ELSEVIER Applied Mathematics and Computation 94 (1998) 227–241

APPLIED MATHEMATICS AND COMPUTATION

A weak order solution to a group ranking and consistency-driven pairwise comparisons

Ryszard Janicki [a,*], Waldemar W. Koczkodaj [b,1]

[a] *Department of Computer Science and Systems, McMaster University, Hamilton, Ont., Canada L8S 4K1*

[b] *Department of Mathematics and Computer Science, Laurentian University, Sudbury, Ont., Canada P3E 2C6*

Abstract

A novel non-numerical algorithm, based on a *partial order*, for solving a *group ranking* problem is compared with the more traditional *consistency-driven* approach. Both approaches are based on the fundamental concept of *pairwise comparisons. Potential applications are discussed.*

Keywords: Group ranking; Pairwise comparisons; Partial and weak orders

1. Introduction

One of the basic problems of knowledge engineering is related to a teamwork in which a knowledge engineer elicits knowledge from human experts, refines it with them, and represents it in the knowledge base. The knowledge may be, for example, expressed by assessing preferences of stimuli (e.g., criteria, factors, or possible alternatives). When devising methods for formulating and assessing preferences, a knowledge engineer has to take into account the limitations in human capabilities for undertaking such an endeavour.

* Corresponding author. E-mail: janicki@mcmaster.ca.

[1] E-mail: waldemar@ramsey.cs.laurentian.ca.

PII: S0096-3003(97)10080-7

Formally, the problem can be formulated as follows (cf. [1,2]). Let us assume that there is a finite set X of objects (e.g., criteria or stimuli) and m experts, numbered $1, 2, \ldots, m$. Each expert i compares objects and constructs a *non-empty weak order* $<_i \subseteq X \times X$. The problem of a group ranking is to define one *group ranking*, $<_g \subseteq X \times X$, which is a "compromise" based on all the individual rankings $\mathrm{IR} = \{<_1, \ldots, <_m\}$. It is possible that some of the individual rankings may collide with other individual rankings. The group ranking $<_g$ is expected to be a *weak order* but it should be somehow *consistent* with most (if not all) of the individual rankings IR.

Various instances of the above problem occur in the Social Choice Theory, Decision Making Process and Knowledge-Based Systems (see [2–4], and many others). Oddly enough *there are not* many diversified solutions to the group ranking problem. Arrow's Impossibility Theorem [2,3], says that under certain assumptions (known as Arrow's Axioms) the above problem has no general solution has been a deterrent in seeking any solution. Arrow's Axioms, perfectly proper in the framework of the Social Choice Theory (for which they were originally formulated [2]) are just too strong and restrictive for many other applications. Nevertheless it had held back the branch of research based on other approaches (e.g., the concept of the weak order) for more than four decades. The recent paper [1] presents a range of fairly general solutions satisfying four of the six Arrow's Axioms. They are in our opinion sufficient for a wide range of applications.

The approach of [1] starts with the observation that if $|X| = 2$ then a variety of solutions exists (for example, a simple majority voting, see [3]). This means that the problem can be reformulated as follows. Let $R \subseteq X \times X$ be a relation (called *a ranking relation*, [1]) such that xRy if y is preferred over x by the experts. (We assume that for $X_{x,y} = \{x, y\}$ the solution for finding the preference is known.) The problem is now reduced to the derivation of $<_g$ from R. The problem of finding *a global ranking* (or *weights*) on the basis of known local ranks (or relative weights) of *all pairs* is called *a pairwise comparisons method.* It originates from Fechner (see [5]) and was formulated for the first time by Thurstone [6] and further enhanced by Saaty who proposed a hierarchical structure (see [7]) and a definition of global inconsistency. The pairwise comparisons method is based on a reasonable observation that while the precise ranking (weighting) of n objects over *unqualified* property is difficult (almost impossible for $n > 7$ where an appropriate hierarchy structure is assumed for larger n), it usually can be done with a respective precision for $n = 2$. The "traditional" pairwise comparisons method ([6,7,4] and others) does not use partial orders. It uses real numbers and linear algebra instead (which in fact can be treated as a special case of a *total order*). It it worthwhile to note that the "traditional" pairwise comparisons method *does not* satisfy Arrow's Axioms (see [2,3]) but it is quite popular in solving practical problems (see [8]). For other approaches (that are not based on the pairwise comparisons) for finding the global ranking (e.g., a utility function), the reader is referred to [3]. In this paper we briefly describe the two above mentioned methods: traditional and weak orders based pairwise comparisons and compare them.

2. An example of group ranking problem

Let us consider seven objects (in this case criteria) $X = \{a, b, c, d, e, f, g\}$, four experts, and their individual rankings $\mathrm{IR} = \{<_1, <_2, <_3, <_4\}$ (provided by Table 1).

The individual rankings $<_i$ are, respectively, represented by the following step-sequences:

$$\{a\}\{b\}\{f\}\{c\}\{d\}\{e\}\{g\},$$
$$\{d\}\{e\}\{g\}\{a\}\{f\}\{c\}\{b\},$$
$$\{b\}\{e\}\{g\}\{a\}\{f\}\{c\}\{d\},$$
$$\{a\}\{f\}\{c, d, e\}\{g\}\{b\}.$$

We are tempted to solve the group ranking problem by simple totals in columns. We receive $(12, 14, 21, 17, 13, 17, 18, 17)$ for (a, b, c, d, e, f, g, h), respectively. It gives the following ranks $(1, 3, 8, 4–6, 2, 4–6, 7, 4–6)$ where 4–6 means a tie for ranks 4, 5, and 6. The above algorithm silently assumes that ordinal numbers $1, 2, \ldots, n$ are corresponding weights for the first, second, etc. ranks which is not derived from any system of axioms or substantiated by any other logical or practical reason. A combined medal classification in Olympic Games, for example, is sometimes based on the following weights: 5 for gold, 3 for silver, 2 for bronze, and 1 for the fourth place (weight 4 has not been used in the above rating). Weight 5 for gold and 3 for silver reflects our emotions that the gold medal is harder to win than the silver medal. Problems with ranking still exist even if we consciously assume ordinal numbers to be weights. Let us look at the expert number 4. He/she has ranked criteria c, d, and e equally and with less-

Table 1
Rankings provided by the individual experts

Expert	Criteria						
	a	*b*	*c*	*d*	*e*	*f*	*g*
1	1	2	4	5	6	3	7
2	4	7	6	1	2	5	3
3	4	1	6	7	2	5	3
4	1	5	3	3	3	2	4

er importance than criterion a and f. This, however, does not mean that criteria c, d, and e are ranked at the third place. In fact, we have to assume that criteria c, d, and e tie for the third, fourth, and fifth place unless there are compelling reasons to the contrary. This also means that the values assigned to criteria c, d, and e should be 4 (that is, $(3+4+5)/3$) rather than 3 and subsequent places for criteria g and b are 6 and 7 (please note that in the original Table 1 no criterion has been ranked as sixth or seventh according to expert 4). After making the above modification in Table 1 (that is substituting 3 with 4, 4 with 6, and 5 with 7 in the last row) we can see that the new column totals are $(10, 17, 20, 17, 14, 15, 19)$. It gives new ranks $(1, 4\text{–}5, 7, 4\text{–}5, 2, 6)$ which are evidently different from the above. For example, criterion b jumped from the position 3–4 to 4–5, d, shares now position 4–5 with b, and f is now in the third place.

Clearly, weights expressing ranks cannot be fixed forever as ordinal numbers since these numbers may be differently assigned in case of a tie. The weights must depend on a case. A silent assumption of ordinal numbers for weights is not reasonable. Arrow's Impossibility Theorem (see Section 3) is quite explicit stating in [2,3] that under certain assumptions (known as Arrow's Axioms) the above problem has no general solution. A similar problem exists with the individual experts. The "straight total" algorithm assumes again that the judgements of all experts are equally important. This simplification does not seem to be reasonable when we assume a different level of expertise or knowledge for individual experts.

In this paper two algorithms (both based on pairwise comparisons) for a group ranking will be analyzed: a non-numerical (based on a weak order approach in [1]) and traditional pairwise comparisons (see [4]). Both approaches assume more relaxed assumptions than Arrow's Axioms but provide constructive solutions.

3. Weakly ordered pairwise comparisons

Let us recall some basic concepts from [1]. A relation $<\subseteq X \times X$ is called a (*sharp*) *partial order* if it is irreflexive and transitive, i.e., if $a < b \Rightarrow \dot{b} < a$, and $a < b < c \Rightarrow a < c$, for all $a, b, c \in X$. We will write $a \sim b$ if $\neg a < b \wedge \neg b < a \wedge a \neq b$ (where $\neg$ denotes *not*), that is if a and b are distinct incomparable elements of X. A partial order is:

- *total* if $\sim$ is empty, i.e., for all $a, b \in X$. $a < b \vee b < a$,
- *weak* [9] if $a \sim b \sim c :\Rightarrow: a \sim c \vee a = c$, i.e., if $\sim \cup \mathrm{id}_X$ is an equivalence relation (id_X is an identity on X).

Evidently, every total order is weak. A wide and complete analysis of axioms for *preference* and *indifference* relations is presented in [3]. It is worthwhile noting that in the majority of cases the preference relation is just a weak order, while indifference corresponds to $\sim \cup \mathrm{id}_X$. The charm of weak orders is that they can be represented uniquely by *value functions* (compare, for example, [3]). More formally, a partial order $<\subseteq X \times X$ is weak if and only if there is a total order $<_t \subseteq Y \times Y$, and a mapping (value function) $v : X \to Y$, such that $a < b \iff v(a) <_t v(b)$. Weak orders can easily be represented by *step-sequences*. For instance, if $X = \{a, b, c, d, e\}$, and $v(a) = v(c) = 0$, $v(e) = 2$, $v(b) = v(d) = 5$, then the weak order $<$ is uniquely represented by the step-sequence $\{a, c\}\{e\}\{b, d\}$.

3.1. Ranking relation

The first step in our approach is to define and construct a *ranking relation*.

Let X be our set of objects to be ranked and $\mathrm{IR} = \{<_1, \ldots, <_m\}$ be the set of individual (weakly ordered) rankings. By a (*binary*) *ranking relation* over the set IR we mean any relation $R \subseteq X \times X$ satisfying the following constrains:

- $\forall x, y \in X$: $|\{i \mid x <_i y\}| \geqslant |\{i \mid y <_i x\}| \Rightarrow \neg yRx$, and
- $\forall x, y \in X$: $(\forall i = 1, \ldots, m$: $x <_i y) \Rightarrow xRy$,

where for every set A, $|A|$ denotes the number of elements of A.

An obvious example of a ranking relation is R_{smv}, *a simple majority voting relation*, where

$$xR_{\mathrm{smv}}y \iff |\{i \mid x <_i y\}| > |\{i \mid y <_i x\}|,$$

i.e., $xR_{\mathrm{smv}}y$ if more experts preferred y over x. The relation R_{smv} exists for every IR, and every ranking relation over IR satisfies $R \subseteq R_{\mathrm{smv}}\cdot R_{\mathrm{smv}}$ is not, however, the only ranking relation one may think of. If for instance, 10 experts would say x is better than y, and 11 just the opposite, the ranking that x and y are of the same value seems to be the most appropriate in many cases (except for sports and politics!).

In general, one would consider the value of $|\{i \mid x <_i y\}| - |\{i \mid y <_i x\}|$ to establish the relation between x and y. If it is "big", we assume xRy. For "small" or "very small", we would rather assume that $\neg xRy$ and $\neg yRx$. Note that in both cases $\neg yRx$.

The second condition says that if all experts prefer y over x then the entire group does, so xRy. The relation R is always irreflexive, i.e., $\neg xRx$.

If, by chance, the relation R is a weak order, our goal has been achieved. Unfortunately the relation R may not even be a partial order. The precise definition of R depends on the specific properties of the set X.

Unanimity has a central place in ranking theory and practice. If every expert prefers y to x, then the group should prefer y to x (see *Pareto's principle* [3]). Let us define the relation $<_U \subseteq X \times X$, as

$$<_U = <_1 \cap \cdots \cap <_m .$$

The relation $<_{\mathrm{U}}$ will be called *unanimous preference*. Unanimous preference is always a partial order (since all $<_i$ are partial orders), although it is not necessarily a weak order. The following corollary follows immediately from the definition of $<_{\mathrm{U}}$ and the second constraint on R. We require that the group ranking $<_{\mathrm{g}}$ satisfies $<_{\mathrm{U}} \subseteq <_{\mathrm{g}}$.

3.2. Approximation of ranking relations by preference orders

As we mentioned above, the ranking relation R may not be a partial order. Our goal is to find a relation $<_{\mathrm{g}}$ which is not only a weak order but also "the best" approximation of R. This will be done in two steps. First we will find $<^{\mathrm{U}}_{R^+}$, "the best" partial order approximation of R. Next we will discuss various weak order approximations of $<^{\mathrm{U}}_{R^+}$.

The problem is that the set X is believed to be partially ordered but the data acquisition process is so influenced by informational noise, imprecision, randomness, or expert ignorance, that the collected data R is only some relation on X. We may say that R gives a fuzzy picture, and to focus it, we must do some pruning and/or extending. If R is not a partial order then either it is not reflexive, or it is not transitive, or both. Suppose that R is not transitive. This could be seen as a limitation of the discriminatory power of the data acquisition process. Let us define the relation $R^+ = \bigcup_{i=1}^{\infty} R^i$, where $R^{i+1} = R^i \circ R$, and $\circ$ denotes composition of relations. Evidently $R \subseteq R^+$ and R^+ is transitive. If R is transitive, then $R^+ = R$. The relation R^+ may not be irreflexive. If R^+ is not irreflexive, we have $\exists x,y \in X.\ xR^+y \wedge yR^+x$. Such a situation could be interpreted as a side effect of the data noise, uncertainty, randomness, etc.

The relation $<_{R^+}$, defined as $x <_{R^+} y \iff xR^+y \wedge \neg yR^+x$, will be called a *partial order approximation of* (*ranking relation*) R (see [1] for details). We have $<_{R^+} \subseteq R^+$, and if R^+ is irreflexive then $R^+ = <_{R^+}$.

However, if R^+ is not irreflexive, it may happen that $<_{\mathrm{U}} \setminus <_{R^+} \neq \emptyset$, so $<_{\mathrm{U}}$ may be not included in $<_{R^+}$.

Let us define $<^{\mathrm{U}}_{R^+}$ as the *smallest partial order* satisfying the following constraint:

$$<_{\mathrm{U}} \cup <_{R^+} \subseteq <^{\mathrm{U}}_{R^+} .$$

The relation $<^{\mathrm{U}}_{R^+}$ will be called the *proper partial order approximation of the ranking relation* R. If $<_{\mathrm{U}} \subseteq <_{R^+}$ then $<_{R^+} = <^{\mathrm{U}}_{R^+}$.

If $<^{\mathrm{U}}_{R^+}$ is a weak order, we can set $<_{\mathrm{g}} = <^{\mathrm{U}}_{R^+}$. The property $\forall x,y \in X.\ x <_{\mathrm{g}} y : \Rightarrow : \neg yRx$ can be interpreted as a weaker version of Arrow's Axiom of Binary Relevance (see [3] for details).

If $<^{\mathrm{U}}_{R^+}$ is not a weak order, we need to find a "good" weak order approximation. To do this we will follow the approach suggested in [9].

Let us assume that set X is believed to be weakly ordered (by, for example, a group ranking relation $<_g$) but the discriminatory power of the data acquisition process, which seeks to uncover this order, is limited. For example, experts may not always know enough about the problem and may make educated guesses or even select choices at random, the number of experts may be insufficient, or instruments may not be precise enough, etc. The acquired data establish only a partial order $<_p$ which is a partial picture of the underlying order. We seek, however, an extension process which is expected to identify correctly the ordered pairs that are not part of the data. To specify this problem in a formal way, we need a new concept. Let $<_p \subseteq X \times X$ be a partial order defined as $<_p = <^{U}_{R^+}$. Let us define $\approx_p \subseteq X \times X$ as follows:

$$x \approx_p y \iff (\forall z \in X.\ z \sim_p x \iff z \sim_p y) \vee x = y$$

Various properties of $\approx_p$ are analyzed in [9]. We need to recall only one here: a partial order $<_p$ is weak if and only if $x \approx_p y \iff x \sim_p y \vee x = y$.

It is worthwhile noting that weak order extensions reflect the fact that if $x \approx_p y$ then *all reasonable methods* for extending $<_p$ will have x equivalent to y in the extension since there is nothing in the data that distinguishes between them (for details see [9]).

We will say that a weak order $<_w \subseteq X \times X$ is *a proper weak order extension of* $<_p$ is and only if:

$$(x <_p y \Rightarrow x <_w y) \quad \text{and} \quad (x \approx_p y \Rightarrow x \sim_w y \vee x = y).$$

If X is finite then for every partial order $<_p$ there always exists a proper weak extension. If $<_p$ is weak, then its only weak extension is $<_w = <_p$. If $<_p$ is not weak, there is usually more than one such extension. We will examine three of such extensions. Let $v_i : X \to \text{Integers}, i = 1, 2, 3$, be the following (value) functions for all $x \in X$:
(1) $v_1(x) = \mid \{y \mid x <_p y\} \mid$,
(2) $v_2(x) = \mid \{y \mid y <_p x\} \mid$,
(3) $v_3(x) = \mid \{y \mid x <_p y\} \mid - \mid \{y \mid y <_p x\} \mid$
Let us define $<_{wi} \subseteq X \times X$, $i = 1, 2, 3$, as follows:
- $\forall x, y \in X.\ x <_{wi} y \iff v_i(x) < v_i(y)$, if $i = 1, 3$, and
- $\forall x, y \in X.\ x <_{w2} y \iff v_2(x) > v_2(y)$,

where $<$ is the standard total order of integers.

The weak order $<_{w3}$ comes from [9], while v_{w1} and $<_{w2}$ correspond to the left- and right-normal forms for partially commutative monoids (cf. [1]). More algorithms for finding various proper weak extensions are presented in [9]. The choice of an appropriate weak order extension depends on the problem considered. If someone is looking for "the best objects", the order $<_{w1}$ seems to be the appropriate choice. For locating "the troublemakers" (that is, for example controversial or extreme opinions), the order $<_{w2}$ is more promising, while $<_{w3}$ corresponds to "the safe opinions".

3.3. The weak order algorithm

Let X be our set of objects and $\mathrm{IR} = \{<_1, \ldots, <_m\}$ be a set of individual rankings of X. We propose the following algorithm for finding the group ranking $<_g$.

- Construct a ranking relation R over IR (as shown in Section 2),
- construct the proper partial order of R as $<^{U}_{R^+}$ (as shown in Section 3),
- if $<^{U}_{R^+}$ is a weak order then define a group ranking as $<_g = <^{U}_{R^+}$,
- otherwise $<_g = <_w$ where $<_w$ is one of the proper weak order extensions of $<^{U}_{R^+}$.

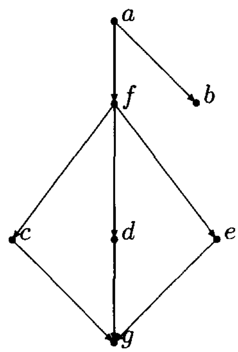
Let us now consider the example from Section 2. The ranking relation R (defined as a simple majority voting, $R = R_{smv}$) for this example is illustrated in Fig. 1. The order $<^{U}_{R^+}$ is illustrated in Fig. 2. The value functions $v_1(x)$, $v_2(x)$, and $v_3(x)$ applied to $<^{U}_{R^+}$ give weak orders $<_{w1}$, $<_{w2}$, and $<_{w3}$, which are, respectively, represented by the following step-sequences:

$$\{a\}\{b,f\}\{c,d,e\}\{g\},$$
$$\{a\}\{f\}\{b,c,d,e\}\{b,g\},$$
$$\{a\}\{f\}\{b,c,d,e\}\{g\}.$$

Table 2 illustrates the three weak extensions of $<^{U}_{R^+}$ constructed in Fig. 2. We may define $<_g$ as $<_{w_i}$, $i = 1,2,3$, or any other proper weak order extension of $<^{U}_{R^+}$.

	a	b	c	d	e	f	g
a		1	1	1		1	
b	-1						
c	-1			1			
d	-1		-1		1	-1	1
e				-1			1
f	-1			1			
g				-1	-1		

Fig. 1. Tabular representation of relation R derived from Table 1 (R is here the simple majority voting, $R = R_{smv}$); $aRb \Rightarrow 1$ in row a and column b, -1 in row b and column a.



Relation $<^U_{R+}$

Fig. 2. Hasse diagrams of the relations $<^U_{R+}$.

Table 2
Three weak extensions of the partial order $<^U_{R+}$ in Fig. 2

Rank	$v_1(x)$	x	$v_2(x)$	x	$v_3(x)$	x
1	0	a	7	a	−6	a
2	1	b,f	5	f	−4	f
3	2	c,d,e	1	e	1	c,e,b,d
4	3	g	0	b	3	g

The compliance of the above algorithm with some of the consistency rules proposed in [2] as well as necessary proofs are shown in [1].

4. Pairwise comparisons

An $n \times n$ pairwise comparisons matrix is defined as a square matrix $A = [a_{ij}]$ such that $a_{ij} > 0$ for every $i, j = 1, \ldots, n$. Each a_{ij} expresses a relative preference of criterion (or stimulus) s_i over criterion s_j for $i, j = 1, \ldots, n$ represented by numerical weights (positive real numbers) w_i and w_j respectively. The quotients $a_{ij} = w_i/w_j$ form a pairwise comparisons matrix

$$A = \begin{vmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \frac{1}{a_{12}} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \cdots & 1 \end{vmatrix}.$$

A pairwise comparisons matrix A is called *reciprocal* if $a_{ij} = 1/a_{ji}$ for every $i,j = 1,\dots,n$ (then automatically $a_{ii} = 1$ for every $i = 1,\dots,n$ because they represent the relative ratio of a criterion against itself). A pairwise comparisons matrix A is called *consistent* if $a_{ij} \cdot a_{jk} = a_{ik}$ holds for every $i,j,k = 1,\dots,n$ since $(w_i/w_j)/(w_j/w_k)$ is expected to be equal to w_i/w_k. Although every consistent matrix is reciprocal, the converse is generally not true. In practice, comparing of s_i to s_j, s_j to s_k, and s_i to s_k often results in inconsistency amongst the assessments in addition to their inaccuracy; however, the inconsistency may be computed and used to improve the accuracy (for details, see [4]).

The first step in pairwise comparisons is to establish the relative preference of each combination of two criteria. A scale from 1 to 5 can be used to compare all criteria in pairs. (Values from the interval $[\frac{1}{5}, 1]$ reflect inverse relationships between criteria since $w_i/w_j = 1/(w_j/w_i)$.) The consistency-driven approach is based on the reasonable assumption that by finding the most inconsistent judgements, one can then reconsider one's own assessments. This in turn contributes to the improvement of judgemental accuracy. Consistency analysis is a dynamic process which is assisted by the software.

Saaty's theorem [7] states that for every $n \times n$ consistent matrix $A = [a_{ij}]$ there exist positive real numbers $w_1,\dots,w_n$ (weights corresponding to criteria $s_1,\dots,s_n$) such that $a_{ij} = w_i/w_j$ for every $i,j = 1,\dots,n$. According to [7], the weights w_i are unique up to a multiplicative constant and the principal eigenvector corresponding to the largest eigenvalue of A provides weights w_i which we wish to obtain from the set of preferences a_{ij}. This is not the only possible solution to the weights problem. In the past, the least-squares solution was known, but it was far more computationally demanding than finding an eigenvector of a matrix with positive elements. Later, a method of row geometric means was proposed, which is the simplest and the most effective method of finding weights. A statistical experiment [10] demonstrated that the accuracy, that is, the distance from the original matrix A and the matrix A' reconstructed from weights with elements $a'_{ij} = [w_i/w_j]$, does not strongly depend on the method. There is, however, a strong relationship between the accuracy and consistency. Consistency analysis is the main focus of the consistency-driven approach.

An important problem is how to begin the analysis. Assigning weights to all criteria (e.g., $A = 18$, $B = 27$, $C = 20$, $D = 35$) seems more natural than the above process. In fact it is a recommended practice to start with some initial values. The above values yield the ratios: $A/B = 0.67$, $A/C = 0.9$, $A/D = 0.51$, $B/C = 1.35$, $B/D = 0.77$, $C/D = 0.57$. Upon analysis, these may look somewhat suspicious because all of them round to 1, which is of equal or unknown importance. This effect frequently arises in practice, and experts are tempted to change the ratios by increasing some of them and decreasing others (depending on the knowledge of the case). The changes usually cause an increase of inconsistency which, in turn, can be handled by the analysis because it contributes to establishing more accurate and realistic weights. There is an increase in accuracy (in excess of 300%) demonstrated by a Monte Carlo experiment with bars of randomly generated lengths [4].

We may assume the values show in Table 3 when we compare them in pairs for example in Table 1:

Table 3

	a	b	c	d	e	f	g
a	1	$\frac{5}{4}$	$\frac{5}{3}$	$\frac{5}{2.5}$	$\frac{5}{2}$	$\frac{5}{1.5}$	$\frac{5}{1}$
b	$\frac{4}{5}$	1	$\frac{4}{3}$	$\frac{4}{2.5}$	$\frac{4}{2}$	$\frac{4}{1.5}$	$\frac{4}{1}$
c	$\frac{3}{5}$	$\frac{3}{4}$	1	$\frac{3}{2.5}$	$\frac{3}{2}$	$\frac{3}{1.5}$	$\frac{3}{1}$
d	$\frac{2.5}{5}$	$\frac{2.5}{4}$	$\frac{2.5}{3}$	1	$\frac{2.5}{2}$	$\frac{2.5}{1.5}$	$\frac{2.5}{1}$
e	$\frac{2}{5}$	$\frac{2}{4}$	$\frac{2}{3}$	$\frac{2}{2.5}$	1	$\frac{2}{1.5}$	$\frac{2}{1}$
f	$\frac{1.5}{5}$	$\frac{1.5}{4}$	$\frac{1.5}{3}$	$\frac{1.5}{2.5}$	$\frac{1.5}{2}$	1	$\frac{1.5}{1}$
g	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2.5}$	$\frac{1}{2}$	$\frac{1}{1.5}$	1

The arbitrarily assumed values presented in Table 3 are for illustration purpose only. In practice we start with values which can be based on statistical research (e.g., questionnaire), group discussions, established by Delphi method, etc. They are further refined by a consistency-driven analysis (see Section 5). We can compute the final weights by taking geometric means of columns and normalizing them. In our case they are 26.32, 21.05, 15.79, 13.16, 10.52, 7.89, and 5.26 (in fact they are quite round in an unnormalized form which we took as $(5, 4, 3, 2.5, 2, 1.5, 1)$ to simplify the computation of the quotients for the pairwise comparisons matrix). After substituting ordinal numbers by their respective weights in Table 1, we received the new ranking: $(1, 3\text{–}4, 7, 5, 2, 3\text{–}4, 6)$ which is different from the last ranking in Section 2, however (and only accidentally) equal to the first ranking. How come? The answer is evident. We received our ranking by comparing criteria in pairs. These local subjective comparisons are done with full consciousness and express our preferences. The global weights are computed on the basis of the local preferences in an objective way. It is quite different. Let us see what happens when we alter our pairwise comparisons table (Table 3) to Table 4:

Table 4

	a	*b*	*c*	*d*	*e*	*f*	*g*
a	1	$\frac{5}{3.5}$	$\frac{5}{3}$	$\frac{5}{2.5}$	$\frac{5}{2}$	$\frac{5}{1.5}$	$\frac{5}{1}$
b	$\frac{3.5}{5}$	1	$\frac{3.5}{3}$	$\frac{3.5}{2.5}$	$\frac{3.5}{2}$	$\frac{3.5}{1.5}$	$\frac{3.5}{1}$
c	$\frac{3}{5}$	$\frac{3}{3.5}$	1	$\frac{3}{2.5}$	$\frac{3}{2}$	$\frac{3}{1.5}$	$\frac{3}{1}$
d	$\frac{2.5}{5}$	$\frac{2.5}{3.5}$	$\frac{2.5}{3}$	1	$\frac{2.5}{2}$	$\frac{2.5}{1.5}$	$\frac{2.5}{1}$
e	$\frac{2}{5}$	$\frac{2}{3.4}$	$\frac{2}{3}$	$\frac{2}{2.5}$	1	$\frac{2}{1.5}$	$\frac{2}{1}$
f	$\frac{1.5}{5}$	$\frac{1.5}{3.5}$	$\frac{1.5}{3}$	$\frac{1.5}{2.5}$	$\frac{1.5}{2}$	1	$\frac{1.5}{1}$
g	$\frac{1}{5}$	$\frac{1}{3.5}$	$\frac{1}{3}$	$\frac{1}{2.5}$	$\frac{1}{2}$	$\frac{1}{1.5}$	1

The new weights are $(27.3, 18.92, 16.20, 13.52, 10.81, 8.11, 5.41)$ and the new ranking is $(1, 2\text{–}4, 7, 5, 2\text{–}4, 2\text{–}4, 6)$ which is different from any of the previously computed rankings.

5. Consistency analysis

Consistency analysis is critical to the approach presented here because the solution accuracy of *not-so-inconsistent* matrices strongly depends on the inconsistency [10]. The challenge to the pairwise comparisons method comes from a lack of consistency in the pairwise comparisons matrices which arises in practice. Given an $n \times n$ matrix A that is not consistent, the theory attempts to provide a consistent $n \times n$ matrix A' that differs from matrix A "as little as possible". Unlike the old eigenvalue-based inconsistency [7,4], the triad-based inconsistency (introduced in [4]) locates the most inconsistent triads. This allows the user to reconsider the assessments included in the most inconsistent triad. It has been shown [11] that the global inconsistency decreases with the reduction of the local inconsistency. It is fair to say that making comparative judgements of rather intangible criteria (e.g., environmental pollution or public safety) results not only in imprecise knowledge, but also in inconsistency in our own judgements. In practice, inconsistent judgements are unavoidable when at least three factors are independently compared against each other.

To illustrate the consistency-driven approach let us consider that assessments in Table 1 of our four experts (labeled A, B, C, and D) are not of equal importance but expressed by a pairwise comparisons matrix in Fig. 3. The pair-

	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
A	1	$\boxed{1}$	$\boxed{5}$	3	1	2.5	5	3	1	$\boxed{2.5}$	5	$\boxed{3}$	1	2.5	5	2.5
B		1	$\boxed{2}$	2		1	$\boxed{2}$	$\boxed{2}$		1	2	$\boxed{1}$		1	2	1
C			1	0.5			1	$\boxed{0.5}$			1	0.5			1	0.5
D				1				1				1				1
	inc=0.6				inc=0.5				inc=0.17							

Fig. 3. Process of consistency improvement.

wise comparisons method can be used to convert their relative importance into weights. Let us concentrate our attention on the ratios of the three experts: A, B, and C. The given estimates $A/B = 1$, $B/C = 2$, and $A/C = 5$ mean that the relative assessments of expert A when compared with expert B is of the same importance (reflected by the value of 1 on a scale 1–5), judgements of expert's B have a relative importance of 2 when compared with C, and relative importance of expert C is 5 when compared with expert A. Evidently something does not "add up" because $(A/B) \cdot (B/C) = 1 \cdot 2 = 2$, which obviously is not equal to 5 (that is, A/C). With an inconsistency index of 0.60, triad (A, B, C) (with "boxed" values of 1, 5, and 2) is the most inconsistent in the entire pairwise comparisons matrix (reciprocal values below the main diagonal are not shown in Fig. 3). A rash judgment may lead us to believe that A/C should indeed be 2, but we do not have any reason to reject the estimation of A/B as 2.5 or A/B as 2. After correcting A/B from 1 to 2.5 (an arbitrary decision which is usually based on additional knowledge gathering), the next most inconsistent triad is $(2, 2, 0.5)$ with an inconsistency index of 0.5. An adjustment of 2 to 1 makes this triad fully consistent ($2 \cdot 0.5$ is 1), but now another triad $(2.5, 3, 1)$ has still an inconsistency of 0.17. By changing 3 to 2.5 the entire table becomes fully consistent. The corrections for real data are done on the basis of professional experience and knowledge of the case by examining all three involved criteria.

An acceptable threshold of inconsistency is assumed to be 0.33 since it reflexes a situation where the worst triad has a judgement which is not more than two grades of the assumed scale 1–5 (or their inverses: 1 to 1/5) apart from the remaining two judgments [4]. We could have stopped the inconsistency improvement process after the third step in Fig. 3. There is no need to continue decreasing the inconsistency, as only a high value is harmful. A very small value may indicate that the artificial data were entered hastily without reconsideration of former assessments. Step 4 in Fig. 3 has been done only for the purpose of illustration.

The weight received for the last pairwise comparisons matrix in Fig. 3 are: $(0.5, 0.2, 0.2, 0.1)$. Using them to multiply weights corresponding to ranks given by experts are $91, 2, 4, 5–6, 5–6, 3, 7)$.

6. Conclusions

Approximation of partial orders by weak orders is useful in knowledge-based systems because of the weak order extensions. Not only are they easy to implement (usually a count of elements satisfying some conditions) but flexible in terms of fitness to the situation such as stress on precision, cautiousness in making decision, reliability, or even hunt for extremes.

The constructions presented above are independent of the form of R. They can be applied to *any* (irreflexive) relation $R \subseteq X \times X$. Hence, even if the definition of the ranking relation R would change, the results of *this section* would hold. This is important from the knowledge-based system viewpoint, since for some specific cases, the definitions of the ranking relations may vary from what we have proposed in this paper.

Under some circumstances it might be attractive to use a method that enables experts to express their priorities in a more refined way. The method of consistency-driven pairwise comparisons allows us to express local preferences and compute the global merit index which can be used for drawing the final conclusions. Consistency-driven pairwise comparisons method processes subjective assessments, therefore any final conclusion is also subjective but controlled by the consistency analysis.

Acknowledgements

This work is partly supported by the Natural Science and Engineering Research Council of Canada (NSERC) grant.

References

[1] R. Janicki, W.W. Koczkodaj, A weak order solution to a group ranking, Comput. Math. Appl. 32 (2) (1996) 51–59.
[2] K.J. Arrow, Social Choice and Individual Values, Wiley, New York, 1951.
[3] S. French, Decision Theory: An Introduction to The Mathematics of Rationality, Ellis Horwood, Chichester, UK, 1986.
[4] W.W. Koczkodaj, A new definition of consistency of pairwise components, Math. Comput. Modelling 18 (7) (1993) 79–84.
[5] G.T. Fechner, Elements of Psychophysics, vol. 1, Holt, Rinehart & Winston, New York, 1965, translation by H.E. Adler of Elemente der Psychophysik, Leipzig: Breitkopf und Härtel, 1860.

[6] L.L. Thurstone, A law of comparative judgments, Psycholog. Rev. 34 (1927) 273–286.
[7] T.L. Saaty, A scaling methods for priorities in hierarchical structure, J. Math. Psychology 15 (1977) 234–281.
[8] J.S. Dyer, Remarks on the analytic hierarchy process, Manag. Sci. 36 (3) (1990) 249–258.
[9] P.C. Fishburn, Interval Orders and Interval Graphs, Wiley, New York, 1985.
[10] M.W. Herman, W.W. Koczkodaj, Monte Carlo study of pairwise comparisons, Inf. Process. Lett. 57 (1) (1996) 25–29.
[11] W. Holsztynski, W.W. Koczkodaj, Convergence analysis of inconsistency algorithms for the pairwise comparisons method, Inf. Process. Lett. 59 (4) (1996) 197–202.