \succ t_2$ and $t_2 \succ_0 t_3$: symmetric to (2). (4) If $t_2 \not\succ_0 t_1$, $t_1 \succ t_2$, $t_3 \not\succ_0 t_2$ and $t_2 \succ t_3$, then $t_3 \not\succ_0 t_1$ (by the negative transitivity of $\succ_0$) and $t_1 \succ t_3$. Thus $t_1 \succ' t_3$. □

Let's turn now to *Pareto composition*. There does not seem to be any simple way to *weaken* the assumptions in Theorem 1 using the notion of 2-compatibility. Assuming that $\succ$, $\succ_0$, or even both are IOs does not sufficiently restrict the possible interleavings of $\succ$ and $\succ_0$ in $TC(\succ_0 \otimes \succ)$ because neither of those two preference relations is guaranteed to be preserved in $TC(\succ_0 \otimes \succ)$. However, we can establish a weaker version of Theorem 3.

Theorem 4 *For every preference relations $\succ_0$ and $\succ$ such that both are WOs, the preference relation $\succ_0 \otimes \succ$ is an SPO.*

Proof Similar to the proof of Theorem 3.

Proposition 2 implies that for all preference relations defined using ERO ipfs, the computation of the preference relations $TC(\succ \cup \succ_0)$, $TC(\succ_0 \rhd \succ)$, as well as $TC(\succ \otimes \succ_0)$ terminates. The computation of transitive closure is done in a completely database-independent way.

Example 5 Consider Examples 1 and 4. We can infer that

$$t_1 = (''\text{VW}'', 2002) \succ_{C_4} (''\text{Kia}'', 1997) = t_3,$$

because $(''\text{VW}'', 2002) \succ_{C_1} (''\text{VW}'', 1999)$, $(''\text{VW}'', 1999) \succ_{C_3} (''\text{Kia}'', 1999)$, and $(''\text{Kia}'', 1999) \succ_{C_1} (''\text{Kia}'', 1997)$. The tuples $(''\text{VW}'', 1999)$ and $(''\text{Kia}'', 1999)$ are *not* in the database.

If the conditions of Theorems 1 and 2 do not apply, Proposition 2 implies that for ERO ipfs the computation of $TC(\succ \cup \succ_0)$, $TC(\succ_0 \rhd \succ)$ and $TC(\succ \otimes \succ_0)$ yields some finite ipf $C(t_1, t_2)$. Thus the irreflexivity of the resulting preference relation reduces to the unsatisfiability of $C(t,t)$, which by Proposition 1 is a decidable problem for ERO ipfs. Of course, the relation, being a transitive closure, is already transitive.

Example 6 Consider Examples 1 and 2. Neither of the preference relations $\succ_{C_1}$ and $\succ_{C_2}$ is an interval order. Therefore, the results established earlier in this section do not apply. The preference relation $\succ_{C*} = TC(\succ_{C_1} \cup \succ_{C_2})$ is defined as follows (this definition is obtained using Constraint Datalog computation):

$$(m,y) \succ_{C*} (m',y') \equiv m = m' \wedge y > y' \vee$$
$$m = ''\mathrm{VW}'' \wedge m' \neq ''\mathrm{VW}'' \wedge y \geq y'.$$

The preference relation $\succ_{C*}$ is irreflexive (this can be effectively checked). It also properly contains $\succ_{C_1} \cup \succ_{C_2}$, because $t_1 \succ_{C*} t_3$ but $t_1 \not\succ_{C_1} t_3$ and $t_1 \not\succ_{C_2} t_3$. The query $\omega_{C*}(Car)$ evaluated in the instance r_1 (Figure 1) returns only the tuple t_1.

4.2 Weak orders

Weak orders are practically important because they capture the situation where the domain can be decomposed into layers such that the layers are totally ordered and all the elements in one layer are mutually indifferent. This is the case, for example, if a preference relation can be represented using a numeric utility function. If a preference relation is a WO, a particularly efficient (essentially single pass) algorithm for computing winnow is applicable (Cho04).

We will see that for weak orders the transitive closure computation is unnecessary and minimal revisions are directly definable in terms of the preference relations involved.

Theorem 5 *For every* 0*-compatible WO preference relations* $\succ$ *and* $\succ_0$*, the preference relations* $\succ \cup \succ_0$ *and* $\succ \otimes \succ_0$ *are WO.*

Proof In view of Lemma 1, we can consider only $\succ' = \succ \cup \succ_0$.

Irreflexivity is obvious. For transitivity, assume $t_1 \succ' t_2$ and $t_2 \succ' t_3$. If $t_1 \succ t_2 \succ t_3$ (resp. $t_1 \succ_0 t_2 \succ_0 t_3$), then $t_1 \succ t_3$ (resp. $t_1 \succ_0 t_3$) and $t_1 \succ' t_3$. If $t_1 \succ_0 t_2$ and $t_2 \succ t_3$, we need 0-compatibility to infer that $t_2 \not\succ t_1$ and thus $t_1 \succ t_2$ or $t_1 \sim t_2$ (where $\sim$ is the indifference relation generated by $\succ$). In both cases, we can infer $t_1 \succ t_3$ and thus $t_1 \succ' t_3$. The last case is symmetric to the previous one.

For negative transitivity, consider $t_1 \not\succ' t_2$ and $t_2 \not\succ' t_3$. Then $t_1 \not\succ_0 t_2$, $t_2 \not\succ_0 t_3$, $t_1 \not\succ t_2$, and $t_2 \not\succ t_3$. Consequently, $t_1 \not\succ_0 t_3$, $t_1 \not\succ t_3$, and thus $t_1 \not\succ' t_3$. □

Note that without the 0-compatibility assumption, WOs are not closed with respect to union and Pareto composition (Cho03).

For prioritized composition, we can relax the 0-compatibility assumption. This immediately follows from the fact that WOs are closed with respect to prioritized composition (Cho03).

Proposition 4 *For every WO preference relations* $\succ$ *and* $\succ_0$*, the preference relation* $\succ_0 \rhd \succ$ *is a WO.*

A basic notion in utility theory is that of *representability* of preference relations using numeric utility functions:

Definition 8 A real-valued function u over a schema R *represents* a preference relation $\succ$ over R iff

$$\forall t_1, t_2 \; [t_1 \succ t_2 \text{ iff } u(t_1) > u(t_2)].$$

Such a preference relation is called *utility-based.*

Being a WO is a necessary condition for the existence of a numeric representation for a preference relation. However, it is not sufficient for uncountable orders (Fis70). It is natural to ask whether the existence of numeric representations for the preference relations $\succ$ and $\succ_0$ implies the existence of such a representation for the preference relation $\succ' = (\succ_0 \; \theta \; \succ)$ where $\theta \in \{\cup, \rhd, \otimes\}$. This is indeed the case.

Theorem 6 *Assume that $\succ$ and $\succ_0$ are WO preference relations such that*

1. *$\succ$ and $\succ_0$ are 0-compatible,*
2. *$\succ$ can be represented using a real-valued function u,*
3. *$\succ_0$ can be represented using a real-valued function u_0.*

Then $\succ' = \succ_0 \; \theta \; \succ$, where $\theta \in \{\cup, \rhd, \otimes\}$, is a WO preference relation that can be represented using any real-valued function u' such that for all x

$$u'(x) = a \cdot u(x) + b \cdot u_0(x) + c$$

where a and b are arbitrary positive real numbers.

Proof Assume $x \succ' y$. Thus $x \succ_0 y$ or $x \succ y$. If $x \succ_0 y$, then $u_0(x) > u_0(y)$. Also, in this case $y \not\succ x$ because of 0-compatibility. This implies $u(x) \geq u(y)$. Consequently, $u'(x) > u'(y)$. The other case is symmetric.

Assume $u'(x) > u'(y)$. Thus $u_0(x) > u_0(y)$ or $u(x) > u(y)$. In the first case, we get $x \succ_0 y$; in the second, $x \succ y$. Consequently, $x \succ' y$. □

Surprisingly, the 0-compatibility requirement cannot in general be replaced by 1-compatibility if we replace $\cup$ by $\rhd$ in Theorem 6.

Example 7 Consider the Euclidean space $\mathscr{R} \times \mathscr{R}$, and the following orders:

$$(x,y) \succ_1 (x',y') \equiv x > x',$$
$$(x,y) \succ_2 (x',y') \equiv y > y',$$

The orders $\succ_1$ and $\succ_2$ are 1-compatible (but not 0-compatible) WOs. It is well known that their prioritized (also called *lexicographic*) composition is not representable using a utility function (Fis70).

Thus, preservation of *representability* is possible only under 0-compatibility, in which case $\succ_0 \cup \succ = \succ_0 \rhd \succ = \succ_0 \otimes \succ$ (Lemma 1). (The results (Fis70) indicate that for countable domains considered in this paper, the prioritized composition of WOs, being a WO, is representable using a utility function. However, that utility function is not definable in terms of the utility functions representing the given orders.)

We conclude this section by showing a general scenario in which the union of orders occurs in a natural way. Assume that we have a numeric utility function u representing a (WO) preference relation $\succ$. The indifference relation $\sim$ generated by $\succ$ is defined as:

$$x \sim y \equiv u(x) = u(y).$$

Suppose that the user discovers that $\sim$ is too coarse and needs to be further refined. This may occur, for example, when x and y are tuples and the function u takes into account only some of their components. Another function u_0 may be defined to take into account other components of x and y (such components are called *hidden attributes* (PFT03)). The revising preference relation $\succ_0$ is now:

$$x \succ_0 y \equiv u(x) = u(y) \wedge u_0(x) > u_0(y).$$

It is easy to see that $\succ_0$ is an SPO 0-compatible with $\succ$ (but not necessarily a WO). Therefore, by Theorem 1 the preference relation $\succ \cup \succ_0$ is an SPO.

5 Incremental evaluation

5.1 Query modification

We show here how the already computed result of the original preference query can be reused to make the evaluation of the modified query more efficient. We will use the following result.

Proposition 5 (Cho03) *If $\succ_1$ and $\succ_2$ are preference relations over a relation schema R and $\succ_1 \subseteq \succ_2$, then for all instances r of R:*

- $\omega_{\succ_2}(r) \subseteq \omega_{\succ_1}(r)$;
- $\omega_{\succ_2}(\omega_{\succ_1}(r)) = \omega_{\succ_2}(r)$ *if $\succ_1$ and $\succ_2$ are SPOs.*

Consider the scenario in which we iteratively modify a given preference query by revising the preference relation using only union in such a way that the revised preference relation is an SPO (for example, if the assumptions of Theorem 1 are satisfied). We obtain a sequence of preference relations $\succ_1, \ldots, \succ_n$ such that $\succ_1 \subseteq \cdots \subseteq \succ_n$.

In this scenario, the sequence of query results is:

$$r_0 = r, r_1 = \omega_{\succ_1}(r), r_2 = \omega_{\succ_2}(r), \ldots, r_n = \omega_{\succ_n}(r).$$

Proposition 5 implies that the sequence $r_0, r_1, \ldots, r_n$ is decreasing:

$$r_0 \supseteq r_1 \supseteq \cdots \supseteq r_n$$

and that it can be computed incrementally:

$$r_1 = \omega_{\succ_1}(r_0), r_2 = \omega_{\succ_2}(r_1), \ldots, r_n = \omega_{\succ_n}(r_{n-1}).$$

To compute r_i, there is no need to look at the tuples in $r - r_{i-1}$, nor to recompute winnow from scratch. The sets of tuples $r_1, \ldots, r_n$ are likely to have much smaller cardinality than $r_0 = r$.

It is easy to see that the above comments apply to all cases where the revised preference relation is a superset of the original preference relation. Unfortunately, this is not the case for revisions that use prioritized or Pareto composition. However, given a specific pair of preference relations $\succ$ and $\succ_0$, one can still effectively check whether $TC(\succ_0 \rhd \succ)$ or $TC(\succ_0 \otimes \succ)$ contains $\succ$ if the validity of preference formulas is decidable, as is the case for ERO formulas (Proposition 1).

5.2 Database update

In the previous section we studied query modification: the query is modified, while the database remains unchanged. Here we reverse the situation: the query remains the same and the database is updated.

We consider first updates that are insertions of sets of tuples. For a database relation r, we denote by $\Delta^+ r$ the set of inserted tuples. We show how the previous result of a given preference query can be reused to make the evaluation of the same query in an updated database more efficient.

We first establish the following result.

Theorem 7 *For every preference relation $\succ$ over R which is an SPO and every instance r of R:*

$$\omega_\succ(r \cup \Delta^+ r) = \omega_\succ(\omega_\succ(r) \cup \Delta^+ r).$$

Proof Assume $t \notin \omega_\succ(\omega_\succ(r) \cup \Delta^+ r)$. Then either $t \notin \omega_\succ(r) \cup \Delta^+ r$ or there exists $t' \in \omega_\succ(r) \cup \Delta^+ r$ such that $t' \succ t$. In the first case, $t \notin \omega_\succ(r)$ and $t \notin \Delta^+ r$. If $t \notin r$ and $t \notin \Delta^+ r$, then $t \notin \omega_\succ(r \cup \Delta^+ r)$. If there exists $t' \in \omega_\succ(r)$ such that $t' \succ t$, then also $t \notin \omega_\succ(r \cup \Delta^+ r)$. In the second case, $t' \in r \cup \Delta^+ r$ and thus $t \notin \omega_\succ(r \cup \Delta^+ r)$.

Assume $t \notin \omega_\succ(r \cup \Delta^+ r)$. Then either $t \notin r \cup \Delta^+ r$ or there exists $t' \in r \cup \Delta^+ r$ such that $t' \succ t$. In the first case, $t \notin \omega_\succ(\omega_\succ(r) \cup \Delta^+ r)$. In the second case, if $t' \in \Delta^+ r$, then $t \notin \omega_\succ(\omega_\succ(r) \cup \Delta^+ r)$. So consider $t' \in r - \Delta^+ r$. If $t \in r$ but $t \notin \Delta^+ r$, then $t \notin \omega_\succ(r) \cup \Delta^+ r$ and $t \notin \omega_\succ(\omega_\succ(r) \cup \Delta^+ r)$. If $t \in \Delta^+ r$, then there exists $t'' \in \omega_\succ(r)$ such that $t'' \succ t$. (t'' may be t' or some element dominating t'.) Therefore, in this case also $t \notin \omega_\succ(\omega_\succ(r) \cup \Delta^+ r)$. □

Consider now the scenario in which we have a preference relation $\succ$, which is an SPO, and a sequence of relations

$$r_0 = r, r_1 = r_0 \cup \Delta^+ r_0, r_2 = r_1 \cup \Delta^+ r_1, \ldots, r_n = r_{n-1} \cup \Delta^+ r_{n-1}.$$

Theorem 7 shows that

$$\begin{aligned}
&\omega_\succ(r_1) = \omega_\succ(\omega_\succ(r_0) \cup \Delta^+ r_0)\\
&\omega_\succ(r_2) = \omega_\succ(\omega_\succ(r_1) \cup \Delta^+ r_1)\\
&\ldots\\
&\omega_\succ(r_n) = \omega_\succ(\omega_\succ(r_{n-1}) \cup \Delta^+ r_{n-1}).
\end{aligned}$$

Therefore, each subsequent evaluation of winnow can reuse the result of the previous one. This is advantageous because winnow returns a subset of the given relation and this subset is often much smaller than the relation itself.

Clearly, the algebraic law, stated in Theorem 7, can be used together with other, well-known laws of relational algebra and the laws specific to preference queries (Cho03; KH03) to produce a variety of rewritings of a given preference query. To see how a more complex preference query can be handled, let's consider the query consisting of winnow and selection, $\omega_{\succ}(\sigma_{\alpha}(R))$. We have

$$\omega_{\succ}(\sigma_{\alpha}(r \cup \Delta^{+}r)) = \omega_{\succ}(\sigma_{\alpha}(r) \cup \sigma_{\alpha}(\Delta^{+}r)) = \omega_{\succ}(\omega_{\succ}(\sigma_{\alpha}(r)) \cup \sigma_{\alpha}(\Delta^{+}r))$$

for every instance r of R. Here again, one can use the previous result of the query, $\omega_{\succ}(\sigma_{\alpha}(r))$, to make its current evaluation more efficient. Other operators that distribute through union, for example projection and join, can be handled in the same way.

Next, we consider updates that are deletions of sets of tuples. For a database relation r, we denote by $\Delta^{-}r$ the set of deleted tuples.

Theorem 8 *For every preference relation $\succ$ over R and every instance r of R:*

$$\omega_{\succ}(r) - \Delta^{-}r \subseteq \omega_{\succ}(r - \Delta^{-}r).$$

Theorem 8 gives an incremental way to compute an approximation of winnow from below. It seems that in the case of deletion there cannot be an exact law along the lines of Theorem 7. This is because the deletion of some tuples from the original database may promote some originally dominated (and discarded) tuples into the result of winnow over the updated database.

Example 8 Consider the following preference relation $\succ = \{(a,b_1),\ldots,(a,b_n)\}$ and the database $r = \{a,b_1,\ldots,b_n\}$. Then $\omega_{\succ}(r) = \{a\}$ but

$$\omega_{\succ}(r - \{a\}) = \{b_1,\ldots,b_n\}.$$

6 Finite restrictions of preference relations

6.1 Restriction

It is natural to consider *restrictions* of preference relations to given database instances (TC02).

Definition 9 Let r be an instance of a relation schema R and $\succ$ a preference relation over R. The *restriction* $[\succ]_r$ *of* $\succ$ *to* r is a preference relation over R, defined as

$$[\succ]_r = \succ \cap\, r \times r.$$

We write $(x,y) \in [\succ]_r$ instead of $x[\succ]_r y$ for greater readability.

The advantage of using $[\succ]_r$ instead of $\succ$ comes from the fact that the former depends on the database contents and can have stronger properties than the latter. For example, $[\succ]_r$ may be an SPO, while $\succ$ is not. Similarly, $[\succ]_r$ may be *i*-compatible with $[\succ_0]_r$, while $\succ$ is not *i*-compatible with $\succ_0$. Therefore, restrictions could be used instead of preference relations in the revision process.

The following is a basic property of restriction. It says that the restriction to an instance does not affect the result of winnow *over the same instance*, so the restriction can be used in place of the original preference relation.

Theorem 9 *Let r be an instance of a relation schema R and $\succ$ a preference relation over R. Then*

$$\omega_{[\succ]_r}(r) = \omega_{\succ}(r).$$

Proof We have $[\succ]_r \subseteq r$ and thus $\omega_{\succ}(r) \subseteq \omega_{[\succ]_r}(r)$. In the other direction, assume $t \notin \omega_{\succ}(r)$. If $t \notin r$, $t \notin \omega_{[\succ]_r}(r)$. If $t \in r$ and there exists $t' \in r$ such that $t' \succ t$, then also $(t',t) \in [\succ]_r$ and $t \notin \omega_{[\succ]_r}(r)$. □

We also establish that restriction distributes over the preference composition operators.

Theorem 10 *If r is an instance of a relation schema R, $\theta \in \{\cup, \rhd, \otimes\}$, and $\succ$ and $\succ_0$ are preference relations over R, then*

$$[\succ_0 \ \theta \ \succ]_r = [\succ_0]_r \theta [\succ]_r.$$

Proof We prove this result for $\theta = \rhd$. The other cases are similar.

We have the following equivalences:

$$\begin{aligned}
&(x,y) \in [\succ_0]_r \rhd [\succ]_r \equiv \\
&(x,y) \in [\succ_0]_r \vee (y,x) \notin [\succ_0]_r \wedge (x,y) \in [\succ]_r \equiv \\
&x \succ_0 y \wedge x \in r \wedge y \in r \vee (y \not\succ_0 x \vee x \notin r \vee y \notin r) \wedge x \succ y \wedge x \in r \wedge y \in r \equiv \\
&x \succ_0 y \wedge x \in r \wedge y \in r \vee y \not\succ_0 x \wedge x \succ y \wedge x \in r \wedge y \in r \equiv \\
&(x \succ_0 y \vee y \not\succ_0 x \wedge x \succ y) \wedge x \in r \wedge y \in r \equiv \\
&(x,y) \in [\succ_0 \rhd \succ]_r.
\end{aligned}$$

The preference revision studied earlier in this paper typically involved the computation of the of the revised preference relation defined as the transitive closure $TC(\succ_0 \ \theta \ \succ)$, where $\theta \in \{\cup, \rhd, \otimes\}$, $\succ$ is the original preference relation, and $\succ_0$ is the revising preference relation. We study several different ways of imposing the restriction of preferences to a relation instance. We consider the following preference relations:

$$\begin{aligned}
\succ_1 &= TC(\succ_0 \ \theta \ \succ), \\
\succ_2 &= [TC(\succ_0 \ \theta \ \succ)]_r, \\
\succ_3 &= TC([\succ_0 \ \theta \ \succ]_r), \\
\succ_4 &= TC([\succ_0]_r \theta [\succ]_r).
\end{aligned}$$

We establish now some fundamental relationships between the preference relations $\succ_1, \succ_2, \succ_3$, and $\succ_4$.

Theorem 11 *Let $\theta \in \{\cup, \rhd, \otimes\}$, and $\succ$ and $\succ_0$ be preference relations over a schema R. Then for every instance r of R:*

$$\succ_4 = \succ_3 \subseteq \succ_2 \subseteq \succ_1,$$

and there are relation instances for which the containments are strict.

Proof The equality of $\succ_4$ and $\succ_3$ follows from Theorem 10. For $\succ_3 \subseteq \succ_2$, we have that

$$[\succ_0 \theta \succ]_r \subseteq \succ_0 \theta \succ,$$

and

$$[\succ_0 \theta \succ]_r \subseteq r \times r.$$

Thus

$$\succ_3 = TC([\succ_0 \theta \succ]_r) \subseteq r \times r,$$

and

$$\succ_3 \subseteq TC(\succ_0 \theta \succ) \cap r \times r = \succ_2 .$$

The containment $\succ_2 \subseteq \succ_1$ follows from the definition of the restriction.

An example where $\succ_3 \subset \succ_2 \subset \succ_1$ is as follows. Let $\succ = \{(a,b)\}$, $\succ_0 = \{(b,c)\}$, $r = \{a,c\}$. Then $[\succ]_r = [\succ_0]_r = \emptyset$. Thus also $[\succ_0 \theta \succ]_r = [\succ_0]_r \theta [\succ]_r = \emptyset$, and $\succ_3 = \emptyset$. On the other hand, $\succ_1 = \{(a,b),(b,c),(a,c)\}$ and $\succ_2 = \{(a,c)\}$. □

Corollary 1 *Let $\theta \in \{\cup, \rhd, \otimes\}$, and $\succ$ and $\succ_0$ be preference relations over a schema R. Then for every instance r of a R:*

$$\omega_{\succ_1}(r) = \omega_{\succ_2}(r) \subseteq \omega_{\succ_3}(r) = \omega_{\succ_4}(r),$$

and for some cases the containment is strict.

Proof Follows from Theorem 9 and Theorem 11. In the example given in the proof of Theorem 11, we obtain $\omega_{\succ_2}(r) = \{a\}$ and $\omega_{\succ_3}(r) = \{a,c\}$. □

We study now the order-theoretic properties of restriction.

Theorem 12 *Let $\theta \in \{\cup, \rhd, \otimes\}$, and $\succ$ and $\succ_0$ be preference relations over a schema R. Then for every instance r of R, $\succ_1$ is an SPO implies that $\succ_2$ is an SPO, which implies that $\succ_3$ is an SPO. There are cases in which the reverse implication does not hold.*

Proof Because $\succ_2 \subseteq \succ_1$, $\succ_2$ is irreflexive. Assume that $x \succ_2 y$ and $y \succ_2 z$. Then $x \succ_1 y$, $y \succ_1 z$, $x \in r$, $y \in r$, and $z \in r$. Therefore, $x \succ_1 z \wedge x \in r \wedge z \in r$, and $x \succ_2 z$.

The preference relation $\succ_1 = \{(a,a)\}$ is not an SPO (and can be obtained from some preference relations $\succ_0$ and $\succ$ using any composition operator). However, its restriction $\succ_2 = [\succ_1]_r$ for $r = \{b\}$ is empty, and thus an SPO.

Assume now $\succ_0 = \{(a,b)\}$ and $\succ = \{(b,a)\}$. Consider $\theta = \cup$ and $r = \{b\}$. Thus, $\succ_1 = \{(a,b),(b,a),(a,a),(b,b)\}$ and $\succ_2 = \{(b,b)\}$ (so it is not an SPO). On the other hand, $[\succ_0 \cup \succ]_r = \emptyset$ and $\succ_3 = \emptyset$, too. Similar examples can be constructed for the other composition operators. □

Unfortunately, for weak orders there is no property analogous to Theorem 12.

Subsequently, we examine the impact of restriction on compatibility.

Theorem 13 *Let $\succ$ and $\succ_0$ be preference relations over a schema R. Then for every instance r of a relation schema R and every $i = 0,1,2$ if $\succ$ is i-compatible with $\succ_0$, then $[\succ]_r$ is i-compatible with $[\succ_0]_r$. There are cases in which the reverse implications do not hold.*

Proof For 0-compatibility the situation is clear. If there are no 0-conflicts between $\succ$ and $\succ_0$, then there are no 0-conflicts between $[\succ]_r$ and $[\succ_0]_r$. However, for higher-level conflicts, the situation is more complicated.

Assume now that $\succ$ is 1-compatible with $\succ_0$ and consider a 1-conflict between $[\succ]_r$ and $[\succ_0]_r$. Then there are elements $t_1, t_2, s_1, \ldots, s_k$ of r such that

$$(t_1,t_2) \in [\succ_0]_r, (t_2,s_1) \in [\succ]_r, \ldots, (s_k,t_1) \in [\succ]_r,$$

and

$$(t_1,s_k) \notin [\succ_0]_r, \ldots, (s_1,t_2) \notin [\succ_0]_r.$$

Consider now any two elements x and y among $t_1, t_2, s_1, \ldots, s_k$ such that $(x,y) \in [\succ]_r$ (resp.$(x,y) \in [\succ_0]_r$). Clearly then also $x \succ y$ (resp., $x \succ_0 y$). Assume $(x,y) \in [\succ]_r$ and $(y,x) \notin [\succ_0]_r$. Thus $y \not\succ_0 x$. So we obtain a 1-conflict between the preference relations $\succ$ and $\succ_0$. 2-conflicts are analyzed in the same fashion.

To see that the lack of 1-conflicts between $[\succ]_r$ and $[\succ_0]_r$ does not imply the lack of 1-conflicts between $\succ$ and $\succ_0$, consider

$$\succ_0 = \{(c,a)\}$$

$$\succ = \{(a,b),(b,c),(a,c)\},$$

and $r = \{(a,c)\}$. Then $[\succ]_r = \{(a,c)\}$ and $[\succ_0]_r = \{c,a\}$. There are no 1-conflicts between $[\succ]_r$ and $[\succ_0]_r$ but there is a 1-conflict between $\succ$ and $\succ_0$. Analogous examples can be constructed for other kinds of conflicts. □

Finally, we compare the computational properties of $\succ_1, \succ_2$ and $\succ_3$. The preference relation $\succ_1$ is recomputed only after preference revisions. The relation $\succ_2$ is recomputed after every revision and every database update. The recomputation after an update uses $\succ_1$ as a selection condition applied to $r \times r$ (where r is the current relation instance). The relation $\succ_3$ is also recomputed after every revision and every database update. However, in the latter case the computation is more involved than for $\succ_2$, because transitive closure of a finite binary relation needs to be computed. Overall, $\succ_1$ represents the most stable and comprehensive preference information. Even if $\succ_2$ is stored, $\succ_1$ needs to be kept up-to-date after preference revisions, since it is used in the recomputation of $\succ_2$ after an update. The preference relation $\succ_3$ can be stored, revised, and updated without any reference to $\succ_1$. However, in this case some preference information is lost, c.f., Corollary 1.

6.2 Non-intrinsic preferences

Non-intrinsic preference relations are defined using formulas that refer not only to built-in predicates.

Example 9 The following preference relation is not intrinsic:

$$x \succ_{Pref} y \equiv (x,y) \in Pref$$

where $Pref$ is a database relation. One can think of such a relation as representing *stored* preferences.

Revising non-intrinsic preference relations looks problematic. First, it is typically not possible to establish the simplest order-theoretic properties of such relations. For instance, in Example 9 it is not possible to determine the irreflexivity or transitivity of $\succ_{Pref}$ on the basis of its definition. Whether such properties are satisfied depends on the contents of the database relation $Pref$. Second, the transitive closure of a non-intrinsic preference relation may fail to be expressed as a finite formula. Again, Example 9 can be used to illustrate this point.

However, it seems that restriction may be able to alleviate the above problems. Suppose $\succ$ is the original and $\succ_0$ the revising preference relations. Computing $TC(\succ_0 \cup \succ)$ may be infeasible, as indicated above. But computing $TC([\succ_0 \cup \succ]_r)$ is not difficult, as $[\succ_0 \cup \succ]_r$ is computed by the first-order query

$$(x \succ_0 y \vee x \succ y) \wedge x \in R \wedge y \in R.$$

For other composition operators, the same approach also works because they are, like union, defined in a first-order way.

7 Weak-order extensions

Theorems 3 and 5, and Proposition 4 demonstrate that for weak orders one can prove stronger properties about revisions than for general partial orders. The 0-compatibility or the interval order requirements may be relaxed, and the transitive closure computation may no longer be necessary.

So it would be advantageous to work with weak orders. Such orders can, for example, be obtained as *extensions* of the given SPOs. We show here how to express the construction of weak order extensions using Datalog$\neg$ rules (AHV95) and the Rule Algebra (IN88). Although not much can be shown in that framework about WO extensions of arbitrary SPOs, the construction of WO extensions of interval orders (IOs) can be guaranteed to terminate.

7.1 Rules

We define the *application* $r(X)$ of a rule r to an input set of facts X in the standard way.

Definition 10 Assume r is of the form

$$A \leftarrow B_1, \ldots, B_n, \neg C_1, \ldots, \neg C_m.$$

Then $r(X)$ consists of all the facts $\tau(A)$ such that $\tau(B_i) \in X$, $i = 1, \ldots, n$, and $\tau(C_j) \notin X$, $j = 1, \ldots, m$, where τ is a ground substitution. In an *inflationary* application $r(X)$ is added to X.

In this paper, we are dealing with infinite sets of facts represented by constraints. However, the above definition of rule application still applies. From this definition, we can obtain a more operational definition that will tell us how to construct the constraints in the head of the rule r from the constraints in the body (KLP00).

Assume that each goal B_i, $i = 1,\ldots,n$ is described by a constraint β_i and each goal C_j, $j = 1,\ldots,m$ by a constraint γ_j. Also denote by V the set of variables that occur only in the body of r. Then A is described by the formula

$$\exists V.\ \beta_1 \wedge \cdots \wedge \beta_n \wedge \neg\gamma_1 \wedge \cdots \wedge \neg\gamma_m.$$

from which negation and quantifiers have been eliminated.

(IN88) present a language called Rule Algebra (RA) which allows rule composition. The syntax of RA expressions is defined as follows:

$$Expr ::= r \,|\, Expr; Expr \,|\, Expr \cup Expr \,|\, Expr^{+},$$

where r is a single rule. The symbol ";" denotes sequential and "$\cup$", parallel composition. The superscript "$+$" denotes unbounded iteration.

The application of RA expressions is defined as follows (IN88):

- for a single rule it is defined as in Definition 10,
- $(F_1;F_2)(X) \doteq F_2(F_1(X))$,
- $(F_1 \cup F_2)(X) \doteq F_1(X) \cup F_2(X)$,
- $F^{+}(X) \doteq \bigcup_{i>0} F^{i}(X)$.

Like rule application, the application of RA expressions comes in two different variants: inflationary and non-inflationary.

Rule Algebra can be implemented directly. However, (IN88) show also how to map Rule Algebra expressions to a class of *locally-stratified* logic programs (Prz88). This class requires a limited use of function symbols to implement counters.

7.2 Strict partial orders

(Fis85) presents a construction of a WO extension of a finite SPO. It is based on a very simple intuition.

Assume we are given that $x \succ y$ and $y \sim z$, or $x \sim y$ and $y \succ z$. In a weak order one needs to be able to have also $x \succ z$ in both cases (see Proposition 3). Therefore, one could produce a WO extension $\succ'$ of a given SPO $\succ$ by supporting the *derivation* of the implied order relationships. Clearly, such derivation should avoid contradiction ($x \succ' y$ and $y \succ' x$).

Example 10 Consider the following order $\succ = \{(a,c),(b,d)\}$. Thus $a \sim d$ and $b \sim c$. So w could derive $a \succ' b$ and $b \succ' a$, a contradiction.

We construct an extension $\succ'$ of a given SPO $\succ$ using a set of rules. Unfortunately, for infinite orders the construction does not always produce a weak order. The input preference relation $\succ$ is described using a set of facts of the relation T of arity $2n$ where n is the arity of the database relation over which $\succ$ is defined. The output preference relation $\succ'$ is also described as a set of facts of the relation T but those facts are computed using rule application.

First, we have two rules P_{11} and P_{12} for deriving new order relationships:

$$P_{11}:\ T(x,z) \leftarrow T(x,y) \wedge \neg T(z,y) \wedge \neg T(y,z).$$

$$P_{12}:\ T(x,z) \leftarrow T(y,z) \wedge \neg T(x,y) \wedge \neg T(y,x).$$

Second, we have the conflict removal rule P_2:

$$P_2:\ T(x,y) \leftarrow T(x,y) \wedge \neg T(y,x).$$

We note that the rules P_{11}, P_{12}, P_2 need to be applied in a specific order. We use the following Rule Algebra expression E_1 (IN88; AHV95)

$$E_1 = ((P_{11} \cup P_{12})\ ;\ P_2)^+,$$

applied to the input preference relation. In the rule P_2 and the expression E_1, the desired semantics is non-inflationary because we want to eliminate conflicts.

Example 11 Consider the preference relation $\succ = \{(a,c),(b,d)\}$ from Example 10. Applying the rules P_{11} and P_{12} we obtain the relation

$$T(x,y) \equiv x = a \wedge y \neq a \vee x = b \wedge y \neq b \vee x \neq c \wedge y = c \vee x \neq d \wedge y = d.$$

This is not an SPO because, for example, we have $T(a,b)$ and $T(b,a)$. Applying the rule P_2, the conflict is removed, yielding

$$\begin{aligned} T(x,y) \equiv\ & x = a \wedge y \neq a \wedge y \neq b \vee x = b \wedge y \neq b \wedge y \neq a \\ & \vee x \neq c \wedge x \neq d \wedge y = c \vee x \neq c \wedge x \neq d \wedge y = d. \end{aligned}$$

which is a weak order. Thus, no further iterations are necessary.

Denote by T_i the preference relation obtained at the end of the i-th stage in the computation of E_1. Clearly, if T_i is a weak order, then nothing new is produced at the next stage, i.e., $T_{i+1} = T_i$. However, the reverse implication does not have to hold for arbitrary SPOs. Therefore, in each stage i, T_i needs to be separately checked for the weak order property (Proposition 1 implies that the appropriate properties are decidable under the assumption that the input preference relation is described by an ERO preference formula).

Example 12 Consider the following rational-order preference relation $\succ$ adapted from (Fis85):

$$x \succ y \equiv x > y \wedge x \neq 0 \wedge y \neq 0.$$

The corresponding indifference relation $\sim$ is defined as

$$x \sim y \equiv x = y \vee x = 0 \vee y = 0.$$

The relation $\succ$ is not a weak order but even the first iteration of the above rules fails to produce anything new. Consider any rational number $b \neq 0$. There are numbers a and c such that $a > b$, $b > c$, $a \sim 0$ and $c \sim 0$. So on the one hand we have initially $T(b,c)$, $\neg T(c,0)$ and $\neg T(0,c)$, and applying the rule P_{11} we get $T(b,0)$. But on the other hand we have $T(a,b)$, $\neg T(a,0)$ and $\neg T(0,a)$. Applying the rule P_{12} we get $T(0,b)$. Therefore, the rule P_2 does not derive $T(b,0)$, $T(0,b)$, or any other new fact.

It is an open question what kind of properties a preference relation should satisfy so that the condition $T_{i+1} = T_i$ implies the weak order property. (Fis85) shows that such an implication holds for SPOs over finite domains. Therefore, it also holds for finite restrictions of arbitrary SPOs (Section 6). For a finite restriction $[\succ]_r$ a different way for constructing a weak order extension of $[\succ]_r$ is available through the use of *ranking* (Cho03). The "best" tuples – those in $\omega_{\succ}(r)$ – receive rank 1, the "second-best" rank 2 etc. Then the weak order extension $\succ'$ is defined as

$$x \succ' y \equiv rank(x) < rank(y).$$

7.3 Interval orders

For interval orders, we can show stronger results about constructing WO extensions. We still use the Datalog¬/Rule Algebra framework but instead of the expression E_1 we use the following expression E_2:

$$E_2 = (P_{11} \,;\, P_{12})^+.$$

We will see that for E_2 the inflationary and non-inflationary semantics coincide.

For simplicity, we identify here a preference relation with the set of facts of the T predicate describing it.

Example 13 Consider Example 12. Applying the rule P_{11} to the preference relation $\succ$ from this example (which is an interval order) yields the following preference relation $\succ'$:

$$x \succ' y \equiv x > y \wedge x \neq 0 \wedge y \neq 0 \vee x \neq 0 \wedge y = 0.$$

This relation is a total order, and thus also a weak order.

Lemma 2 *For every irreflexive preference relation X, $X \subseteq P_{11}(X)$, $X \subseteq P_{12}(X)$, and $X \subseteq P_{12}(P_{11}(X))$.*

Lemma 3 *Assume X is an interval order preference relation. Then $P_{11}(X)$ and $P_{12}(X)$ are interval order preference relations.*

Proof WLOG, consider $Y = P_{11}(X)$. Clearly, Y is irreflexive. For transitivity, consider $T(x,y) \in Y$ and $T(y,z) \in Y$. Then there is a z' such that $T(x,z') \in X$, $T(z',y) \notin X$, and $T(y,z') \notin X$. Similarly, there is a z'' such that $T(y,z'') \in X$, $T(z'',z) \notin X$, and $T(z,z'') \notin X$. Because X is an interval order, we have $T(x,z'') \in X$ or $T(y,z') \in X$. Assume the former. Then $T(x,z) \in Y$. The preservation of the interval order condition can be shown in a similar way. □

Lemma 4 *Let $F = (P_{11};P_{12})$ and Y be an SPO. Then $F(Y) \subseteq Y$ iff Y is a WO.*

Proof If Y is a WO, then

$$Y = P_{11}(Y) = P_{12}(P_{11}(Y)).$$

If Y is not a WO but an SPO, then there are x, y and z such that $T(x,y) \in Y$, $T(x,z) \notin Y$, $T(z,x) \notin Y$, $T(y,z) \notin Y$ and $T(z,y) \notin Y$. Thus $T(x,z) \in P_{11}(Y)$ and by Lemma 2, $T(x,z) \in P_{12}(P_{11}(Y))$. Thus $P_{12}(P_{11}(Y)) \not\subseteq Y$. □

The following theorem shows that finite termination of the evaluation of E_2 is equivalent to the weak order property.

Theorem 14 *Let X be an IO. For every $i > 0$, $E_2(X) = (P_{11};P_{12})^+(X)$ equals $(P_{11};P_{12})^i(X)$ iff $(P_{11};P_{12})^i(X)$ is a WO.*

Proof Follows from Lemmas 2, 3, and 4. Note that for $j < i$, $(P_{11};P_{12})^j(X) \subseteq (P_{11};P_{12})^i(X)$. It is essential that the given preference relation be an IO. Otherwise, an application of $P_{11};P_{12}$ may produce preference relations which are not SPOs and the equivalence in Lemma 4 may stop to hold. □

To explore the possible implementations of the Rule Algebra expression E_2, we note first that Lemma 2 implies that for the rules P_{11} and P_{12} inflationary and non-inflationary semantics coincide. Therefore, we can use inflationary or non-inflationary languages for the implementation of E_2. (AHV95) indicate that Rule Algebra expressions can be translated to Inflationary Datalog$\neg$ (GS86), a variant of Datalog that allows unstratified negation (necessary here because of the rules P_{11} and P_{12}) at the price of having inflationary semantics. It is clear that Inflationary Datalog$\neg$ programs terminate on finite inputs. However, preference relations are typically infinite. Still, they are finitely representable using preference formulas, and thus we are dealing with the problem of termination of Inflationary Constraint Datalog$\neg$ programs. Fortunately, there are positive results established in this area in (KKR95), which, together with Theorem 14, imply the following:

Theorem 15 *Every interval order preference relation $\succ$, defined using an ERO formula, has a weak order extension $\succ'$, defined using an ERO formula. The formula defining $\succ'$ can be computed in exponential time.*

8 Related work

8.1 Preference change

(Han95) presents a general framework for modeling change in preferences. Preferences are represented syntactically using sets of ground preference formulas, and their semantics is captured using sets of preference relations. Thanks to the syntactic representation preference revision is treated similarly, though not identically, to belief revision (GR95), and some axiomatic properties of preference revisions are identified. The result of a revision is supposed to be minimally different from the original preference relation (using a notion of minimality based on symmetric difference) and satisfy some additional background postulates, for example specific order axioms. (Han95) does not address the issue of constructing or defining revised relations, nor does it study the properties of specific classes of preference relations. On the other hand, (Han95) discusses also preference contraction, and domain expansion and shrinking.

In our opinion, there are several fundamental differences between belief and preference revision. In belief revision, propositional theories are revised with propositional formulas, yielding new theories. In preference revision, binary preference relations are revised with other preference relations, yielding new preference relations. Preference relations are single, finitely representable (though possibly infinite) first-order structures, satisfying order axioms. Belief revision focuses on

axiomatic properties of belief revision operators and various notions of revision minimality. Preference revision focuses on axiomatic, order-theoretic properties of revised preference relations and the definability of such relations (though still taking revision minimality into account).

(Wil97) considers revising a ranking (a WO) of a finite set of tuples with new information, and shows that a new ranking, satisfying the AGM belief revision postulates (GR95), can be computed in a simple way. (Rev97) describes a number of different revision operators for constraint databases. However, the emphasis is on the axiomatic properties of the operators, not on the definability of revised databases. (PFT03) formulates various scenarios of preference revision and does not contain any formal framework. (Won94) studies revision and contraction of finite WO preference relations by single pairs $t_1 \succ_0 t_2$. (Fre04) describes minimal change revision of *rational* preference relations between propositional formulas.

8.2 Preference queries

Two different approaches to preference queries have been pursued in the literature: qualitative and quantitative. In the *qualitative* approach, preferences are specified using binary *preference relations* (LL87; GJM00; Cho02; Cho03; Kie02; KK02). In the *quantitative* utility-based approach, preferences are represented using *numeric utility functions* (AW00; HP04), as shown in Section 4. The qualitative approach is strictly more general than the quantitative one, since one can define preference relations in terms of utility functions. However, only WO preference relations can be represented by numeric utility functions (Fis70). Preferences that are not WOs are common in database applications, c.f., Example 1.

Example 14 There is no utility function that captures the preference relation described in Example 1. Since there is no preference defined between t_1 and t_3 or t_2 and t_3, the score of t_3 should be equal to the scores of both t_1 and t_2. But this implies that the scores of t_1 and t_2 are equal which is not possible since t_1 is preferred over t_2.

This lack of expressiveness of the quantitative approach is well known in utility theory (Fis70). The paper (Cho03) contains an extensive discussion of the preference query literature.

In the earlier work on preference queries (Cho03; Kie02), one can find positive and negative results about closure of different classes of orders, including SPOs and WOs, under various composition operators. The results in the present paper are, however, new. Restricting the relations $\succ$ and $\succ_0$ (for example, assuming the interval order property and compatibility) and applying transitive closure where necessary make it possible to come up with positive counterparts of the negative results in (Cho03). For example, (Cho03) shows that SPOs and WOs are in general not closed w.r.t. union, which should be contrasted with Theorems 1 and 5. In (Kie02), Pareto and prioritized composition are defined somewhat differently from the present paper. The operators combine two preference relations, each defined over some database relation. The resulting preference relation is defined over the Cartesian product of the database relations. So such operators are not useful in the context of revision of preference relations. On the other hand, the careful design

of the language guarantees that every preference relation that can be defined is an SPO.

Probably the most thoroughly studied class of qualitative preference queries is the class of *skyline* queries. A skyline query partitions all the attributes of a relation into DIFF, MAX, and MIN attributes. Only tuples with identical values of all DIFF attributes are comparable; among those, MAX attribute values are maximized and MIN values are minimized. The query in Example 1 is a very simple skyline query (BKS01), with *Make* as a DIFF and *Year* as a MAX attribute. Without DIFF attributes, a skyline is a special case of n-ary Pareto composition.

Various algorithms for evaluating qualitative preference queries are described in (Cho03; TC02), and for evaluating skyline queries, in (BKS01; PTFS03; BGZ04). (BG04) describes how to implement preference queries that use Pareto compositions of utility-based preference relations. In Preference SQL (KK02) general preference queries are implemented by a translation to SQL. (HP04) describes how materialized results of utility-based preference queries can be used to answer other queries of the same kind.

8.3 CP-nets

CP-nets (BBD$^+$04) are an influential recent formalism for reasoning with conditional preference statements under *ceteris paribus* semantics (such semantics is also adopted in other work (MD04; WD91)). We conjecture that CP-nets can be expressed in the framework of preference relations of (Cho03), used in the present paper, by making the semantics explicit. If the conjecture is true, the results of the present paper will be relevant to revision of CP-nets.

Example 15 The CP-net $M = \{a \succ \overline{a}, a : b \succ \overline{b}, \overline{a} : \overline{b} \succ b\}$ where a and b are Boolean variables, captures the following preferences: (1) prefer a to $\overline{a}$, all else being equal; (2) if a, prefer b to $\overline{b}$; (3) if $\overline{a}$, prefer $\overline{b}$ to b. We construct a preference relation $\succ_{C_M}$ between worlds, i.e., Boolean valuations of a and b:

$$\begin{aligned}(a,b) \succ_{C_M} (a',b') \equiv\ & a = 1 \wedge a' = 0 \wedge b = b' \\ \vee\ & a = 1 \wedge a' = 1 \wedge b = 1 \wedge b' = 0 \\ \vee\ & a = 0 \wedge a' = 0 \wedge b = 0 \wedge b' = 1.\end{aligned}$$

Finally, the semantics of the CP-net is fully captured as the transitive closure $TC(\succ_{C_M})$. Such closure can be computed using Constraint Datalog with Boolean constraints (KLP00).

CP-nets and related formalisms cannot express preference relations over infinite domains which are essential in database applications.

9 Conclusions and future work

We have presented a formal foundation for an iterative and incremental approach to constructing ans evaluating preference queries. Our main focus is on *query*

modification, a query transformation approach which works by revising the preference relation in the query. We have provided a detailed analysis of the cases where the order-theoretic properties of the preference relation are preserved by the revision. We have considered a number of different revision operators: union, prioritized and Pareto composition. We have also formulated algebraic laws that enable incremental evaluation of preference queries. Finally, we have studied the strengthening of the properties of preference relations through finite restriction and weak-order extension.

Tables 1 and 2 summarize the closure properties of preference revision under union and prioritized composition. There is no separate table for Pareto composition, because there are only few results specific to this kind of composition.

	$\succ$ SPO	$\succ$ *IO*	$\succ$ WO
$\succ_0$ SPO	not closed	TC SPO if 0-compat.	SPO if 0-compat.
$\succ_0$ IO	TC SPO if 0-compat.	TC SPO if 0-compat.	SPO if 0-compat.
$\succ_0$ WO	SPO if 0-compat.	SPO if 0-compat.	WO if 0-compat.

Table 1 Revision using union

	$\succ$ SPO	$\succ$ *IO*	$\succ$ WO
$\succ_0$ SPO	not closed	TC SPO if 0-compat.	SPO if 0-compat.
$\succ_0$ IO	TC SPO if 1-compat.	TC SPO if 1-compat.	TC SPO if 1-compat.
$\succ_0$ WO	SPO	SPO	WO

Table 2 Revision using prioritized composition

Future work includes the integration of our results with standard query optimization techniques, both rewriting- and cost-based. Semantic query optimization techniques for preference queries (Cho04) can also be applied in this context. Another possible direction could lead to the design of a *revision language* in which richer classes of preference revisions can be specified (GMR97).

One should also consider possible courses of action if the original preference relation $\succ$ and $\succ_0$ lack the property of compatibility, for example if $\succ$ and $\succ_0$ are not 0-compatible in the case of revision by union. Then the target of the revision is an SPO which is the closest to the preference relation $\succ \cup \succ_0$. Such an SPO will not be unique. Moreover, it is not clear how to obtain ipfs defining the revisions. Similarly, one could study *contraction* of preference relations. The need for contraction arises, for example, when a user realizes that the result of a preference query does not contain some expected tuples.

Finally, one can consider preference query transformations which go beyond preference revision, as well as more general classes of preference queries that involve, for example, ranking (Cho03).

References

AHV95. S. Abiteboul, R. Hull, and V. Vianu. *Foundations of Databases*. Addison-Wesley, 1995.

AW00. R. Agrawal and E. L. Wimmers. A Framework for Expressing and Combining Preferences. In *ACM SIGMOD International Conference on Management of Data*, pages 297–306, 2000.

BBD+04. C. Boutilier, R. I. Brafman, C. Domshlak, H. H. Hoos, and D. Poole. CP-nets: A Tool for Representing and Reasoning with Conditional Ceteris Paribus Preference Statements. *Journal of Artificial Intelligence Research*, 21:135–191, 2004.

BG04. W-T. Balke and U. Güntzer. Multi-objective Query Processing for Database Systems. In *International Conference on Very Large Data Bases (VLDB)*, pages 936–947, 2004.

BGZ04. W-T. Balke, U. Güntzer, and J. X. Zhang. Efficient Distributed Skylining for Web Information Systems. In *International Conference on Extending Database Technology (EDBT)*, pages 256–273, 2004.

BKS01. S. Börzsönyi, D. Kossmann, and K. Stocker. The Skyline Operator. In *IEEE International Conference on Data Engineering (ICDE)*, pages 421–430, 2001.

Cho02. J. Chomicki. Querying with Intrinsic Preferences. In *International Conference on Extending Database Technology (EDBT)*, pages 34–51. Springer-Verlag, LNCS 2287, 2002.

Cho03. J. Chomicki. Preference Formulas in Relational Queries. *ACM Transactions on Database Systems*, 28(4):427–466, December 2003.

Cho04. J. Chomicki. Semantic Optimization of Preference Queries. In *International Symposium on Constraint Databases*, pages 133–148, Paris, France, June 2004. Springer-Verlag, LNCS 3074.

Cho06. J. Chomicki. Iterative Modification and Incremental Evaluation of Preference Queries. In *International Symposium on Foundations of Information and Knowledge Systems (FOIKS)*, pages 63–82. Springer, LNCS 3861, 2006.

Fis70. P. C. Fishburn. *Utility Theory for Decision Making*. Wiley & Sons, 1970.

Fis85. P. C. Fishburn. *Interval Orders and Interval Graphs*. Wiley & Sons, 1985.

Fre04. M. Freund. On the Revision of Preferences and Rational Inference Processes. *Artificial Intelligence*, 152:105–137, 2004.

GJM00. K. Govindarajan, B. Jayaraman, and S. Mantha. Preference Queries in Deductive Databases. *New Generation Computing*, 19(1):57–86, 2000.

GMR97. G. Grahne, A. O. Mendelzon, and P. Z. Revesz. Knowledge-base Transformations. *Journal of Computer and System Sciences*, 54(1):98–112, 1997.

GR95. P. Gärdenfors and H. Rott. Belief Revision. In D. M. Gabbay, J. Hogger, C, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 4, pages 35–132. Oxford University Press, 1995.

GS86. Y. Gurevich and S. Shelah. Fixed-Point Extensions of First-Order Logic. *Annals of Pure and Applied Logic*, 32:265–280, 1986.

GSW96. S. Guo, W. Sun, and M.A. Weiss. Solving Satisfiability and Implication Problems in Database Systems. *ACM Transactions on Database Systems*, 21(2):270–293, 1996.

Han95. S. O. Hansson. Changes in Preference. *Theory and Decision*, 38:1–28, 1995.

HP04. V. Hristidis and Y. Papakonstantinou. Algorithms and Applications for Answering Ranked Queries using Ranked Views. *VLDB Journal*, 13(1):49–70, 2004.

IN88. T. Imieliński and S. Naqvi. Explicit Control of Logic Programs through Rule Algebra. In *ACM Symposium on Principles of Database Systems (PODS)*, Austin, Texas, 1988.

ISWGA04. I. F. Ilyas, R. Shah, and A. K. Elmagarmid W. G. Aref, J. S. Vitter. Rank-aware Query Optimization. In *ACM SIGMOD International Conference on Management of Data*, pages 203–214, 2004.

KH03. W. Kießling and B. Hafenrichter. Algebraic Optimization of Relational Preference Queries. Technical Report 2003-1, Institut für Informatik, Universität Augsburg, 2003.

Kie02. W. Kießling. Foundations of Preferences in Database Systems. In *International Conference on Very Large Data Bases (VLDB)*, pages 311–322, 2002.

KK02. W. Kießling and G. Köstler. Preference SQL - Design, Implementation, Experience. In *International Conference on Very Large Data Bases (VLDB)*, pages 990–1001, 2002.

KKR95. P. C. Kanellakis, G. M. Kuper, and P. Z. Revesz. Constraint Query Languages. *Journal of Computer and System Sciences*, 51(1):26–52, August 1995.

KLP00. G. Kuper, L. Libkin, and J. Paredaens, editors. *Constraint Databases*. Springer-Verlag, 2000.

LL87. M. Lacroix and P. Lavency. Preferences: Putting More Knowledge Into Queries. In *International Conference on Very Large Data Bases (VLDB)*, pages 217–225, 1987.

MD04. M. McGeachie and J. Doyle. Utility Functions for Ceteris Paribus Preferences. *Computational Intelligence*, 20(2), 2004.

PFT03. P. Pu, B. Faltings, and M. Torrens. User-Involved Preference Elicitation. In *IJCAI Workshop on Configuration*, 2003.

Prz88. T. C. Przymusinski. On the Declarative Semantics of Deductive Databases and Logic Programs. In J. Minker, editor, *Foundations of Deductive Databases and Logic Programming*, pages 193–216. Morgan Kaufmann Publishers, 1988.

PTFS03. D. Papadias, Y. Tao, G. Fu, and B. Seeger:. An Optimal and Progressive Algorithm for Skyline Queries. In *ACM SIGMOD International Conference on Management of Data*, pages 467–478, 2003.

Rev97. P. Z. Revesz. Model-Theoretic Minimal Change Operators for Constraint Databases. In *International Conference on Database Theory (ICDT)*, pages 447–460. Springer-Verlag, LNCS 1186, 1997.

TC02. R. Torlone and P. Ciaccia. Which Are My Preferred Items? In *Workshop on Recommendation and Personalization in E-Commerce*, May 2002.

WD91. M. P. Wellman and J. Doyle. Preferential Semantics for Goals. In *National Conference on Artificial Intelligence*, pages 698–703, 1991.

Wil97. Mary-Anne Williams. Belief Revision via Database Update. In *International Intelligent Information Systems Conference*, 1997.

Won94. S. T. C. Wong. Preference-Based Decision Making for Cooperative Knowledge-Based Systems. *ACM Transactions on Information Systems*, 12(4):407–435, 1994.