RAPPORT DE RECHERCHE

L R I

COMBINING PREFERENCE RELATIONS : COMPLETENESS AND CONSISTENCY

SPYRATOS N / MEGHINI C

Unité Mixte de Recherche 8623
CNRS-Université Paris Sud – LRI

12/2009

Rapport de Recherche N° 1530

CNRS – Université de Paris Sud
Centre d'Orsay
LABORATOIRE DE RECHERCHE EN INFORMATIQUE
Bâtiment 490
91405 ORSAY Cedex (France)

Combining Preference Relations: Completeness and Consistency

Nicolas Spyratos[1] and Carlo Meghini[2]

[1] Université Paris-Sud, Laboratoire de Recherche en Informatique, Orsay Cedex, France spyratos@lri.fr
[2] Consiglio Nazionale delle Ricerche, Istituto della Scienza e delle Tecnologie della Informazione, Pisa, Italy meghini@isti.cnr.it

Abstract. We introduce two criteria for judging "goodness" of the result when combining preference relations in information systems: completeness and consistency. Completeness requires that the result must be the union of all preference relations, while consistency requires that the result must be an acyclic relation. In other words, completeness requires that the result contain all pairs appearing in the preference relations, and only those pairs; while consistency requires that for every pair (x, y) in the result, it must be able to decide which of x and y is preferred to the other. Obviously, when combining preference relations, there is little hope for the result to satisfy both requirements. In this paper, we classify the various methods for combining preference relations, based on the degree to which the result satisfies completeness and consistency. Our results hold independently of the nature of preference relations (quantitative or qualitative); and also independently of the preference elicitation method (*i.e.* whether the preference relations are obtained by the system using query-log analysis or whether the user states preferences explicitly). Moreover, we assume no constraints whatsoever on the preference relations themselves (such as transitivity, strict ordering and the like).

1 Introduction

The problem considered in this paper is how to combine preferences to arrive at a consensus in the context of databases and information systems. Combining preferences to arrive at a consensus is a problem known in the literature as Social Choice Theory [20]. Variants of this problem, such as voting schemes, have been studied since the 18th century (by Condorcet and Borda). More recently, preferences have been used since the 50s in decision making for ranking alternative choices (see [18] for an extensive survey). In databases and information systems, however, the use of preferences started only in the late 90s and it is mainly concerned with the ranking of query answers [13, 7, 19, 1, 6, 4, 5, 12, 15, 14]. Indeed, with the explosion of the amount of information available today (*e.g.*, on the web) query results may be very large, and ranking these results according to users' preferences is of great help to users.

In the area of information systems, user preferences are classified from different points of view, as follows.

In terms of their nature, preferences can be:

- *Quantitative* (or absolute), expressed by a number on a scale (thus capturing intensity of desire). For example, "I like BMW cars 80%" and "I like VW cars 70%". Quantitative preferences are difficult to express by the casual user, but easy to compute by a machine from query logs.
- *Qualitative* (or relative), expressed by comparison. For example, "I like BMW more than VW". Qualitative preferences express no intensity of desire; they are easy to express by the casual user and also easy to infer by a machine (from quantitative preferences).

In terms of their duration in time, preferences can be:

- *Long-term* preferences; these may be either discovered by the system (unobtrusively, from query logs) or declared explicitly by the user. In both cases long-term preferences are stored in the so-called user profile.
- *Short-term* preferences; these are expressed explicitly by the user, on-line, together with a query (in which case one usually talks about *preference queries*).

We note that the nature and the duration in time are orthogonal characteristics of preferences.

In this paper, we assume a set of objects O and we model a preference over O as a pair (o, o') meaning that o is preferred to o'; moreover, we refer to a set of preferences over O as a *preference relation.*

We stress the fact that we model preference relations as binary relations over O *without* any particular constraint such as transitivity, strict ordering and the like. The reason for doing so is that there is no general consensus in the literature as to what properties a preference relation should satisfy. For instance, although transitivity is generally considered as a desirable property, in several situations preference relations are assumed to be non-transitive [9–11, 22].

Our approach considers only positive preference statements of the kind "x is preferred to y"; in other words, we do not take into consideration indifference relations [18].

When O represents a set of alternative choices, one or more experts may be asked to express their opinion directly on the alternatives, thus leading to a set of preference relations over O. In other cases, the experts may be asked to express their opinion on one or more features of the alternatives, each one inducing a preference relation over O. In both cases, we end up with a set of preference relations over O, and the problem is how to combine them into a single preference relation incorporating as best as possible the opinions of all experts.

In the area of databases and information systems, there are two general methods for combining a set of preference relations: the Prioritized method and the Pareto method, in their restricted and unrestricted versions. Several studies in the literature use these methods, for example for defining preference queries and sky-line queries [3]. However, to our knowledge, no previous work has addressed the problem of how well the information content of the individual preferences is incorporated in the combined relation by these methods.

In this paper, we propose to evaluate Prioritized and Pareto, and their variants, by introducing two criteria: completeness and consistency; and we classify these methods based on the degree to which the combined relations satisfy the two criteria.

Our results hold independently of the nature of preference relations (*i.e.* whether they are qualitative or whether they have been inferred from quantitative rankings); and also independently of the preference elicitation method (*i.e.* whether the preference relations have been obtained by the system, using query-log analysis, or whether they have been stated by the user states, explicitly).

Section 2 gives the formal statement of the problem, as it appears in the area of information systems, also providing some rationale; section 3 introduces the classical methods for combining preference relations; section 4 analyzes how these methods behave with respect to the measures introduced earlier, namely consistency and completeness. Finally, Section 5 summarizes the results and draws some conclusions. Proofs can be found in the appendix.

2 Statement of the problem and rationale

We say that a binary relation P is *complete* with respect to $n > 1$ given binary relations $P_1, \ldots, P_n$ iff P is the union of $P_1, \ldots, P_n$. Moreover, we say that P is *consistent* iff it is acyclic. We note that both completeness and consistency [8] can be tested efficiently.

Given a set of preference relations $P_1, \ldots, P_n$, the problem that we consider is how to find a *combined* preference relation P that satisfies the following requirements:

1. *Completeness.* This property requires that the result should contain all pairs appearing in the preference relations, and only those pairs (*i.e.*, no preference expressed by the user is lost, and no extraneous preference is introduced in the result).
2. *Consistency.* This property requires that the result must be an acyclic relation; that is, for every pair (x, y) appearing in the result, it must be able to decide which of x and y is preferred to the other.

Clearly, when the result is complete, all expressed preferences are taken into account, but there may be contradictions among the individual preference relations, generating cycles in their union. On the other hand, consistency expresses absence of contradictions. So the presence of both completeness and consistency characterizes the optimal situation, where all preferences are taken into account and there is no contradiction.

We note that the presence of contradictions is natural, as they are the result of putting together preferences which either come from different users, independently, or come from the same user but address different aspects of the objects. As we mentioned earlier, such non-contradiction is expressed by the absence of cycles. Indeed, acyclicity allows the ranking of objects, as follows [21, 17]. Let P be an acyclic binary relation (viewed as a digraph), and define the rank of an object o as follows:

- if o is a root of P then $rank(o) = 0$ (a root is a node with no incoming edge);
- else $rank(o)$ is the length of a maximal path among all paths from a root of P to o.

The intuition behind this definition of rank is that the farther an object o is from the best objects (represented by the roots) the less preferred it is. We note that the computational complexity of computing ranks following this definition is linear in the size of the preference graph (understood as usual as the number of nodes plus the number of arcs). In reality, the size of the preference relations is quite small as users often express just a few preferences. Clearly, the definition of rank is sound only if P has at least one root, and this is guaranteed only when P is acyclic.

Now, let us denote by B_i the set of objects with rank i, and let m be the maximal path length among all paths starting from a root. Then it is easy to prove the following proposition:

Proposition 1. *The sequence $B_0, B_1, \ldots, B_m$ defined above has the following properties:*

- *$B_0, B1, \ldots, B_m$ is a partition of the set of objects appearing in P;*
- *for each $i = 0, 1, 2, \ldots, m$, there is no arc of P connecting two objects of B_i;*
- *for each $i = 1, 2, \ldots, m$ and each object $t \in B_i$ there is an object $s \in B_{i-1}$ such that there exists an arc from s to t in P.*

In general, when combining preference relations, there is little hope for the result to satisfy both completeness and consistency. Since ranking is important in information systems and acyclicity is a sufficient condition in order to do ranking, a reasonable approach to follow when combining preference relations is to satisfy acyclicity while *minimizing* the loss of completeness. In order to achieve this goal, we must select as the result P of combining the given preference relations a largest acyclic subset of the union. In other words, we must select P in such a way that there is no proper superset of P which is an acyclic subset of the union.

If we view preference relations as digraphs, finding such a P is equivalent to solving the maximum acyclic sub-graph problem, which is stated as follows: Given a digraph $G = (V, E)$, find a maximum cardinality subset $E' \subseteq E$ such that (V, E') is acyclic. This problem is known to be NP-hard [16].

Now, since there have been several proposals in the literature for combining preference relations, one wonders how these proposals relate to the maximum acyclic sub-graph problem. This is the question that we undertake in the rest of this paper.

More specifically, we show that the classical methods for combining preference relations provide a clever trade-off between maximality of the result and efficiency of computation. As these methods were introduced long before the theory of complexity was formulated, this is a rather surprising result that provides an *a posteriori* justification for the introduction of these methods.

3 Classical methods for combining preference relations

There are several well-known methods for combining a set of preference relations $P_1, \ldots, P_n$, into a single preference relation, most notably, the so-called Prioritized and Pareto (and their variants). In order to formally define these methods, we use the following terminology from [2]. Let P be a preference relation:

- we use interchangeably the notations $(x, y) \in P$ and xPy; if $(x, y) \notin P$, we write $x\overline{P}y$;
- x and y are said to be *equivalent* with respect to P, written $xP^{\equiv}y$, iff both xPy and yPx hold;
- if $x\overline{P}y$ and $y\overline{P}x$, we say that x and y are *incomparable* with respect to P, written $xP^{\#}y$;
- finally, if xPy and $y\overline{P}x$, x is said to be *strictly preferred* to y with respect to P, written $xP^{<}y$.

We first recall the definition of $P_{\cup}$, the union of the given preference relations $P_1, \ldots, P_n$:

$$xP_{\cup}y \iff \exists i.(xP_iy)$$

A basic difference between Prioritized and Pareto is that the former, apart from the given preference relations $P_1, \ldots, P_n$, requires some additional information in order to be applied, whereas the latter requires no additional information. Indeed, in the case of Prioritized, one assumes that the preference relations $P_1, \ldots, P_n$, are ordered by a priority relation $\prec$, which is a strict partial order. We recall that a strict partial order is a binary relation which is irreflexive and transitive, and consequently asymmetric (asymmetric means that if $a < b$ holds then $b < a$ does not hold).

Prioritized and Pareto each come into two variants:

1. *Restricted Prioritized* (RPR, for short).
   This is the "classical" Prioritized rule, also used in [2]. Given a set of preference relations $P_1, \ldots, P_n$ strictly ordered by $\prec$, the RPR rule defines a binary relation $Pr_r(P_1, \ldots, P_n, \prec)$, or simply Pr_r when there is no ambiguity, as follows:

$$xPr_ry \iff \forall i.(xP_iy \ \vee \ \exists j.(j \prec i \wedge xP_j^{<}y))$$

$$xPr_r^{\#}y \iff \exists i.(xP_i^{\#}y\&\forall j$$

   Notice that x and y are incomparable in the combined relation (*i.e.*, $xPr_r^{\#}y$) if and only if x and y are incomparable on the preference relation with the highest priority for which they are not equivalent.

We note that the well-known lexicographic ordering is a special case of Restricted Prioritized. Indeed, the lexicographic ordering is a Prioritized ordering, in which:

- O stands for a set of words of finite length over some finite alphabet A, where A is totally ordered by some strict total order $<_A$ (*i.e.*, given any two distinct letters a and a', either $a <_A a'$ or $a' <_A a$ but not both); the minimum element of $<_A$ is the special character *blank;*
- we have as many preference relations P_i as the number of characters in the longest word in O;
- the i-th preference relation captures preference based on the i-th letter of words (starting from the left), according to $<_A$; so, $carlo P_1 nicolas$ while $nicolas P_3 carlo$;
- the priority relation over the P_i's is the natural total order $<_N$ over the indices.

Under these assumptions, given two words $w = a_1 a_2 \dots a_m$ and $w' = a'_1 a'_2 \dots a'_n$, the RPR rule becomes:

$$wPr_r w' \iff \forall i.(a_i <_A a'_i \ \vee \ \exists j.(j <_N i \wedge a_j <_A^{<} a'_j))$$

Now, since $<_A$ is a strict total order, $<_A = <_A^{<}$, therefore the rule becomes:

$$wPr_r w' \iff \forall i.(a_i <_A a'_i \ \vee \ \exists j.(j < i \wedge a_j <_A a'_j))$$

The last rule is clearly the rule used to order words in a dictionary.

2. *Unrestricted Prioritized* (UPR, for short).
This method extends RPR by allowing the combined preference to hold even in presence of incomparability in some preference relation, as long as there exists comparability in some other relation. The UPR rule defines a binary relation Pr_u as follows:

$$xPr_u y \iff \forall i.(xP_i y \ \vee \ (xP_i^{\#} y \wedge \exists k.(xP_k y)) \ \vee \ \exists j.(j \prec i \wedge xP_j^{<} y))$$

Notice that x and y are incomparable if and only if x and y are incomparable in *all* given preference relations.

3. *Restricted Pareto* (RPA, for short).
In this method, x is preferred to y in the combined relation if and only if for every i, x is preferred to y, and for at least one P_i x is strictly preferred to y. Formally, the RPA rule defines a binary relation Pa_r as follows:

$$xPa_r y \iff \forall i.(xP_i y) \ \wedge \ \exists j.(xP_j^{<} y)$$

4. *Unrestricted Pareto* (UPA, for short).
In this method, x is preferred to y if and only if for at least one P_i x is strictly preferred to y, and for no j y is strictly preferred to x (in all other preferences x and y can be comparable in the same direction, incomparable or equivalent). The UPA rule defines a binary relation Pa_u as follows:

$$xPa_u y \iff \forall i.(y\overline{P_i^{<}}x) \ \wedge \ \exists j.(xP_j^{<} y)$$

$$
\begin{aligned}
xP_{\cup}y &\iff \exists i.(xP_iy)\\
xPr_ry &\iff \forall i.(xP_iy \ \vee \ \exists j.(j \prec i \wedge xP_j^{<}y))\\
xPr_uy &\iff \forall i.(xP_iy \ \vee \ (xP_i^{\#}y \wedge \exists k.(xP_ky)) \ \vee \ \exists j.(j \prec i \wedge xP_j^{<}y))\\
xPa_ry &\iff \forall i.(xP_iy) \ \wedge \ \exists j.(xP_j^{<}y)\\
xPa_uy &\iff \forall i.(y\overline{P_i^{<}}x) \ \wedge \ \exists j.(xP_j^{<}y)
\end{aligned}
$$

Fig. 1. Definitions of the classical methods

Based on the definitions given so far (and summarized in Figure 1 for convenience), we can prove the following proposition (we recall that proofs can be found in the appendix).

Proposition 2. *Let $P_1, \ldots, P_n$, be n preference relations and let $\prec$ be any strict ordering on them. Then, the following hold:*

1. $Pa_r \subseteq Pr_r \subseteq Pr_u \subseteq P_{\cup}$
2. $Pa_r \subseteq Pa_u \subseteq Pr_u \subseteq P_{\cup}$
3. Pa_u *and* Pr_r *are incomparable with respect to set-containment.*

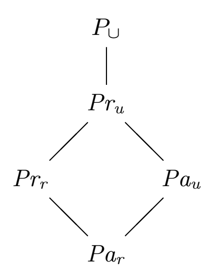

Fig. 2. Classical methods for combining preference relations

Figure 2 summarizes graphically the set-containment relations established by proposition 2.

4 Completeness and consistency of classical methods

In order to characterize the classical methods with respect to completeness and consistency, we consider three cases as follows:

Case 1 the union $P_\cup$ is acyclic.

In this case, consistency is guaranteed and we show that completeness holds only for the unrestricted methods.

Case 2 the union $P_\cup$ is cyclic but each individual preference relation is acyclic.

In this case, we show that the restricted methods lead to an acyclic result, but they are incomplete in the sense that they may not produce a maximum acyclic sub-graph; on the other hand, the unrestricted methods may produce a cyclic result.

Case 3 one or more individual preference relations are cyclic.

In this case, all methods may produce a cyclic result.

4.1 Acyclic union

It follows from proposition 2 that whenever the union of the preference relations is acyclic, so are all the variants of Prioritized and Pareto defined above. In this case, then, consistency is satisfied. It remains to be seen whether completeness is satisfied as well.

The next lemma states a consequence of the acyclicity of the union that will be used in what follows.

Lemma 1. *Let $P_1, \ldots, P_n$, be preference relations whose union $P_\cup$ is acyclic. If $xP_\cup y$ then for some i we have that $xP_i^< y$, and for all $k \neq i$ we have that either $xP_k y$ or $xP_k^\# y$.*

We can now prove the following.

Proposition 3. *Let $P_1, \ldots, P_n$ be preference relations whose union $P_\cup$ is acyclic. Then,*

1. $P_\cup = Pa_u$
2. $P_\cup = Pr_u$ *for any strict order $\prec$ of the preference relations.*

Notice that in proving that $Pr_u \subseteq P_\cup$, the acyclicity of the union is not required.

Concerning Pr_r and Pa_r, we have that in some cases $P_\cup$ is acyclic and that $Pr_r \subset P_\cup$ and $Pa_r \subset P_\cup$. In proof, let us consider the example shown in Figure 3; using RPR to combine P_1 and P_2 leads to the incomparability of 3 and 4 (if $P_1 \prec P_2$) or the incomparability of 1 and 2 (if $P_2 \prec P_1$), while both these pairs are in $P_\cup = P_1 \cup P_2$. Analogously, it can be seen that $(3, 4) \in Pa_r^\#$ and $(1, 2) \in Pa_r^\#$.

We can then conclude that when the union of the given preference relations is acyclic, both RPR and RPA are incomplete, while UPR and UPA are "optimal", *i.e.* both complete and consistent.

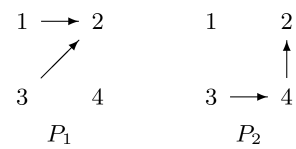

Fig. 3. Preference relations

4.2 Cyclic union with acyclic preference relations

We now consider the case in which the union of the preference relations is cyclic while each individual preference relation is acyclic. In this case, we are interested to know (a) whether Prioritized and Pareto produce an acyclic combined preference relation, and, if yes, (b) how close to the union $P_\cup$ they are.

In order to answer these questions, we first derive necessary and sufficient conditions for the acyclicity of the result in each of the classical methods. We then apply these conditions to answer the above questions.

We recall that a cycle in a binary relation P is a sequence of objects $C = (o_0, o_1, \ldots, o_k)$ with $k \geq 2$, such that

- $o_0 = o_k$,
- $o_{i-1} P o_i$ for each $i = 1, \ldots, k$, and
- there is no repetition in $o_0, o_1, \ldots, o_{k-1}$.

Proposition 4. *Let $P_1, \ldots, P_n$, be n preference relations and let $\prec$ be any strict ordering on them. Then, the Restricted Prioritized Pr_r is acyclic iff for each cycle $C = (o_0, \ldots, o_k)$ in $P_\cup$ there exists an arc (o_{u-1}, o_u) such that for some preference relation P_i, $o_{u-1}\overline{P_i}o_u$ and for each j such that $o_{u-1}P_j o_u$, either $o_u P_j o_{u-1}$ or $i \prec j$.*

The last Proposition states the condition under which an undesired preference (o, o') (*i.e.*, any one of the preferences found in a cycle in $P_\cup$) does not end up in Pr_r. The condition is derived by combining two facts: oP_io' and $o\overline{Pr_r}o'$ (the latter in turn obtained by negating the RPR rule). By a similar technique, we obtain the analogous conditions for the other considered variants of Prioritized and Pareto. The proofs of the corresponding Propositions are omitted, as they are very similar to that of the previous Proposition.

Proposition 5. *Let $P_1, \ldots, P_n$, be n preference relations and let $\prec$ be any strict ordering on them. Then, the Unrestricted Prioritized Pr_u is acyclic iff for each cycle $C = (o_0, \ldots, o_k)$ in $P_\cup$ there exists one arc (o_{u-1}, o_u) such that for some preference relation P_i, $o_u P_i^{<} o_{u-1}$ and for each j such that $o_{u-1}P_j o_u$, either $o_u P_j o_{u-1}$ or $i \prec j$.*

Proposition 6. *Let $P_1, \ldots, P_n$, be n preference relations. Then, the Restricted Pareto preference relation Pa_r is acyclic iff for each cycle $C = (o_0, \ldots, o_k)$ in $P_\cup$ there exists one arc (o_{u-1}, o_u) such that either $o_{u-1}\overline{P_i}o_u$ for some preference relation P_i, or $o_u P_j o_{u-1}$ for all preference relations P_j.*

Proposition 7. *Let $P_1, \ldots, P_n$, be n preference relations. Then, the Unrestricted Pareto Pa_u is acyclic iff for each cycle $C = (o_0, \ldots, o_k)$ in $P_\cup$ there exists one arc (o_{u-1}, o_u) such that $o_u P_i^{<} o_{u-1}$ for some preference relation P_i or $o_{u-1} P_j o_u$ implies $o_u P_j o_{u-1}$) for all j.*

Having established under which condition each of the considered methods produces a cyclic relation, we now examine whether this condition can ever occur.

Proposition 8. *The Restricted Prioritized Pr_r of n acyclic preference relations $P_1, \ldots, P_n$ is acyclic for any strict order $\prec$ on $P_1, \ldots, P_n$.*

The last Proposition says that when each one of the given preference relations is acyclic, the necessary and sufficient conditions for the acyclicity of Pr_r established by Proposition 4 are always met. This can be verified by considering that under the circumstances, every cycle in $P_\cup$ involves at least two preference relations, one of which has necessarily a lower priority than the other; thus, any arc (x, y) in the cycle coming from a preference relation with a lower priority can play the role of (o_{u-1}, o_u) in Proposition 4.

The last Proposition shows that in presence of a cyclic union, the RPR rule produces an acyclic relation. However, the rule turns out to be too restrictive in that it may exclude from the result preferences that are unproblematic (*i.e.*, preferences which do not produce any cycle if included in the result). In other words, RPR may not produce a maximum acyclic sub-graph of $P_\cup$.

Consider for instance the following preference relations: $P_1 = \{(2,3)\}$, $P_2 = \{(1,2)\}$ and $P_3 = \{(3,2)\}$, together with the following strict order: $P_1 \prec P_2 \prec P_3$. In this case, we have $Pr_r = \{(2,3)\}$, which is a non-maximum acyclic sub-graph of $P_\cup$, since it misses the preference $(1,2)$.

Furthermore, there may be maximum acyclic sub-graphs of $P_\cup$ that RPR cannot define under any $\prec$. The example presented in Figure 4 is a case in point: we have $Pr_r = P_1$ (if $P_1 \prec P_2$) or $Pr_r = P_2$ (if $P_2 \prec P_1$). But for no ordering of the preference relations we can have $Pr_r = R_1$ or $Pr_r = R_2$ (also shown in Figure 4).

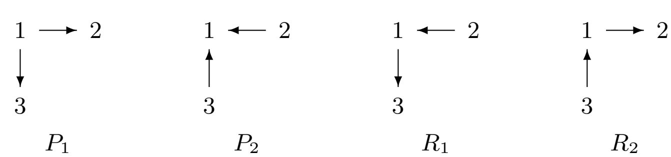

Fig. 4. Prioritized combinations

Let us now consider Pr_u. It is easy to construct an example that falsifies the condition of Proposition 5 (*i.e.*, a cyclic $P_\cup$ leading to a cyclic Pr_u): Assume $P_1 \prec P_2$ and $P_1 = \{(1,2)\}$ whereas $P_2 = \{(2,3),(3,1)\}$. In this case, none of the arcs making up the cycle $(1, 2, 3, 1)$ in $P_\cup$ has a reverse arc as required by Proposition 5. In fact, $Pr_u = P_\cup$ and so we have a cyclic Pr_u.

For Pareto, the situation is identical.

Corollary 1. *The Restricted Pareto Pa_r of n acyclic preference relations $P_1, \ldots, P_n$ is acyclic.*

As for Proposition 8, the acyclicity of each preference relation P_i implies that the condition established by Proposition 6 are always met. In particular, each cycle in $P_\cup$ comes from preferences belonging to different relations; then, any arc (x, y) in the cycle is missed at least by one P_i, *i.e.* $x\overline{P_i}y$, and as such it plays the role of (o_{u-1}, o_u) in Proposition 6.

Since Pa_r is a subset of Pr_r (Proposition 2), we have that the above remarks related to the completeness of Pr_r carry over to Pa_r.

Finally, the previous example can be used also to show that a cyclic $P_\cup$ may lead to a cyclic Pa_u.

We may then conclude that when each preference relation is acyclic but their union is cyclic, both Pr_r and Pa_r are acyclic but incomplete. On the other hand, their unrestricted versions Pr_u and Pa_u may be cyclic.

4.3 Cyclic preference relations

In considering cyclic preference relations, we distinguish between two cases:

- the cycle involves only two objects.
  We call these objects *equivalent* in the sense defined earlier.
- the cycle involves at least three objects.

In the classical methods, object equivalence arises only in very special circumstances or not at all, as shown by the following corollary (that follows from Propositions 4 to 7):

Corollary 2. *Let $P_1, \ldots, P_n$, be n preference relations and let $\prec$ be any strict ordering on them. Then, for any two objects x, y*

- *$xPr_r^{\equiv}y$ iff $xP_i^{\equiv}y$ for all $i = 1, \ldots, n$, and*
- *$xPr_u^{\equiv}y$ iff the following hold:*
  1. *for some i we have $xP_i^{\equiv}y$, and*
  2. *for all $j = 1, \ldots, n$ we have either $xP_j^{\equiv}y$ or $xP_j^{\#}y$.*

Moreover, both Pa_r and Pa_u are asymmetric.

Thus, two objects are equivalent with respect to Restricted Prioritized if they are equivalent with respect to each preference relation. In this case they are incomparable for both versions of Pareto. For Unrestricted Prioritized, two objects are equivalent if they are comparable on some dimension but in no dimension one of the two is strictly preferred to the other. Pareto rules out object equivalence by being asymmetric in both its variants.

For cycles involving three or more objects, the situation is much worse. If some of the $P_1, \ldots, P_n$ contain such a cycle, then all variants of Prioritized and Pareto may produce a cyclic result.

The following example shows that this is indeed the case for RPA. Suppose $P_1 = P_2 = \{(1,2),(2,3),(3,1)\}$. Then, it is easy to see that $Pa_r = P_1$. Notice that if we assume transitivity of the P_i's, then for all objects x and y involved in a cycle, we have $xP_i^{\equiv}y$ and therefore $(x,y) \notin Pa_r$. In the example, transitivity of the P_i's implies that $P_i^{<} = \emptyset$ for all i and consequently $Pa_r = \emptyset$.

Since Pa_r is the smallest relation among those produced by the classical methods, we have that all methods can produce a cyclic result.

5 Concluding remarks

From an information system perspective, there are two basic requirements that must be satisfied when combining preference relations into a single preference relation: completeness and consistency. Completeness means that the result must be the union of all preference relations, while consistency means that the result must be an acyclic relation.

However, it is rarely possible to satisfy these two requirements simultaneously, due to the fact that the union of the given preference relations may be cyclic. Given the importance of acyclicity in ranking the objects, we have argued that a reasonable compromise is to aim at a maximum acyclic sub-graph of the union of the given relations.

We have analyzed two classical methods for combining preference relations, Prioritized and Pareto, each in two variants: restricted and unrestricted. We have shown that all four methods are inadequate with respect to the above requirements, as they may produce a result that is either incomplete or cyclic. In particular,

- In the (fully unproblematic) case when the union of the given preference relations is acyclic, both restricted approaches are incomplete, while the unrestricted ones are optimal (*i.e.*, complete and acyclic).
- When each preference relation is acyclic but their union is cyclic, both restricted approaches produce an acyclic result, which however looses more preferences than needed to obtain acyclicity; under the same circumstances, the unrestricted methods may produce a cyclic result.
- Finally, when one or more preference relations have cycles involving at least three objects, all four methods may produce a cyclic result.

Thus, all four classical methods are for various reasons unsatisfactory. However, if we look at the problem from a purely computational perspective, we are faced with the fact that computing a maximum acyclic sub-graph of a given graph is known to be NP-hard. This fact sheds a different light on the classical approaches.

In fact we may conclude that both Prioritized and Pareto trade off efficiency to optimality (unless P=NP). In their unrestricted variants, both methods achieve optimality only when it is computationally easy to do so, (*i.e.*, when the union is acyclic). In all other cases, they retain efficiency while loosing optimality.

From a practical point of view, there are several options:

If the size of the problem is not prohibitive, one can go for the optimal solution and compute a maximum acyclic sub-graph of the union of the given preference relations. In this respect, we make the following observation. Let A be a binary relation and A^o stand for the set of all maximum acyclic sub-graphs of A. When some of the P_i's are cyclic, for any $X \in (P_\cup)^o$ and for any $Y \in \cup(P_i^o)$ we have that $X \subseteq Y$, but not necessarily $Y \subseteq X$ (we recall that $P_\cup$ is the union of the preference relations $P_1, \ldots P_n$). As a consequence, it is more efficient to compute a minimum acyclic sub-graph from $P_\cup$ rather than from maximum acyclic sub-graphs of the individual preference relations.

If the size of the problem is prohibitive, three alternatives are possible. First, one might want to go for an approximation technique for computing the maximum acyclic sub-graph [16]. Second, one can use a classical method. As our results show, Unrestricted Prioritized is clearly the best candidate. Finally, we mention the following method, based on the premise that objects on the same preference cycle can be regarded as indistinguishable. Define the following equivalence relation $\equiv$ over the objects appearing in $P_\cup : x \equiv y$ iff either $x = y$ or $x \neq y$ and x and y are on the same cycle in the graph of $P_\cup$. Next, define a binary relation P over the set of equivalence classes as follows: $[x]P[y]$ iff $x'P_\cup y'$ for some $x' \in [x]$ and $y' \in [y]$. Finally, apply the algorithm given in Section 2 to rank the equivalence classes in P.

In this paper we have considered preference relations simply as binary relations without any additional constraint. We are currently investigating how different constraints on preference relations (such as transitivity) might influence the classification of the classical methods with respect to completeness and consistency. We also plan to investigate the trade-offs between completeness and consistency with respect to more general classes of schemes, such as schemes following Condorcet's Principle.

References

1. R. Agrawal and E. L. Wimmers. A framework for expressing and combining preferences. In *Proc. of ACM SIGMOD*, pages 297–306, Dallas, USA, 2000.
2. H. Andreka, M. Ryan, and P.-Y. Schobbens. Operators and laws for combining preference relations. *Journal of Logic and Computation*, 12(1):13–53, 2002.
3. R.I. Brafman and C. Domshlak. Preference handling - an introductory tutorial. *AI Magasine*, 30(1), 2009.
4. J. Chomicki. Querying with intrinsic preferences. In *Proceedings of the 8th International Conference on EDBT*, pages 34–51, Prague, Czech Rep., 2002.
5. J. Chomicki. Semantic optimization of preference queries. In *Proc. of 1st Int. Sym. on Appl. of Constraint Databases*, number 3074 in LNCS, 2004.
6. J. Chomicki. Iterative modification and incremental evaluation of preference queries. In *Proc. of FoIKS*, pages 63–82, 2006.
7. Jan Chomicki. Preference formulas in relational queries. *ACM Trans. Database Systems*, 28(4):427–466, 2003.

8. T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein. *Introduction to algorithms.* The MIT Press, 2nd edition, 2001.
9. Peter C. Fishburn. Nontransitive preferences in decision theory. *Journal of Risk and Uncertainty*, 4(2):113–124, April 1991.
10. Oswald Huber. Nontransitive multidimensional preferences: Theoretical analysis of a model. *Theory and Decision*, 19:147–165, January 1979.
11. J. Kacprzyk and M. Roubens, editors. *Non-conventional preference relations in decision making.* Springer-Verlag, 1988.
12. W. Kiessling and G. Köstler. Preference SQL design, implementation, experiences. In *In Proceedings of 28th International Conference on Very Large Data Bases*, pages 990–1001, Hong Kong, China, 2002.
13. Werner Kiessling. Foundations of preferences in database systems. In *Proc. of VLDB*, pages 311–322, 2002.
14. Georgia Koutrika and Yannis E. Ioannidis. Constrained optimalities in query personalization. In *Proc. of ACM SIGMOD*, pages 73–84, 2005.
15. Georgia Koutrika and Yannis E. Ioannidis. Personalized queries under a generalized preference model. In *Proc. of ICDE*, pages 841–852, 2005.
16. Alantha Newman. Approximatng the maximum acyclic subgraph. Master's thesis, Massachussets Institute of Technology, Dept. of Electrical Engineering and Computer Science, 2000.
17. E.M. Nguer and N. Spyratos. A user-friendly interface for evaluating preference queries over tabular data. In *Proc. of SIGDOC08, the 26th ACM International Conference on Design of Communication*, Lisbon, Portugal, Sept 2008.
18. M. Öztürk, A. Tsoukiàs, and Ph. Vincke. Preference modelling. In J. Figueira M. Ehrgott, S. Greco, editor, *Multiple Criteria Decision Analysis: State of the Art Surveys,*, pages 27–73. Springer Verlag, 2005.
19. P.Georgiadis, I.Kapantaidakis, M.Nguer, N.Spyratos, and V.Christophides. Efficient rewriting algorithms for preference queries. In *Proc. of ICDE08, the 24th International Conference on Data Engineering*, Cancun, Mexico, April 2008.
20. Amartya Sen. Social choice theory. In K. J. Arrow and M.D. Intriligator, editors, *Handbook of Mathematical Economics*, volume 3 of *Handbook of Mathematical Economics*, chapter 22, pages 1073–1181. Elsevier, September 2005.
21. N. Spyratos and V. Christophides. Querying with preferences in a digital library. In *Dagstuhl Seminar (N 05182), Federation over the Web*, number 3847 in LNAI, May 2005.
22. Tversky. Intransitivity of preferences. *Psychological review*, 76(1), 1969.

A Proofs

Proposition 2. *Let $P_1, \ldots, P_n$, be n preference relations and let $\prec$ be any strict ordering on them. Then, the following hold:*

1. $Pa_r \subseteq Pr_r \subseteq Pr_u \subseteq P_\cup$
2. $Pa_r \subseteq Pa_u \subseteq Pr_u \subseteq P_\cup$
3. *Pa_u and Pr_r are incomparable with respect to set-containment.*

Proof. *1 and 2 follow immediately from the definitions, except for $Pa_u \subseteq Pr_u$. In order to prove it, we show that for any x, y $x\overline{Pr_u}y$ implies $x\overline{Pa_u}y$. By unfolding the definition, we have that $x\overline{Pr_u}y$ iff:*

$$\exists i\{x\overline{P_i}y \bigwedge [x\overline{P_i^{\#}}y \vee \forall k(x\overline{P_k}y)] \bigwedge \forall j(j \prec i \rightarrow x\overline{P_j^{<}}y)\}$$

or equivalently:

$$\exists i\{x\overline{P_i}y \bigwedge x\overline{P_i^{\#}}y \bigwedge \forall j(j \prec i \rightarrow x\overline{P_j^{<}}y)\} \bigvee \tag{1}$$
$$\{x\overline{P_i}x \bigwedge \forall k(x\overline{P_k}y) \bigwedge \forall j(j \prec i \rightarrow x\overline{P_j^{<}}y)\}. \tag{2}$$

Now, the first two conjuncts in (1) are equivalent to $yP_i^{<}x$, while (2) is equivalent to $\forall k(x\overline{P_k}y)$, so we have

$$x\overline{Pr_u}y \Longleftrightarrow \exists i\{yP_i^{<}x \bigwedge \forall j(j \prec i \rightarrow x\overline{P_j^{<}}y)\} \bigvee \forall k(x\overline{P_k}y).$$

On the other hand,

$$x\overline{Pa_u}y \Longleftrightarrow \exists i\{yP_i^{<}x \bigvee \forall j(x\overline{P_j^{<}}y)\}$$

or equivalently:

$$x\overline{Pa_u}y \Longleftrightarrow \exists i\{yP_i^{<}x \bigvee \forall j(x\overline{P_j}y \vee yP_jx)\}.$$

Clearly, $x\overline{Pr_u}y$ implies $x\overline{Pa_u}y$, hence $Pa_u \subseteq Pr_u$.
Finally, in order to show 3, we give counterexamples. In order to show that $Pa_u \subset Pr_r$, let us consider the two preference relations P_1 and P_2 shown in Figure 5.(a), with $P_1 \prec P_2$. In this case, we have $\emptyset = Pa_u \subset Pr_r = \{(1,2)\}$. On the other hand, with the preference relations P_1 and P_2 shown in Figure 5.(b), with $P_1 \prec P_2$, we have $\emptyset = Pr_r \subset Pa_u = \{(1,2)\}$.

(a)	(b)
$P_1 : 1 \rightarrow 2$	$P_1 :$
$P_2 : 1 \leftarrow 2$	$P_2 : 1 \rightarrow 2$

Fig. 5. Preference relations

For any binary relation A, $A^\star$ denotes the transitive closure of A.

Lemma 1. *Let $P_1, \ldots, P_n$, be preference relations whose union $P_\cup$ is acyclic. If $xP_\cup y$ then for some i we have that $xP_i^{<}y$, and for all $k \neq i$ we have that either xP_ky or $xP_k^{\#}y$.*
Proof. *If $xP_\cup y$, then for some i xP_iy. Acyclicity implies that $(y,x) \notin (P_\cup)^\star$. In particular, $y\overline{P_j}x$ for all k. This means that for all k, either xP_ky or $xP_k^{\#}y$. Moreover, xP_iy and $y\overline{P_i}x$ means $xP_i^{<}y$.*

Proposition 3. *Let $P_1, \ldots, P_n$ be preference relations whose union $P_\cup$ is acyclic. Then,*

1. $P_\cup = Pa_u$
2. $P_\cup = Pr_u$ *for any strict order* $\prec$ *of the preference relations.*

Proof. *We have already observed that* $Pr_u \subseteq P_\cup$. $P_\cup \subseteq Pr_u$ *follows from Lemma 1 and the UPR rule. The proof that* $P_\cup = Pa_u$ *is very similar.*

Proposition 4. *Let* $P_1, \ldots, P_n$, *be* n *preference relations and let* $\prec$ *be any strict ordering on them. Then, the Restricted Prioritized* Pr_r *is acyclic iff for each cycle* $C = (o_0, \ldots, o_k)$ *in* $P_\cup$ *there exists an arc* (o_{u-1}, o_u) *in* C *such that for some preference relation* P_i, $o_{u-1}\overline{P_i}o_u$ *and for each* j *such that* $o_{u-1}P_jo_u$, *either* $o_uP_jo_{u-1}$ *or* $i \prec j$.
Proof. $(\rightarrow)$ *Suppose for each cycle* C *in* $P_\cup$ *there exists one arc* (o_{u-1}, o_u) *satisfying the hypotheses. Then by definition of* Pr_r, $o_{u-1}\overline{Pr_r}o_u$, *therefore* Pr_r *is acyclic.*
$(\leftarrow)$ *Conversely, if* Pr_r *and* $P_\cup$ *are both acyclic, the proposition is vacuously satisfied. Now suppose that* Pr_r *is acyclic while* $P_\cup$ *is cyclic. As a consequence, on each cycle* C *in* $P_\cup$ *there exists some arc* $(o, o') \in P_\cup \setminus Pr_r$. *Since* $o\overline{Pr_r}o'$, (o, o') *does not satisfy the RPR rule. By negating the rule, the proposition follows.*

Proposition 8. *The Restricted Prioritized* Pr_r *of* n *acyclic preference relations* $P_1, \ldots, P_n$ *is acyclic for any strict order* $\prec$ *on* $P_1, \ldots, P_n$.
Proof. *Since* $Pr_r \subseteq P_\cup$, *if* $P_\cup$ *is acyclic so is* Pr_r. *Then, suppose* C *is a cycle in* $P_\cup$. *Then, suppose* C *is a cycle in* $P_\cup$, *and let* P_i *be a* $\prec$*-minimal preference relation which must exist because* $\prec$ *is finite. As all preference relations are acyclic,* P_i *must fail on some arc of the given cycle* C. *Then* Pr_r *evidently also fails on this arc. It follows that it is acyclic.*

Corollary 1. *The Restricted Pareto* Pa_r *of* n *acyclic preference relations* $P_1, \ldots, P_n$ *is acyclic.*
Proof. *As already observed,* $Pa_r \subset Pr_r$. *By the previous Proposition,* Pr_r *is acyclic, so* Pa_r *is acyclic too.*