A Mixed Model for Pairwise Comparisons and Its Applications

A MIXED MODEL FOR PAIRWISE COMPARISONS AND ITS APPLICATIONS

BY

ABEER RASHAD MIRDAD, M.Sc

A THESIS

SUBMITTED TO THE DEPARTMENT OF COMPUTING AND SOFTWARE

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

Master of Science (2015)	McMaster University
(Computer Science)	Hamilton, Ontario, Canada

TITLE:	A Mixed Model for Pairwise Comparisons and Its Applications
AUTHOR:	Abeer Rashad Mirdad B.Sc. Umm Al-Qura University, Makkha, Saudi Arabia
SUPERVISOR:	Dr. Ryszard Janicki
NUMBER OF PAGES:	xvi, 88

To my Parents, Husband, and kids

Abstract

The method of Pairwise Comparisons was first described by Ramon Llull in the end of XIII century [45] . At present, this method is identified with the controversial Saaty's Analytic Hierarchy Process [52] that was first proposed in 1977. The Analytic Hierarchy Process is a formal method to derive ranking orders from pairwise comparisons and it is used around the world in a wide variety of decision making, in fields such as education, industry, and government. However many researchers consider it as a flawed procedure that might produce arbitrary rankings [14].

It the last two decades alternative models, also based on pairwise comparisons paradigm, that appear to work better in many cases and have better theoretical fundamentals. One of them is 'qualitative pairwise comparisons based ranking', first proposed in [29] and later developed and refined in [26, 32], the other one is a quantitative pairwise comparisons ranking but based on the concept of 'distance-based consistency', proposed in [35], refined and used relatively often in the last decade (c.f. [8, 20, 36]).

In this thesis we will substantially refine '*mixed model*', roughly proposed in [28, 30], and then apply this model for several 'real world' cases. The '*mixed model*' is a systematic composition of the quantitative model of [35, 20, 36] and the qualitative model of [26, 32]. In this thesis we clarify and provide some formal foundations for scales and assignment of numerical values for qualitative factors and based on these we provide a formal process to

be followed.

Seven applications in such fields as software evaluation, software quality in use, quality in use in video games, software reuse, smart grid analysis, healthcare quality and quality analysis of medical devices are provided and analysed. They show the *mixed method* to be useful and appreciated by users.

Acknowledgements

First of all I want to take this opportunity to express my thanks to Dr. Ryszard Janicki for supporting me and guiding me throughout the completion of this thesis. Also, I would like to thank Mohammed Alqarni, Sahar Kokaly and Jason Jaskolka for their time and feedback. I would like to thank also everyone who gave me time helped me finish this thesis, their invaluable contributions have made this thesis possible. I would like to thank my parents for bearing me and encouraging me through tough times. Finally, I am thankful to my wonderful husband for his patience and support during my studies I also wish to thank many friends and relatives who were always there for me during the hard moments I had.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis goes under the Artificial Intelligence category and classification theory of computer science. More specifically, this thesis is about merging both qualitative and quantitative Pairwise Comparisons in one model which we call *Mixed Model* and it is about designing and conducting an experiment to prove that the Mixed Model is better than the standard Geometric Model. In this thesis, by applying the Mixed Model to a number of applications in different areas, we will be showing and describing how the Mixed Model works perfectly.

1.1 Motivation

The *Pairwise Comparisons* method is based on the observation that it is much easier to rank the importance of *two* objects than it is to rank the importance of *several* objects. This very old idea goes back to Ramon Llull in the end of XIII century [45]. Its modern version is due to the 1785 influential paper by Marquis de Condorcet [10], where he used this method in the election process where voters rank candidates based on their preference, and 1860 paper by Fechner [15]. However it was Thurstone in 1927 [44] and Saaty in 1977 [52] who provided mathematical foundations for this method to be effectively used in multicriteria decision making and analysis.

At present, this method is often identified with Saaty's *Analytic Hierarchy Process* (or AHP) [52, 53]. AHP is a formal method to derive ranking orders from pairwise comparisons. The value of paired comparisons can be obtained from actual measurements such as length, weight etc., or from one's subjective opinion such as preference.
On one hand, AHP has many respected practical applications. However, it is still considered by many researchers to be a flawed procedure that produces arbitrary rankings [14]. The rank reversal occurs quite often [14]. There are also many other problems [14, 35], for example, for the AHP method, even though the input numbers are given roughly, the results are treated very precisely. If for example, after AHP processing, the object A obtained 'weight' 20.01 and B obtained 'weight' 19.99, it is usually stated that *A is preferred to B*. However in most cases '*A* and *B*' *are indistinguishable* or *equivalent* seems to be more appropriate.

In the middle 1990's, Ryszard Janicki and Waldemar W. Koczkodaj, proposed and discussed pairwise comparisons based non-numerical solutions for ranking [29]. This model has significantly been revised, enhanced and modified by Ryszard Janicki and Yun Zhai in [24, 25, 26, 31, 32, 61]. The solutions did not use numbers at all, instead some partial orders of qualitative values are used. A non-numerical model of ranking based on pairwise comparisons was proposed using the concept of partial orders and emphasizing the importance of indifference and weak ordering. The model exploits known, but not often used results from partial order theory [17].

Recently, in [30] and [28], a 'mixed' approach to pairwise comparisons was suggested. It appears that two quantitative models and a qualitative model of pairwise comparisons could be merged together, resulting in a convenient procedure for finding *consistent* weights and provide *consistent* ranking.

In this thesis we will refine, analyze and apply the 'mixed' model of pairwise comparisons. The main objectives of this thesis can be described as follows:

- Discussing, refining and analyzing the properties of *Mixed Model*, and comparing with the properties of a classical quantitative approach (i.e. for example [53, 14, 35]) and a qualitative only approach of [26, 32].
- Applying Mixed Model in different fields to show both its strength and weakness.

1.2 Contributions

The contributions of this thesis are two fold.

- The first part of the contribution is to formalize and revise the 'mixed model' proposed in [30] and [28]. In particular, some foundations for using proper scale are provide and some formal justifications for assigning relevant numbers to appropriate qualifications are also given. This part ends with a relatively formal procedure based on our version of 'mixed model' for weight assignments when judgments are subjective. To justify this procedure we first had to analyze in detail both the qualitative model of [26, 32] and the quantitative model of [53] and [35].

- The second part is a collection of applications in different areas applied to them for instance, the 'mixed model' technique, source code quality, quality in use in software, quality in use in video game, reusing software, smart energy grid, quality in health-care in Canada, and quality systems in medical devices. This part was very labour consuming as it involved plenty of interaction with specialists in other domains.

All calculations were made by using open source software called *JConcluder* [60].
We think the second contribution shows that the mixed model is superior to both pure qualitative and pure quantitative approaches.

A conference version of the results presented in this thesis has been accepted and will be presented in July 2015 at ICAI'2015 in Las Vegas [49].

Chapter 2

Partial Orders

Since *ranking* or *preference* is usually defined as a *weakly ordered* relationship between a set of items such that, for any two items, the first is either "less preferred", "more preferred" or "indifferent" to the second one, we need to provide basics concepts of partial order theory. This chapter is mainly based on [17] and [26].

2.1 Partial, Total and Weak Orders

Let X be a finite set. A relation $\lhd \subset X \times X$ is a (sharp) partial order if it is *irreflexive* and *transitive*, i.e., if:

$$a \lhd b \Rightarrow \neg(b \lhd a) \text{ and } a \lhd b \lhd c \Rightarrow a \lhd c$$

for all a,b,c $\in$ X.

A pair $(X, \lhd)$ is called a *partially ordered set* or *poset*.

Let $(X, \lhd)$ be a partially ordered set. Two elements a and b of X, are said to be *incomparable*, written $a \sim_{\lhd} b$, if they satisfy

$$a \sim_{\lhd} b \Longleftrightarrow \neg(a \lhd b) \wedge \neg(b \lhd a).$$

Note that for any $a \in X$, $a \sim_{\lhd} a$.

We say that the two elements a and b of X are *equivalent* with respect to $\lhd$, written $a \approx_{\lhd} b$ [18], if they satisfy

$$a \approx_{\lhd} b \Longleftrightarrow \left(\left\{x \middle| a \lhd x\right\} = \left\{x \middle| b \lhd x\right\}\right) \wedge \left(\left\{x \middle| x \lhd a\right\} = \left\{x \middle| x \lhd b\right\}\right)$$

The relation $a \approx_{\lhd} b$ is an equivalence relation (i.e. it is reflexive, symmetric and transitive) and it is called *the equivalence with respect to* $\lhd$, since if $a \approx_{\lhd} b$, there is nothing in $\lhd$ that can distinquish between a and b [17].

Note also that for a poset $(X, \lhd)$, and $a, b \in X$ we have

- $a \approx_{\lhd} b \Rightarrow a \sim_{\lhd} b$
- $a \approx_{\lhd} b \Longleftrightarrow \left\{x \middle| x \sim_{\lhd} a\right\} = \left\{x \middle| x \sim_{\lhd} b\right\}$

The partial order set $(X, \lhd)$ is said to be

- *total* or *linear*, if $\sim_{\lhd}$ is the identity relation, i.e., for all $a,b \in X$

$$a \lhd b \vee b \lhd a \vee a = b$$

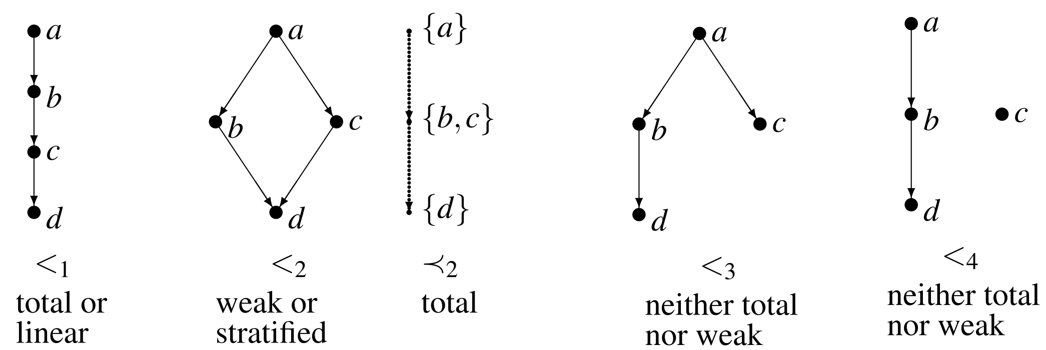


Figure 2.1: Various types of partial orders (represented as Hasse diagrams). The total order $\prec_2$ represents the weak order $<_2$., The partial orders $<_3$ and $<_4$ are neither total nor weak.

- *weak* or *stratified*, if $a \sim_\lhd b \sim_\lhd c \Rightarrow a \sim_\lhd c$, i.e. $\sim_\lhd$ is an equivalence relation. (see Figure 2.1).

Evidently, every total order is weak. Note also that if the order $\lhd$ is weak, then

$$a \approx_\lhd b \iff a \sim_\lhd b.$$

Weak orders are often defined in an alternative way (c.f. [17]):

- a partial order set $(X, \lhd)$ is a weak order iff there exists a total order $(Y, \prec)$ and a mapping $\phi : X \to Y$ such that $\forall x, y \in X, x \lhd y \iff \phi(x) \prec \phi(y)$.

This definition is illustrated in Figure 2.1, let $\phi : \{a,b,c,d\} \to \{\{a\},\{b,c\},\{d\}\}$ and $\phi(a) = \{a\}, \phi(b) = \phi(c) = \{b,c\}, \phi(d) = \{d\}$. Note that for all $x, y \in \{a,b,c,d\}$ we have: $x <_2 y \iff \phi(x) \prec_2 \phi(y)$.

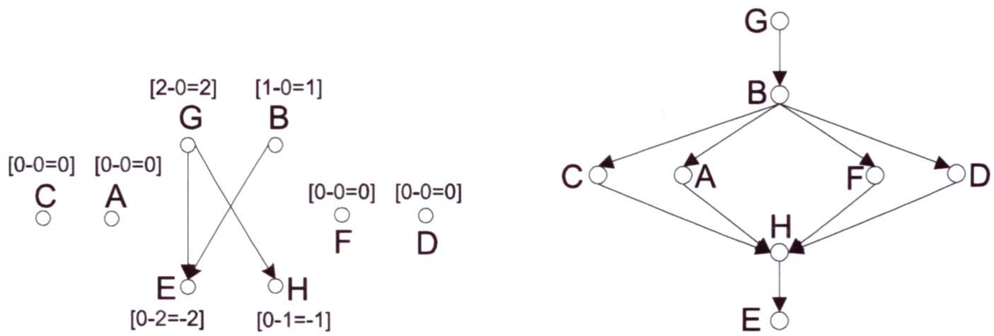
The preferable outcome of any ranking is a total order (c.f. [19]). Note that for any total order $\lhd$, both $\sim_\lhd$ and $\approx_\lhd$ reduce to the identity relation.

A total order has two natural models, both deeply embedded in the human perception of reality, namely: time and numbers, this is the main reason of its popularity.
Unfortunately in many cases it is not reasonable to insist that everything can or should be totally ordered. We may not have sufficient knowledge or such a perfect ranking may not even exist [1]. Quite often insisting on a totally ordered ranking results in an artificial and misleading "global index"

Weak (stratified) orders are a very natural generalization of total orders. They allow the modelling of some regular indifference, their interpretation is very simple and intuitive, and they are finally, although reluctantly, accepted by decision makers. Although not as much as one might expect given the huge theory of such orders [17].

If $\lhd$ is a weak order, then $a \approx_{\lhd} b \iff a \sim_{\lhd} b$, hence *equivalent* (w.r.t. $\lhd$) is equivalent to *incomparable*, and the relation $\lhd$ can be reinterpreted as a total order on the equivalence classes of $\sim_{\lhd}$ (or $\approx_{\lhd}$). For example, the partial order $<_2$ in Figure 2.1, we can interpret as a total order of equivalence classes:

$$\{a\}$$
$$\downarrow$$
$$\{b,c\}$$
$$\downarrow$$
$$\{d\}$$



(a) Partial order with global scores (b) Weak order extension of (a)

Figure 2.2: Computing global scores to construct weak order extension

2.2 Weak order approximations

For various reasons a ranking procedure may not immediately produce a weak order, just a partial one. What would be the ”best” weak order approximation? We will discuss it below.

Let X be a set and let $\lhd$ be a partial order relation on X. The relation $\lhd$ may or may not be a weak order. We are looking for the “best” weak order extension of $\lhd$. It appears that in this case the solution may not be unique.

Note that weak order extensions reflect the fact that if $x \approx_{\lhd} y$ then all reasonable methods for extending $\lhd$ will have x equivalent to y in the extension since there is nothing in the data that distinguishes between them [17], which leads to the definition below.

A weak (or total) order $\lhd_w \subseteq X \times X$ *is a proper weak (or total) order extension of* $\lhd$ if and only if:

$$(x \lhd y \implies x \lhd_w y) \text{ and } (x \sim_{\lhd_w} y \implies x \approx_{\lhd} y).$$

If X is finite, then every partial order $\lhd$ has a weak order extension. If $\lhd$ is weak, then its only proper weak order extension is $\lhd_w = \lhd$. If $\lhd$ is not weak, there are usually many such extensions [17]. For non-numerical ranking purposes, the best such extension seems to be the one found according to the concept of a *global score function*[17], defined by:

$$g_{\lhd}(x) = \left|\{z|z \lhd x\}\right| - \left|\{z|x \lhd z\}\right|$$

(where $|X|$ denotes the cardinality of a finite set X).

Given the global score function $g_{\lhd}(x)$, the relation $\lhd_w^g \subseteq X \times X$ is defined as

$$a \lhd_w^g b \Longleftrightarrow g_{\lhd}(a) < g_{\lhd}(b).$$

This is a proper weak extension of the partial order $\lhd$. Weak order approximations are discussed in detail in [18], and it was argued in [26] that the global score weak approximation is the most suitable for our purposes. An example of constructing a proper weak extension is shown in Figure 2.2.

Chapter 3

Classical Quantitative Pairwise Comparisons

In this chapter we will discuss the classical quantitative pairwise comparisons as proposed and described in for example [14, 53, 47, 35].

3.1 Pairwise Comparison Matrix and Consistency

Let $C_1, \ldots, C_n$ be a finite set of objects to be judged and/or analyzed. These objects are usually called *entities*, *criteria*, *alternatives*, *attributes*, etc., in this paper we will use the name *criteria*.

The relationship between features may be *qualitative* (*relational*) or *quantitative* (*numerical*). In this section we assume the relationship quantitative, and this quantitative relationship between entities C_i and C_j is represented by a positive number a_{ij}. We assume $a_{ij} > 0$ and $a_{ij} = \frac{1}{a_{ji}}$, for $i, j = 1, \ldots, n$ (which implies $a_{ii} = 1$ for all i).

If $a_{ij} > 1$ then C_i is more important (preferred, better, etc.) than C_j and a_{ij} is a measure of this relationship (the bigger a_{ij}, the bigger the difference), if $a_{ij} = 1$ then C_i and C_j are indifferent.

We also will call this model *multiplicative* since a_{ij} is interpreted as C_i *is* a_{ij} *times preferred* (*more important, etc.*) *than* C_j.

The matrix of such (multiplicative) relative comparison coefficients,

$$A = [a_{ij}]_{n\times n},$$

is called a (multiplicative) ***pairwise comparison matrix***.

Since the criteria $C_1, ..., C_n$ are not random, on contrary, they are usually carefully chosen and interrelated, the values of a_{ij} are not random, they should be somehow *consistent*.
A pairwise comparison matrix $A = [a_{ij}]_{n\times n}$ is *consistent* [52] if and only if

$$a_{ij}a_{jk} = a_{ik}, \tag{3.1}$$

for $i, j, k = 1, ..., n$. The equation (3.1) additionally justifies the name *multiplicative* as $a_{ij}a_{jk} = a_{ik}$ is interpreted as:

if C_j *is* a_{ij} *times preferred over* C_i *and* C_k *is* a_{jk} *times preferred over* C_j*, then* C_k *is* a_{ik} *times preferred over* C_i.

Saaty's Theorem [52] states that a pairwise comparison matrix A is consistent if and only if there exists positive numbers $w_1, ..., w_n$ such that $a_{ij} = w_i/w_j$, $i, j = 1, ..., n$. The values w_i are unique up to a multiplicative constant. They are often called *weights* and interpreted as a measure of importance. They are also often scaled to $w_1 + .. + w_n = 1$ (or 100%) and they obviously create 'natural' ranking.

For a consistent pairwise comparison matrix A, the values w_i create a ranking (i.e. a weak order):

$$x_i < x_j \iff w_i < w_j \text{ and } x_i \approx x_j \iff w_i = w_j.$$

There are *three problems here*. *First*, what kind of numerical values should be assign to a_{ij}? I am convinced that C_1 is slightly more important than C_3, but C_1 is significantly more important than C_4. What are the values of a_{13} and a_{14}? This is the problem of *scale*.
Second, even if we find some convincing way of transforming qualitative relationship between C_i and C_j into the numerical value of a_{ij}, it is very unlikely that the values a_{ij} will be consistent. Hence, *some measurement of inconsistency, and a guidance what level of inconsistency is accepted, are needed.*
Third, no matter how an inconsistency index is defined, the perfect consistency may be unrealistic or even not required, so how can we obtain weights $w_1, \ldots, w_n$ (which for consistent matrix satisfy $\frac{w_i}{w_j} = a_{ij}$, so they can easily be obtained starting from an arbitrary nonzero a_{12}) if the matrix is inconsistent, but within acceptable limit.

3.2 Saaty's Scale

In [52] and more formally in [53], Saaty has proposed nine different point intensity scales of importance between two compared elements. This scale is given below.

a_{ij}	Relationship between C_i and C_j
1	Equal Importance
3	Moderate Importance
5	Strong importance
7	Very strong or demonstrated importance
9	Extreme importance
2,4,6,8	For compromise between the above values
Reciprocals of above	If activity i has one of the above nonzero numbers assigned to it when compared with activity j, then j has the reciprocal value when compared with i

Saaty's scale, even though it has been often used, has also been widely criticized for example in [8, 14, 31, 35]. The main point of criticism was that it does not take into account natural limitations of the human mind as presented in [11, 48]. Nine point is just beyond the abilities of most humans and the more proper scale should have 4-7 points. Alternative scales, proposed in [35] and [31] will be discussed later in this thesis.

3.3 Saaty's Inconsistency Index

Since pairwise Comparison matrices of real life decision problems are rarely consistent, Saaty [52] has proposed an inconsistency index based on the value of the largest eigenvalue of A.

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RI_n	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49	1.51	1.48	1.56	1.57	1.59

Table 3.1: Comparative values

More precisely his method is based on computing the largest eigenvalue λ_{max} of the pairwise comparisons matrix $A = [a_{ij}]_{n \times n}$. In his seminal paper [52], he has shown that $\lambda_{max} \geq n$, $\lambda_{max} = n$ if and only if A is consistent. Based on this observation he defined *the inconsistency index*, denoted CI_n as:

$$CI_n = \frac{\lambda_{max} - n}{n - 1} \tag{3.2}$$

Notice that for $\lambda_{max} \geq n$, the value of CI_n is always positive. In Saaty's model *the inconsistency index CI_n on its own has no meaning, unless it is compared with some benchmark to determine the magnitude of the deviation from consistency.*

Let RI_n denote the average value of the *randomly obtained inconsistency indices* from a set of pairwise comparison matrices of size $n \times n$. Table 3.1 presents a sample of 15 such values, which is more than enough for most of the practical applications (c.f. [53]).

The inconsistency index CR_n of a given pairwise comparison matrix A is now defined by a ratio:

$$CR_n = \frac{CI_n}{RI_n} \tag{3.3}$$

If the matrix is consistent, then $\lambda_{max} = n$, so $CI_n = 0$ and $CR_n = 0$; if $RI_n \neq 0$, and we assume that $CR_n = 0$ if $CI_n = 0$ and $CR_n = 0$. In [52], Saaty concluded that an inconsistency ratio of about 10% or less could be considered acceptable. After that, Saaty has made some changes and improvement on his inconsistency ratio in 1990. Then, Saaty has proposed a threshold of 5% for 3×3, and 8% for 4×4 matrices. But all these thresholds are still controversial and debatable (c.f. [14, 35, 31]). As it was stated in [30]:

> '*Acceptable levels of inconsistency depend on the inconsistency index definition and particular interpretation of* $C_1, \ldots, C_n$. *It is highly doubtful that it can ever be "set in stone" as it represents our level of ignorance (or lack of the precise knowledge) and as such depends on a particular application*'.

However : The basic problem with Saaty's approach is that *it does not give any clue where most inconsistent values of A are located* [8, 35, 31]. Removing inconsistencies or lowering them to an acceptable level is a kind of art when the eigenvalue-based inconsistency index is used.

3.4 Koczkodaj's Scale and Inconsistency Index

In principle Koczkodaj's scale [35] is similar to that of Saaty's; however only five points are used. The scale is presented below.

a_{ij}	Relationship between C_i and C_j
1	Equal Importance
2	Moderate Importance
3	Strong importance
4	Very strong or demonstrated importance
5	Extreme importance
Reciprocals of above	If activity i has one of the above nonzero numbers assigned to it when compared with activity j, then j has the reciprocal value when compared with i

The scale is still linear but take into account natural limitations of the human mind as presented in [11, 48], so it is more trustworthy in practical applications.

In 1993, Koczkodaj [35] provided an inconsistency index based on the analysis of all triads a_{ij}, a_{jk} and a_{ik} from $A = [a_{ij}]_{n\times n}$. It is now called *distance-based inconsistency* and it is defined as follows:

$$cm_A = \max_{(i,j,k)} \left(\min \left(\left| 1 - \frac{a_{ij}}{a_{ik}a_{kj}} \right|, \left| 1 - \frac{a_{ik}a_{kj}}{a_{ij}} \right| \right) \right) \tag{3.4}$$

In this case the most inconsistent triad is localized, which is very helpful for the process of inconsistency reduction. In Equation (3.4), the outer $\max_{(i,j,k)}$ is self-explanatory, the inner min just follows from the fact that we want a measure of the inconsistency of the triple a_{ij}, a_{jk} and a_{ik} to be in the interval from zero to one.

While removing inconsistencies or lowering them to an acceptable level is a kind of art when the eigenvalue-based inconsistency index is used; when distance-based inconsistency index (Equation 3.4) is used, since the biggest 'troublemakers' are localized, we can improve consistency step by step, by small changes of values of the triple that results in the maximal inconsistency index. It was proved in [39] that this process converges. In most cases it converges pretty fast initially, and since there is no practical reason to continue decreasing the inconsistency indicator to zero (only the relatively high values of the inconsistency indicator are considered as unacceptable and harmful), a matrix with acceptable level of inconsistency can be found in a small number of steps by following a common sense heuristic. However in some cases it may take some time to get the matrix consistent (up to the inconsistency threshold) [36]. Recently a fast systematic algorithm for inconsistency reduction has been proposed in [38].

Since lowering the distance-base inconsistency can easily be algorithmized, it will be used in the rest of this thesis.

When distance based consistency index cm_A (equation 3.4) is used, an acceptable level of inconsistency is often set as 0.33 or $\frac{1}{3}$. It mainly follows from the use of triads, and it was justified [37] by a detailed analysis of 3×3 and 4×4 cases.

The paper [8] gives a very thorough analysis of Saaty's and Koczkodaj's approaches to the issue of consistency, and it appears the Koczkodaj measure, even though much less popular, has more merits and justification.

3.5 Weights and Inconsistent Pairwise Comparisons Matrices

There are two main approaches for deriving a suitable value w_i from an inconsistent, but with acceptable level of inconsistency, matrix A.

In the first, older one, first proposed by Saaty in [52], the weights $w_1, \ldots, w_n$ are defined as the principal eigenvector of the matrix A.

In the second one, proposed by Barzilai in [3], the weights $w_1, \ldots, w_n$ are defined as the geometric means of columns (or equivalently, rows) of the matrix A, i.e. for $i = 1, ..., n$,

$$w_i = \sqrt[n]{\prod_{j=1}^{n} a_{ij}} \tag{3.5}$$

The geometric means are used in this thesis. For small values of the inconsistency index, both methods produce very similar results [8].

3.6 Number of Criteria

The mathematics of pairwise comparisons method does not impose any limit on the number of attributes. However, for most of the practical applications the number of attributes (on one level) is relatively small, seldom bigger than ten, and it depends on particular domain of application [52, 35, 54]. The reasons are again limitations of human mind ([11, 48]) and the fact that finding a measure of relationship between distant attributes is almost always problematic [54].

Instead, for more complex cases, a hierarchical approach is used [52, 53].

Chapter 4

Non-numerical or Qualitative Pairwise Comparisons

When mostly subjective judgment is involved, providing quantitative relationship between two entities is usually difficult and almost always controversial. It is not easy to justify statements like 'C_i *is* 1.5 *better than* C_j', etc. It is much more convincing and trustworthy just to provide a *qualitative* assessment like 'C_i *is much better than* C_j' or 'C_i *is only slightly better than* C_j', etc.

A qualitative or nonnumerical pairwise comparison based ranking was proposed in 1996 by Janicki and Koczkodaj [29]. From 2007, this model has significantly been revised, enhanced and modified by Ryszard Janicki and Yun Zhai in [24, 25, 26, 31, 32]. The model is based on the concepts of a partial order and an approximation of an arbitrary binary relation by a partial order [27]. No numerical weights or quantitative mutual relationships are used.

Recall that a *ranking* is usually defined as a weakly order relationship between a set of items such that, for any two items a and b, the first is either "less preferred" (often denoted by $a < b$), "more preferred" ($a > b$), or "indifferent" ($a \approx b$) to the second one. However, when a pairwise comparisons approach is used, the initial result of judgments could be, in general, any relation. In particular the case: $a < b$, $b < c$ and $c < a$ might happen. Hence, some approximation techniques had to be developed.

In this chapter we will discuss basis of pairwise comparisons based qualitative ranking techniques proposed in [26] and [32].

4.1 Basic Model

The simplest and the most rough case is when for two different objects a and b we may either say a is preferred to b ($a \rightarrow b$ in a graph), or b is preferred to a ($a \leftarrow b$), or a and b are indifferent (no graphical connection)

The initial outcome of such pairwise comparisons process could be a relation as the one from Figure 4.1. Such relation, in general case, may not even be a partial order, so it cannot directly be interpreted as any ranking. Which leads us to the problem of *an approximation of an arbitrary relation by a partial order*.

We start with defining some specialized relations that will help us to define the problem and provide a formal solution (c.f. [25, 26, 27]). Let X be a set, $R \subseteq X \times X$ be a relation and let $id = \{(x,x) \mid x \in X\}$

- $R^+ = \bigcup_{i=1}^{\infty} R^i$, is the *transitive closure* of R.

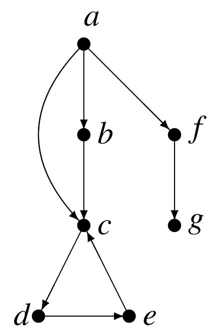


Figure 4.1: An example of a possible initial result of pairwise comparisons.

- $R^{\circ} = R \cup id$ is the *reflexive closure* of R.
- For $a \in X$, let:
  - $Ra = \{x \mid xRa\}$, and
  - $aR = \{x \mid aRx\}$.
- The relation $R^{\subset}$, an *inclusion property kernel*, is defined as:

$$aR^{\subset}b \iff bR^{\circ} \subset aR^{\circ} \wedge R^{\circ}a \subset R^{\circ}b.$$

- The relation R^{cyc} is defined as follows :

$$aR^{cyc}b \iff aR^{+}b \wedge bR^{+}a,$$

  If $aR^{cyc}b$ we will say that a and b belong to some cycle(s).

- The relation $R^{\bullet}$, an acyclic refinement of R, is defined as:

$$aR^{\bullet}b \iff aRb \wedge \neg(aR^{cyc}b).$$

- We define the relation $R^{\subset\wedge\bullet}$ as follows:

$$aR^{\subset\wedge\bullet}b \iff aR^{\subset}b \wedge aR^{\bullet}b.$$

- The relation $\equiv_R$ defined as:

$$a \equiv_R b \iff aR = bR \wedge Ra = Rb$$

  is the *equivalence relation with respect to R*, since if $a \equiv_R b$, there is nothing in R that can distinguish between a and b (the relation $\approx_{\lhd}$ from Chapter 2.1 is just $\equiv_{\lhd}$ where $\lhd$ is a partial order).

The following definition of a property driven partial order approximation of an arbitrary relation has been proposed in [25], analyzed in detail in [27], and applied to ranking procedures in [26].

Definition 1 ([25]). *A partial order $< \subseteq X \times X$ is a (property driven)* partial order approximation *of a relation $R \subseteq X \times X$ if it satisfies the following four conditions:*

1. $a < b \implies aR^{+}b$,
2. $a < b \implies \neg(aR^{cyc}b)$ *(or, equivalently* $a < b \implies \neg(bR^{+}a)$*),*
3. $aR^{\subset\wedge\bullet}b \implies a < b$,
4. $a \equiv_R b \implies a \approx_{<} b$. □

In principle it is a formal description of standard intuitions about approximation by a partial order.

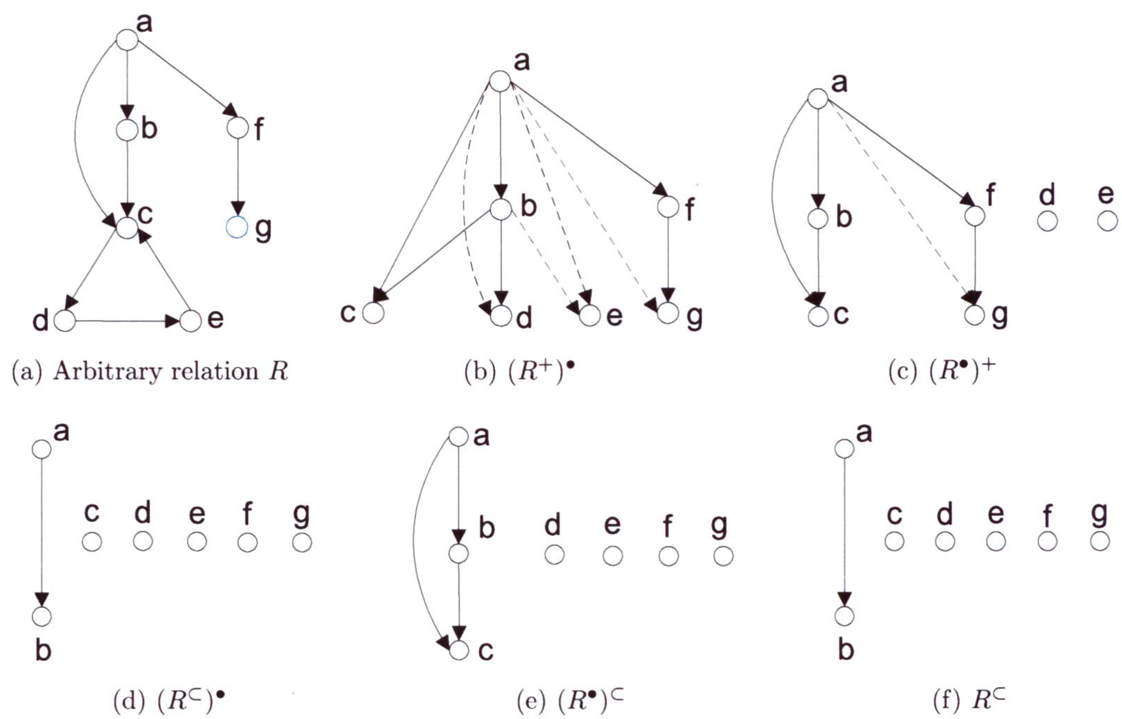


Figure 4.2: An example of arbitrary relation R and its partial order approximations

Theorem 2 ([25])**.** *$R^{\subset\wedge\bullet}$, $(R^{\bullet})^{\subset}$, $(R^{\bullet})^{+}$, $(R^{+})^{\bullet}$ are partial orders and partial order approximations of R, and $R^{\subset\wedge\bullet} \subseteq (R^{\bullet})^{\subset} \subseteq (R^{\bullet})^{+} \subseteq (R^{+})^{\bullet}$ and $R^{\subset\wedge\bullet} \subseteq (R^{\bullet})^{\subset} \subseteq R$.* □

Theorem 2 is illustrated in Figure 4.1.

The relation $(R^{+})^{\bullet}$ is a partial order approximation proposed implicitly by Shröder in 1895 [56]. It is often regarded as the only partial order approximation of R, which is not true, compare [25] and [27].

The operations form Theorem 2 do not guarantee weak ordering, only partial ordering. If the outcome is not a weak order, and weak ordering is required, we can use for example the global score function and the procedure described in Chapter 2.2.

For this basic model, a qualitative or non-numerical ranking procedure, consists of the following three steps:

1. Define some relation R that represent initial empirical ranking data of pairwise comparisons judgments.
2. Construct appropriate partial order approximation of R, dependently on the case you may choose one from $R^{\subset\wedge\bullet}$, $(R^{\bullet})^{\subset}$, $(R^{\bullet})^{+}$, $(R^{+})^{\bullet}$
3. Build appropriate weak order approximation of the relation from step 3.

This method is analyzed in detail in [26]. The initial 1996 model of [29] is in principle a basic model with $(R^{+})^{\bullet}$ as a partial order approximation of R.

4.2 Full Model

The full model, first proposed in 2007 in [24], and later refined and extended in [26, 32] is conceptually closer to the classical pairwise comparisons as described in Chapter 3.

Instead of numerical values a_{ij}, the binary relations $\approx, \sqsubset, \subset, <, \prec$ and $\sqsupset, \supset, >, \succ$ are used. The relations are interpreted as

- $a \approx b$: a and b are *indifferent*,
- $a \sqsubset b$: *slightly in favor* of b,
- $a \subset b$: *in favor* of b,

- $a < b$: b is *strongly better*,
- $a \prec b$: b is *extremely better*.

The number of relations has been limited to five because of the known restrictions of human mind when it comes to subjective judgments (magical numbers 4, see [11]; and 7, see [48]).

The relations $\approx, \sqsubset, \subset, <, \prec, \sqsupset, \supset, >, \succ$ are disjoint and cover the all cases, i.e. for every C_i, C_j we have $C_i \; R \; C_j$ where R is one from $\approx, \sqsubset, \subset, <, \prec, \sqsupset, \supset, >, \succ$. The relation $\approx$ is symmetric and includes identity. The relations $\sqsupset, \supset, >, \succ$ are inversions of appropriate $\sqsubset$, $\subset, <, \prec$.

Definition 3. *If no other properties are assumed, a tuple* $\mathscr{R}_d = (X, \approx, \sqsubset, \subset, <, \prec)$ *is called a* Pairwise Comparison Ranking Data. □

The ranking data $\mathscr{R}_d$ is created from experts reports, so no reasonable consistency in any sense is expected, for example the case like $C_i < C_j < C_k < C_i$ *might* happen.

We now define the relations $\widehat{\sqsubset}$, $\widehat{\subset}$, $\widehat{<}$, and $\widehat{\prec}$ as follows:

$$\widehat{\prec} \; = \; \prec \qquad \widehat{<} \; = \; \prec \cup <$$
$$\widehat{\subset} \; = \; \prec \cup < \cup \subset \qquad \widehat{\sqsubset} \; = \; \prec \cup < \cup \subset \cup \sqsubset$$

The relations $\widehat{\sqsubset}$, $\widehat{\subset}$, $\widehat{<}$, and $\widehat{\prec}$ are interpreted as *combined preferences*, i.e.

- $a \widehat{\sqsubset} b$: *at least slightly in favour* of b,
- $a \widehat{\subset} b$: *at least in favour* of b,
- $a \widehat{<} b$: *at least strongly in favour* of b, and

- $a\widehat{\prec}b$: *at least b is far superior than a.*

Definition 4 ([26])**.** *The tuple* $\mathscr{R}_s = (X, \approx, \sqsubset, \subset, <, \prec)$ *is a* Pairwise Comparison Ranking System *if the following two simple rules are satisfied:*

1. $\widehat{\sqsubset}$, $\widehat{\subset}$, $\widehat{<}$, $\widehat{\prec}$ *are partial orders (*$\widehat{\sqsubset}$ *is often required to be weak)*
2. $\approx = \sim_{\widehat{\sqsubset}}$, *i.e.* $\approx \cup \widehat{\sqsubset} \cup \widehat{\sqsubset}^{-1} = X \times X$. □

Note that it $\widehat{\sqsubset}$ is a weak order, then $\approx$ is an equivalence relation.

Not every ranking system $\mathscr{R}_s$ is automatically consistent. Intuitively consistency means that the relationships C_i vs C_j and C_j vs C_k influences the relationship C_i vs C_k. For quantitative multiplicative pairwise comparisons it is given by the formula $a_{ij} \cdot a_{jk} = a_{ik}$.

For the qualitative ranking system $\mathscr{R}_s$, the consistency is defined by a set of axioms it must satisfy, and the basic idea on which all those axioms are constructed is very simple:

composition of relations should be relatively continuous and must not change preferences in a drastic way.

Consider the following composition of preferences: $a \approx b \wedge b \sqsubset c$. What relationship between a and c is consistent? Intuitively, $a \approx c$ and $a \sqsubset c$ are for sure consistent, $a \subset c$ is debatable, while $a < c$ and $a \prec c$ are definitively inconsistent. This leads to the following definition:

Definition 5 ([32])**.** *The tuple* $\mathscr{R}_s = (X, \approx, \sqsubset, \subset, <, \prec)$ *is a* pairwise comparison consistent ranking system *if it satisfies the axioms for a pairwise comparison ranking data and the following rules (called* consistency rules*) are also satisfied:*

1. $(a \approx b \wedge b \approx c) \Rightarrow (a \approx c \vee a \sqsubset c \vee c \sqsubset a)$

2.1. $(a \approx b \wedge b \sqsubset c) \vee (a \sqsubset b \wedge b \approx c) \Rightarrow (a \approx c \vee a \sqsubset c \vee a \subset c)$

2.2. $(a \approx b \wedge b \subset c) \vee (a \subset b \wedge b \approx c) \Rightarrow (a \sqsubset c \vee a \subset c \vee a < c)$

2.3. $(a \approx b \wedge b < c) \vee (a < b \wedge b \approx c) \Rightarrow (a \subset c \vee a < c \vee a \prec c)$

3. $(a \approx b \wedge b \prec c) \vee (a \prec b \wedge b \approx c) \Rightarrow (a < c \vee a \prec c)$

4.1. $(a \sqsubset b \wedge b \sqsubset c) \Rightarrow (a \sqsubset c \vee a \subset c \vee a < c)$

4.2. $(a \subset b \wedge b \sqsubset c) \vee (a \sqsubset b \wedge b \subset c) \Rightarrow (a \subset c \vee a < c \vee a \prec c)$

4.3. $(a < b \wedge b \sqsubset c) \vee (a \sqsubset b \wedge b < c) \Rightarrow (a < c \vee a \prec c)$

5.1. $(a \sqsubset b \wedge b \prec c) \vee (a \prec b \wedge b \sqsubset c) \Rightarrow (a \prec c)$

5.2. $(a \subset b \wedge b \prec c) \vee (a \prec b \wedge b \subset c) \Rightarrow (a \prec c)$

5.3. $(a < b \wedge b \prec c) \vee (a \prec b \wedge b < c) \Rightarrow (a \prec c)$

5.4. $(a \prec b \wedge b \prec c) \Rightarrow (a \prec c)$

6.1. $(a \subset b \wedge b < c) \vee (a < b \wedge b \subset c) \Rightarrow (a \prec c)$

6.2. $(a < b \wedge b < c) \Rightarrow (a \prec c)$

7. $(a \subset b \wedge b \subset c) \Rightarrow (a < c \vee a \prec c)$

8.1. $(a \approx b \wedge b \sqsupset c) \vee (a \sqsupset b \wedge b \approx c) \Rightarrow (a \approx c \vee a \sqsupset c \vee a \supset c)$

8.2. $(a \approx b \wedge b \supset c) \vee (a \supset b \wedge b \approx c) \Rightarrow (a \sqsupset c \vee a \supset c \vee a > c)$

8.3. $(a \approx b \wedge b > c) \vee (a > b \wedge b \approx c) \Rightarrow (a \supset c \vee a > c \vee a \succ c)$

8.4. $(a \approx b \wedge b \succ c) \vee (a \succ b \wedge b \approx c) \Rightarrow (a > c \vee a \succ c)$

9.1. $(a \sqsupset b \wedge b \sqsupset c) \Rightarrow (a \sqsupset c \vee a \supset c \vee a > c)$

9.2. $(a \sqsupset b \wedge b \supset c) \vee (a \supset b \wedge b \sqsupset c) \Rightarrow (a \supset c \vee a > c \vee a \succ c)$

9.3. $(a \sqsupset b \wedge b > c) \vee (a > b \wedge b \sqsupset c) \Rightarrow (a > c \vee a \succ c)$

9.4. $(a \sqsupset b \wedge b \succ c) \vee (a \succ b \wedge b \sqsupset c) \Rightarrow (a \succ c)$

10.1. $(a \supset b \wedge b \supset c) \Rightarrow (a > c \vee a \succ c)$

10.2. $(a \supset b \wedge b > c) \Rightarrow (a \succ c)$

10.3. $(a \supset b \wedge b \succ c) \Rightarrow (a \succ c)$

11.1. $(a > b \wedge b > c) \Rightarrow (a \succ c)$

11.2. $(a > b \wedge b \succ c) \vee (a \succ b \wedge b > c) \Rightarrow (a \succ c)$

12. $(a \succ b \wedge b \succ c) \Rightarrow (a \succ c)$

13.1. $(a \sqsupset b \wedge b \sqsubset c) \vee (a \sqsubset b \wedge b \sqsupset c) \Rightarrow (a \approx c \vee a \sqsubset c \vee a \sqsupset c)$

13.2. $(a \sqsupset b \wedge b \subset c) \vee (a \subset b \wedge b \sqsupset c) \Rightarrow (a \approx c \vee a \sqsubset c \vee a \subset c)$

13.3. $(a \sqsupset b \wedge b < c) \vee (a < b \wedge b \sqsupset c) \Rightarrow (a \sqsubset c \vee a \subset c \vee a < c)$

13.4. $(a \sqsupset b \wedge b \prec c) \vee (a \prec b \wedge b \sqsupset c) \Rightarrow (a \subset c \vee a < c \vee a \prec c)$

14.1. $(a \supset b \wedge b \sqsubset c) \vee (a \sqsubset b \wedge b \supset c) \Rightarrow (a \approx c \vee a \sqsupset c \vee a \supset c)$

14.2. $(a \sqsupset b \wedge b \subset c) \vee (a \subset b \wedge b \sqsupset c) \Rightarrow (a \approx c \vee a \sqsubset c \vee a \sqsupset c)$

14.3. $(a \supset b \wedge b < c) \vee (a < b \wedge b \supset c) \Rightarrow (a \approx c \vee a \sqsubset c \vee a \subset c)$

14.4. $(a \supset b \wedge b \prec c) \vee (a \prec b \wedge b \supset c) \Rightarrow (a \sqsubset c \vee a \subset c \vee a < c \vee a \prec c)$

15.1. $(a > b \wedge b \sqsubset c) \vee (a \sqsubset b \wedge b > c) \Rightarrow (a \sqsupset c \vee a \supset c \vee a > c)$

15.2. $(a > b \wedge b \subset c) \vee (a \subset b \wedge b > c) \Rightarrow (a \approx c \vee a \sqsupset c \vee a \supset c)$

15.3. $(a > b \wedge b < c) \vee (a < b \wedge b > c) \Rightarrow (a \approx c \vee a \sqsubset c \vee a \sqsupset c)$

15.4. $(a > b \wedge b \prec c) \vee (a \prec b \wedge b > c) \Rightarrow$
$(a \approx c \vee a \sqsubset c \vee a \subset c \vee a < c \vee a \prec c)$

16.1. $(a \succ b \wedge b \sqsubset c) \vee (a \sqsubset b \wedge b \succ c) \Rightarrow (a \supset c \vee a > c \vee a \succ c)$

16.2. $(a \succ b \wedge b \subset c) \vee (a \subset b \wedge b \succ c) \Rightarrow (a \sqsupset c \vee a \supset c \vee a > c \vee a \succ c)$

16.3. $(a \succ b \wedge b < c) \vee (a < b \wedge b \succ c) \Rightarrow$
$(a \approx c \vee a \sqsupset c \vee a \supset c \vee a > c \vee a \succ c)$

16.4. $(a \succ b \wedge b \prec c) \vee (a \prec b \wedge b \succ c) \Rightarrow$
$(a \approx c \vee a \sqsupset c \vee a \sqsubset c \vee a \supset c \vee a \subset c \vee a > c \vee a < c \vee a \succ c \vee a \prec c)$ □

Note that it is not immediately obvious that a consistent ranking system is also a ranking system of Definition 4, however it was proved in [32] that the properties of Definition 5 imply the properties of Definition 4, so a consistent ranking system is also a ranking system.

Since the components of starting pairwise comparison data $\mathscr{R}_d = (X, \approx', \sqsubset', \subset', <', \prec')$ are seldom even partial orders, we need some credible algorithms that could transform $\mathscr{R}_d$

into an appropriate pairwise comparison ranking system or pairwise comparison consistent ranking system $\mathscr{R}_s = (X, \approx, \sqsubset, \subset, <, \prec)$.

The following two algorithms have been proposed in [26] and [32].

Algorithm 1 ([26])**.** *Every ranking data* $\mathscr{R}_d = (X, \approx', \sqsubset', \subset', <', \prec')$ *can be transformed into appropriate ranking system* $\mathscr{R}_s = (X, \approx, \sqsubset, \subset, <, \prec)$ *in* $O(|X|^3)$ *time. The algorithm always terminates.* □

Algorithm 2 ([32])**.** *Every ranking data* $\mathscr{R}_d = (X, \approx', \sqsubset', \subset', <', \prec')$ *can be transformed into appropriate consistent ranking system* $\mathscr{R}_s = (X, \approx, \sqsubset, \subset, <, \prec)$ *in* $O(|X|^3)$ *time. The algorithm may end with* $\approx = X \times X$*, but always terminates.* □

Algorithm 2 does not require input to be a ranking system, however it is recommended to used the above algorithm in the sequence Algorithm 1 followed by Algorithm 2. Algorithm 2 alone may, and usually does, result it unreasonably big relation $\approx$. The problem is discussed in details in [32] and [61].

If the relation $\widehat{\sqsubset}$ is not a weak order, and weak ordering is required, we can use for example the global score function and the procedure described in Chapter 2.2, before applying Algorithm 2. The details again can be found in [32].

For this full model, a qualitative or non-numerical ranking procedure, consists of the following four steps:

1. Define a tuple $\mathscr{R}_d$ that represent empirical ranking data.

2. Construct appropriate $\mathscr{R}_s$ using Algorithm 1 (which uses Theorem 2).

3. If necessary construct appropriate weak order approximation of the relation $\widehat{\sqsubset}$.

4. Apply Algorithm 2 to guarantee consistency.

4.3 Testing Qualitative rankings

Testing procedures that involve subjective judgments is always problematic. How can we test something that is subjective?

Testing means that there are some data and results that are known to be correct, and then a given technique is applied to the same data. The differences between the correct results and those obtained by using a given technique are used to judge the value of this technique. Hence testing models such as the one presented above is problematic since *it is not obvious what should be tested against*. What are the correct results for given data? If the object has measurable attributes and there is a precise algorithm to calculate the value, the whole problem disappears.

Nevertheless a proper tests for these kinds of ranking techniques have been designed and executed [61].

> *A blindfolded person compared the weights of stones. The person put one stone in his left hand and another in his right hand, and then decided which of the relations* $\approx, \sqsubset, \subset, <, or \prec$ *held. The experiment was repeated for the same set of stones by various people, and then again for different stones and different number of stones, and again for various subsets of* $\{\approx, \sqsubset, \subset, <, or \prec\}$.

Those experiments have most likely been carried out by prehistoric man. Our ancestors probably used this technique to decide which stone was better to kill an enemy or an animal.

> *The important fact is that in this experiment the stones can be weighted using a precise scale, so we have the precise results to test against.*

The results of those experiments (60 stones, 50 experiments, comparison subsets from 4 to 60 stones) and their partial analysis can be found in [61]. In principle they validate the techniques described in this Chapter.

Chapter 5

‘Mixed’ Model, Non-linear Scale and Additive Quantitative Pairwise Comparisons

This chapter is a part of original contributions of this thesis. It follows from two observations.

First, in reality in almost all cases the process starts with providing some *qualitative* judgments first. So, why not make this fact a more formal part of the process? From this observation a *‘Mixed’ Model* was derived. Moreover, transformation of qualitative relations into quantitative values is still the most controversial problem. We will argue that the *non-linear* scale proposed in [31] should rather be used instead any of these discussed in Chapters 3.2 and 3.4.

The second observation is the following: When applying pairwise comparisons to various problems it was noticed that experts often felt much more comfortable and more confident when they were asked to divide 100 quality points between entities C_i and C_j than to provide multiplicative relationship, i.e. ratio a_{ij} [28]. This lead to the concept of *Additive Quantitative Pairwise Comparisons*.

We will start with the latter concept first. Then we will discuss the scale problem, and at the end we will introduce the 'Mixed' Model.

5.1 Additive Quantitative Pairwise Comparisons

Dividing 100 quality points between the entities C_i and C_j works usually better when the difference of their importance/influence is small. Then judgments as 45 for C_i and 55 for C_j are probably more precise and easier to justify than when the concept of ratio a_{ij} is used.

Dividing of 100 between C_i and C_j means that we are replacing the multiplicative relationship $a_{ij}a_{ji} = 1$, with the *additive*[1] relationship $b^i_j + b^j_i = 1$.

In this approach, for every $i, j = 1, \ldots, n$, we model the mutual relationship between C_i and C_j by two numbers b^i_j and b^j_i, where:

- b^i_j measures the importance of C_i in comparison with C_j assuming that their total importance is 1.0 (or 100%), similarly
- b^j_i measures the importance of C_j in comparison with C_i.

[1]The standard relationship between multiplication and addition is often defined by the log function as $\log(ab) = \log(a) + \log(b)$, however if $a < 1$ then $\log(a) < 0$ and clearly $\log(a_{ij}) + \log(a_{ji}) \neq 1$, so this transformation does not work for our purposes.

Formally we assume that, for all $i, j = 1, \ldots, n$,

$$b^i_j \geq 0, b^j_i \geq 0 \text{ and } b^i_j + b^j_i = 1 \tag{5.6}$$

Clearly $b^i_i = 0.5$ (or 50%) for all $i = 1, \ldots, n$.

The matrix of such (additive) relative comparison coefficients,

$$B = [b^i_j]_{n \times n},$$

is called a (additive) pairwise comparison matrix.

When b^i_j is interpreted as the probability that judges would prefer the entity C_i over C_j, the equation (5.6) is exactly the same as in Bradley-Terry model [9], however at this point similarity stops. The Bradley-Terry model does not have a concept of consistency similar to the one considered in this paper.

We want a mapping $\phi : \langle 0, 1 \rangle \to \langle 0, \infty)$ such that $a + b = 1$ implies $\phi(a) \cdot \phi(b) = 1$. When a and b are interpreted as *two parts of one whole* (which equals 1.0, or 100%), then $\frac{a}{b}$ represents *ratio* between a and b, and $\frac{a}{1-a}$ represents ratio between a and its complement. Hence, the most natural mapping seems to be

$$\phi(a) = \frac{a}{1-a}$$

This mapping has many different applications (c.f. [31, 41]), and in our case leads to the following transformation of ‘additive’ model into ‘multiplicative’ model.

For all $i, j = 1, \ldots, n$:

$$a_{ij} = \frac{b^i_j}{1 - b^i_j} = \frac{b^i_j}{b^j_i} \tag{5.7}$$

From equation (5.7) we immediately get that for all $i, j = 1, \ldots, n$, we have:

$$b^i_j = \frac{a_{ij}}{a_{ij} + 1} \tag{5.8}$$

$$a_{ij} a_{ji} = 1 \iff b^i_j + b^j_i = 1 \tag{5.9}$$

We may now analyze and reduce inconsistency by using the formula from equation (3.4). If going back to the additive model is needed, for example due to better interpretation of the results obtained, we can do it by using the equation (5.8).

5.2 Linear and Non-linear Scales

The scales presented in Chapters 3.2 and 3.4 (from [52] and [35] respectively) have only intuitive, heuristic and experimental justifications.

The idea of qualitative-quantitative relationship is based on the assumption any transformation in either way should preserve consistency. The mutual relationship between quantitative and qualitative pairwise comparisons has been analyzed in [31], which provides the following result.

Proposition 6 ([31]). *If the matrix $[a_{ij}]_{n\times n}$ is consistent (w.r.t. $a_{ik}a_{kj} = a_{ij}$ as the consistency definition) and each a_{ij} is transformed into R_{ij} by using ranges from Column I of Table 5.1 (i.e. for example if $1.28 \le a_{ij} \le 1.94$ then $R_{ij} = \sqsupset$, etc.), then the resulting set of* *relations is consistent with respect to qualitative consistency of* [32]. □

The more immediately applicable result where qualitative consistency is a premise and appropriate values of a_{ij} that guarantee acceptable inconsistency index, are conclusions does not exist yet. Nevertheless Proposition 6 is still very useful. First, it can be used to check if the final values of a_{ij}'s, obtained after reducing inconsistency to some acceptable level, fit the intuitions and feelings that are modeled by the relations $\approx, \sqsubset, \subset, <, \prec$. Secondly, it provides some guidance, based on more than just intuition and experience, what values should be attached to qualitative judgments. Thirdly it provides useful and convincing suggestions on what kind of *scale* should be used. Finally, the random tests have shown that when the numbers from Column I or III of Table 5.1 are used, a qualitatively consistent ranking system is transformed into a pairwise comparisons matrix with relatively, often acceptable, consistency index.

Using a proper scale is crucial as they supply the first, probably most important, link between qualitative and quantitative classification. Table 5.1 provides scales that are used or discussed later in this thesis. All scales can be seen as five points scales as they assume five different qualifying predicates. The scales with larger sets of qualifying predicates (for example nine point scale of [53] discussed in Chapter 3.3), but as described by [35] larger sets of qualifying predicates lose meaning in the comparison process. Moreover, in case of subjective judgments, limitations of human mind needs to be taken into consideration. According to [11, 48], the length of the scale should be between four (in [11]) and seven alternatives (in [48]). Smaller scales are also supported by recent result [20], which demonstrates that use of smaller scales, rather than larger ones, has good mathematical

Table 5.1: Relationship between *additive*, *multiplicative* and *relational* scales. Formulas $a_{ij} = \frac{b^i_j}{1-b^i_j}$ and $b^i_j = \frac{a_{ij}}{a_{ij}+1}$ are used to calculate the relationships between Columns I and III and Columns II and V. If no other data is available, default values are recommended.

quantitative scales [31, 35, 53]							**qualitative (relational) scale**	
additive scale of [31]		**multiplicative scales**					relation symbols for R_{ij} [31, 26]	**definition of intensity or importance** (E_i **vs** E_j) [26, 31, 35, 53]
range of b^i_j		**from [31]** range of a_{ij}		defaults derived from Column II	defaults **from**			
range	defaults	range	defaults		[35]	[53]		
0.44-0.55	0.5	0.79-1.27	1.0	1.0	1	1	$E_i \approx E_j$	*indifferent/equal/unknown*
0.56-0.65	0.6	1.28-1.94	1.6	1.5	2	3	$E_i \sqsupset E_j$	*slightly in favour/weak imp.*
0.66-0.75	0.7	1.95-3.17	2.6	2.3	3	5	$E_i \supset E_j$	*in favour/moderate importance*
0.76-0.85	0.8	3.18-6.14	4.7	4.0	4	7	$E_i > E_j$	*strongly better/demonstrated imp.*
0.86-1.00	0.9	6.15-	7.0	9.0	5	9	$E_i \succ E_j$	*extremely better/absolute imp.*
Column I	Col. II	Col. III	Col. IV	Column V	Col. VI	Col. VII	Col. VIII	Column IX

foundations. Using scales with a large set of qualifiers makes particular values of a_{ij} untrustworthy [26, 35] and may occasionally be a cause of known problems associated with multiplicative pairwise comparisons (c.f. [14]).

The linear multiplicative scale in Column VII of Table 5.1 is the scaled down to five points the oldest nine point scale proposed in [52, 53] and discussed in Chapter 3.3. The linear multiplicative scale in Column VI is the scale discussed in Chapter 3.4 and proposed in [35] Both scales have only intuitive, heuristic and experimental justifications. The fact that theses scale are in fact additive but is used in multiplicative pairwise comparisons is a little bit suspicious and may often lead to unnecessary big initial inconsistency.

The practically linear *additive* ranges in Column I were proposed in [31] in order to prove Proposition 6. This is not the only one scale that satisfies Proposition 6 but it is most likely the most intuitive one.

The ranges in Column III were derived from Column I by using the formula $a_{ij} = \frac{b^i_j}{1-b^i_j}$ (equation 5.7), however they can also be used directly to represent the relations from Column VII.

The default values in first four rows of Columns II and IV are middle points of appropriate ranges in Columns I and III, while the default values in the fifth row of Columns II and IV are educated guesses. The default values in Column V are derived from the values in Column II by using the formula $a_{ij} = \frac{b^i_j}{1-b^i_j}$. The default values are to be used when no other data or information or 'feelings' are available.

5.3 'Mixed' Model

The results from previous chapters and sections suggest the following weights assignment or ranking procedure.

Procedure 1.

1. Experts provide *qualitative* judgments using relations from Columns VIII and IX of Table 5.1.

2. Experts transform their qualitative judgments into *quantitative* judgments either using *additive* scale (Columns I and II), or *multiplicative* scale (Columns IV, V and VII) from Table 5.1, or *both* scales. It is recommended to use the additive scale for

the relations $\approx$ and $\sqsupset$. Sometimes the multiplicative scale (from Columns III and IV) works better for the relation $\succ$.

3. If the additive scale has been used, the values b^i_j are transformed into a_{ij} using the equation (5.7).

4. A standard procedure for inconsistency reduction is used (preferably with distance based consistency, see equation (3.4)).

5. The outcome, which is a multiplicative pairwise comparison matrix with acceptable inconsistency, is transformed back into *qualitative* matrix using ranges from Column III of Table 2, i.e. a_{ij}'s are transformed into appropriate R_{ik} from Column VIII.

6. In some cases the outcome is also transformed into additive pairwise comparison matrix by using the formula $b^i_j = \frac{a_{ij}}{a_{ij}+1}$ (equation (5.8)).

7. All final qualitative and quantitative tables are sent back to experts for final adjustments and potential changes.

8. The whole process is repeated as many times as necessary.

9. After the results are accepted, the weights are calculated as the geometric means of columns of the final matrix $A = [a_{ij}]_{n\times n}$, i.e. $w_i = \sqrt[n]{\prod_{j=1}^{n} a_{ij}}$, for $i = 1, \ldots, n$.

10. If ranking is required, it is derived from the weights in a standard manner.

The process described above is called a *'Mixed' Model*.

Chapter 6

Applications of 'Mixed' Model

Procedure 1 from the previous chapter is fairly general, it may be different concrete instantiations. The following specialized version of Procedure 1 will be used in all applications in this chapter.

Weights Assignment Process

Procedure 2.

1. A small group of experts were given a number of attributes and they were asked to provide *qualitative* values using relations from Table 5.1.

2. Then the same experts were asked to provide *quantitative* evaluation using the *additive scale* and ranges that provided in Column I of Table 5.1. *In all cases the experts decided to use used default values from Column II.*

3. Next the pairwise comparisons matrix $[a_{ij}]_{n\times n}$ was produced by using the formula $a_{ij} = \frac{b_{ij}}{1-b_{ij}}$. The distance based consistency index cm_A (equations (5.6)) is calculated,

and, if necessary, reduced to the level smaller than 0.3. *JConcluder* software [60] was used to reduce inconsistency to an acceptable level.

4. After a few days the same experts were asked to transform their qualitative judgments into quantitative form, but now by using *multiplicative* scale and ranges from Columns III and IV of from Table 5.1. *Again in all cases the experts decided to use default values, in this case it was Column IV. JConcluder* [60] was again used to reduce inconsistency to an acceptable level. *We observed that it is easier for experts to differentiate between different, say, indifferences, when b_{ij} were used; than when a_{ij} were used.*

5. Next the matrices we have got from step number 3 and step number 4 went through the standard process of inconsistency reduction.

6. The matrices from steps (4) and (5) were then translated back into qualitative form using Column II of Table 5.1. Even though the quantitative matrices were slightly different, their qualitative representations were identical (*which can be interpreted as yet another validation of the distance based consistency*).

7. For all four applications the final result was different than the initial one. The final results were sent back to the experts for final analysis and acceptance. □

6.1 Application 1: Source Code Quality

The quality of source code has been discussed and analyzed by two experts from Computing and Software Engineering at McMaster University into six key attributes [23, 43]: the *readability* and the logical structuring of the code, *ease of maintenance* to fix errors and the ability for adding new features easily, *low complexity* with fewer functions to reduce number of the decisions that the code takes, *number of compilation* which largely controls of source code effective, *robust input validation* to ensure that a program operates correctly, and finally *low resource consumption* and its negative effects in prevent valid users from accessing the software as well as it is impact on the surrounding environment [2]. However the importance of the mentioned attributes is open for discussion due to the need of given attributes for clarity and prioritization of the content. [2]

Using pairwise comparisons based techniques can provide weights assignments and ranking that are more trustworthy. Procedure 2 has been used by the experts and the results of applying this procedure have presented from table 6.1 to table 6.7.

After analyzing and studying the obtained results we have found that the pairwise comparisons matrix, presented in Table 6.3, has the inconsistency coefficient $cm_A = 0.96$ which is considered unacceptable for the distance-based inconsistency [35]. After reducing the inconsistency index to $cm_A = 0.11$, as presented in Table 6.4 we found that the difference of appropriate weights from Tables 6.4 and 6.6 is rather small which could also be seen as validation of our method.

The difference between initial and final judgments were not severe, but the initial judgments were not consistent, while the final judgments were consistent (i.e. with acceptable level of inconsistency), hence they were much more trustworthy. The relation $>$ replaced $\succ$ for the relationship between Readability and Robust Input Validation, $>$ replaced $\succ$ for the relationship between Readability and Low Resource consumption, $\sqsupset$ replaced $\supset$ for the relationship between Ease of Maintenance and Low Complexity, $\subset$ replaced $\sqsubset$ for the relationship between Ease of Maintenance and Robust Input Validation, $\subset$ replaced $<$ for the relationship between Ease of Maintenance and Low Resource consumption, $\sqsubset$ replaced

$<$ for the relationship between Number of Compilation and Low Resource consumption, and $\sqsupset$ replaced $\supset$ for the relationship between Robust Input Validation and Low Resource consumption.

Our finding could help to develop initiatives to address specific quality problems in source code, and ultimately lead to better outcomes. The scaled weights (calculated to 100%) have the following values:

Readability =58% , Ease of Maintenance = 5.5% , Low Complexity= 4% , Number of Compilation= 9% , Robust Input Validation=11.5%, Low Resource Consumption=13.5%.

Note that *Readability* has scored the highest weight and it is about 60% of the total weights. This can be interpreted that the *Readability* plays a key role, especially in large systems. This results is not a surprise, but putting some numbers, that are justified and trustworthy because the outcome is consistent, and the relationship between numbers used for weights computation and appropriate qualitative factors was not 'ad hoc' but somehow justified, make the result more convincing and useful in decision making and evaluations. In contrast *Low Complexity* and *Ease of Maintenance* have scored as the two lowest weights, which was a some surprised results for the author.

The experts have approved our final results, both qualitative and quantitative. The final qualitative results and with a bottom row containing the final weights (calculated as averages from Tables 6.4 and 6.6) are presented in Table 6.7.

Table 6.1: **Source Code**. Initial *qualitative* judgments of experts about mutual relationship of key attributes.

Relations	Name	C_1	C_2	C_3	C_4	C_5	C_6
Readability	C_1	$\approx$	$\succ$	$\succ$	$\succ$	$\succ$	$\succ$
Ease of Maintenance	C_2	$\prec$	$\approx$	$\supset$	$\subset$	$\sqsubset$	$<$
Low Complexity	C_3	$\prec$	$\subset$	$\approx$	$<$	$\subset$	$\sqsupset$
Number of Compilation	C_4	$\prec$	$\supset$	$>$	$\approx$	$\approx$	$<$
Robust Input Validation	C_5	$\prec$	$\sqsupset$	$\supset$	$\approx$	$\approx$	$\supset$
Low Resource Consumption	C_6	$\prec$	$>$	$\sqsubset$	$>$	$\subset$	$\approx$

Table 6.2: **Source Code**. Initial *quantitative* judgments when experts using values of b_{ij} from Column I of Table 5.1 (*additive* model) to represent appropriate relations from Column VII. The experts decided to use the default values from Column II of Table 5.1.

b_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6
Readability	C_1	0.5	0.9	0.9	0.9	0.9	0.9
Ease of Maintenance	C_2	0.1	0.5	0.7	0.3	0.4	0.2
Low Complexity	C_3	0.1	0.3	0.5	0.2	0.3	0.6
Number of Compilation	C_4	0.1	0.7	0.8	0.5	0.5	0.2
Robust Input Validation	C_5	0.1	0.6	0.7	0.5	0.5	0.7
Low Resource Consumption	C_6	0.1	0.8	0.4	0.8	0.3	0.5

Table 6.3: **Source Code**. The values of a_{ij} obtained from Table 6.2.

$a_{ij} = \frac{b_{ij}}{1-b_{ij}}$	Name	C_1	C_2	C_3	C_4	C_5	C_6
Readability	C_1	1.0	9.0	9.0	9.0	9.0	9.0
Ease of Maintenance	C_2	0.1	1.0	2.3	0.43	0.67	0.25
Low Complexity	C_3	0.1	0.43	1.0	0.25	0.43	1.5
Number of Compilation	C_4	0.1	2.3	4.0	1.0	1.0	0.25
Robust Input Validation	C_5	0.1	1.5	2.3	1.0	1.0	2.3
Low Resource Consumption	C_6	0.1	4.0	0.67	4.0	0.43	1.0
inconsistency coefficient $cm_A = 0.96 > 0.3$.							

Table 6.4: **Source Code**. Consistent (i.e. with acceptable inconsistency) pairwise comparisons matrix derived from Table 6.3, using distance-based consistency. The weights were calculated using the geometric means.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6
Readability	C_1	1.0	11.10	13.02	6.4	5.8	4.5
Ease of Maintenance	C_2	0.1	1.0	1.32	0.57	0.50	0.43
Low Complexity	C_3	0.1	0.75	1.0	0.46	0.40	0.31
Number of Compilation	C_4	0.15	1.8	2.2	1.0	0.82	0.69
Robust Input Validation	C_5	0.17	2.0	2.5	1.2	1.0	0.81
Low Resource Consumption	C_6	0.22	2.3	3.2	1.4	1.2	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5	w_6
values		**58%**	**5%**	**4%**	**9%**	**12%**	**14%**
inconsistency coefficient $cm_A = 0.11 < 0.3$.							

Table 6.5: **Source Code**. Initial *quantitative* judgments of the same experts when using the default values from Column IV of Table 5.1 (multiplicative model), to represent appropriate relations from Column VII of Table 5.1.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6
Readability	C_1	1.0	7.0	7.0	7.0	7.0	7.0
Ease of Maintenance	C_2	0.14	1.0	2.6	0.38	0.63	0.21
Low Complexity	C_3	0.14	0.38	1.0	0.21	0.38	1.6
Number of Compilation	C_4	0.14	2.6	4.7	1.0	1.0	0.21
Robust Input Validation	C_5	0.14	1.6	2.6	1.0	1.0	2.6
Low Resource Consumption	C_6	0.14	4.7	0.63	4.7	0.38	1.0
inconsistency coefficient $cm_A = 0.97 < 0.3$.							

Table 6.6: **Source Code**. Consistent (i.e. with acceptable inconsistency) pairwise comparisons matrix derived from Table 6.5, using distance-based consistency. The weights were calculated using the geometric means. This table is almost identical as Table 6.4. The different cells are shaded.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6
Readability	C_1	1.0	11.10	13.02	6.4	5.8	4.5
Ease of Maintenance	C_2	0.1	1.0	1.32	0.57	0.50	0.43
Low Complexity	C_3	0.1	0.75	1.0	0.46	0.42	0.31
Number of Compilation	C_4	0.15	1.8	2.2	1.0	0.82	0.69
Robust Input Validation	C_5	0.17	2.0	2.4	1.2	1.0	0.81
Low Resource Consumption	C_6	0.22	2.3	3.2	1.4	1.2	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5	w_6
values		**58%**	**6%**	**4%**	**9%**	**11%**	**13%**
inconsistency coefficient $cm_A = 0.11 < 0.3$.							

Table 6.7: **Source Code**. Final *qualitative* judgments of key attributes derived from Table 6.4 by using the intervals from Column III of Table 5.1 for a_{ij} .Corrected cells are shaded.

Relations	Name	C_1	C_2	C_3	C_4	C_5	C_6
Readability	C_1	$\approx$	$\succ$	$\succ$	$\succ$	$>$	$>$
Ease of Maintenance	C_2	$\prec$	$\approx$	$\sqsupset$	$\sqsubset$	$\subset$	$\subset$
Low Complexity	C_3	$\prec$	$\sqsubset$	$\approx$	$\subset$	$\subset$	$<$
Number of Compilation	C_4	$\prec$	$\sqsupset$	$\supset$	$\approx$	$\approx$	$\sqsubset$
Robust Input Validation	C_5	$<$	$\supset$	$\supset$	$\approx$	$\approx$	$\sqsupset$
Low Resource Consumption	C_6	$<$	$\supset$	$>$	$\sqsupset$	$\sqsubset$	$\approx$
Final weights		**58%**	**5.5%**	**4%**	**9%**	**11.5%**	**13.5%**

6.2 Application 2: Software Quality in Use

SO/IEC 9126-1and ISO/IEC 14598-1 have defined *Quality in Use* as *the extent to which a product used by specified users meets their needs to achieve specified goals with effectiveness, productivity and satisfaction in a specified context of use*[4, 6]. Which means quality in use is the end user perspective of software quality, it is one of the three perspectives of software quality (internal quality, external quality, quality in use).

ISO (c.f. [23, 5]) has recently developed a new more comprehensive definition of *quality in use*, which has usability, flexibility and safety as sub-characteristics that can be quantified from the perspectives of different stakeholders, including users, managers and maintainers. Quality in use depends not only on the software or computer system, but also on the particular context in which the product is being used and it can be assessed by observing representative users carrying out representative tasks in a realistic context of use [50].

The standard ISO/IEC 25010 [23] proposes the following attributes (called 'characteristics') for the assessment of software product quality in use: *Effectiveness, Efficiency, Satisfaction, Freedom from Risk* and *Context Coverage* with which users can achieve goals in a specified context of use. In detail:

- *Effectiveness* means accuracy and completeness with which users achieve specified goals.
- *Efficiency* relates to resources used to obtain the accuracy and completeness.
- *Satisfaction* describes a degree that the user needs are satisfied when a product or system is used in a specified context of use.
- *Freedom From Risk* describes a degree that a product or system mitigates the potential risk to economic status, human life, health, or the environment.

- *Context coverage* describes a degree to which a product or system can be used in both specified contexts of use and in contexts beyond those initially explicitly identified.

To provide trustworthy assignment importance indicators, Procedure 2 has been used by three software experts from Computing and Software Department at McMaster University. The results are presented in Tables 6.8–6.14. Note that in this case the weights from Table 6.11 and Table 6.13 are identical.
The difference of appropriate weights from Table 6.11 and Table 6.13 is rather small and the experts have approved our final results, both qualitative and quantitative.
The difference between initial and final judgments were not severe, but the initial judgments were not consistent, while the final judgments were consistent (i.e. with acceptable level of inconsistency), hence they were much more trustworthy.

As in the previous case, the difference between initial and final judgments were not severe, but the initial judgments were not consistent, while the final judgments were consistent (i.e. with acceptable level of inconsistency), hence they were much more trustworthy.
The weights obtained by our analysis (see Table 6.14) have the following values (scaled to 100%):
Effectiveness =21%, Efficiency = 12%, Satisfaction = 26% , Freedom From Risk = 25%, Context Coverage = 16%.

The results appear in Table 6.14 the entities weights fall into two categories: Effectiveness, Satisfaction, and Freedom From Risk, weighted between 21% and 26%, while Efficiency, Context Coverage, weighted between 12% and 16%, roughly 10% smaller.

Table 6.8: **Software Quality In Use**. Initial *qualitative* judgments.

Relations	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	$\approx$	$\sqsupset$	$\sqsupset$	$\sqsubset$	$\sqsupset$
Efficiency	C_2	$\sqsubset$	$\approx$	$\subset$	$\sqsubset$	$\sqsubset$
Satisfaction	C_3	$\sqsubset$	$\supset$	$\approx$	$\supset$	$\supset$
Freedom From Risk	C_4	$\sqsupset$	$\sqsupset$	$\subset$	$\approx$	$\supset$
Context Coverage	C_5	$\sqsubset$	$\sqsupset$	$\subset$	$\subset$	$\approx$

Table 6.9: **Software Quality In Use**. Initial *quantitative* judgments derived from Table 6.8 by using defaults values from Column II of Table 5.1.

b_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	0.5	0.6	0.6	0.4	0.6
Efficiency	C_2	0.4	0.5	0.3	0.4	0.4
Satisfaction	C_3	0.4	0.7	0.5	0.7	0.7
Freedom From Risk	C_4	0.6	0.6	0.3	0.5	0.7
Context Coverage	C_5	0.4	0.6	0.3	0.3	0.5

Table 6.10: **Software Quality In Use**. The values of a_{ij} obtained from Table 6.9.

$a_{ij} = \frac{b_{ij}}{1-b_{ij}}$	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	1.0	1.5	1.5	0.67	1.5
Efficiency	C_2	0.67	1.0	0.43	0.67	0.64
Satisfaction	C_3	0.67	2.3	1.0	2.3	2.3
Freedom From Risk	C_4	1.5	1.5	0.43	1.0	2.3
Context Coverage	C_5	0.67	1.5	0.43	0.43	1.0
inconsistency coefficient $cm_A = 0.81 > 0.3$.						

Table 6.11: **Software Quality In Use**. Consistent pairwise comparisons matrix derived from Table 6.10.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	1.0	1.8	0.80	0.80	1.4
Efficiency	C_2	0.57	1.0	0.43	0.5	0.8
Satisfaction	C_3	1.15	2.3	1.0	1.03	1.64
Freedom From Risk	C_4	1.0	1.8	0.97	1.0	1.64
Context Coverage	C_5	0.67	1.14	0.43	0.6	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5
values		**21%**	**12%**	**26%**	**25%**	**16%**
inconsistency coefficient $cm_A = 0.12 < 0.3$.						

Table 6.12: **Software Quality In Use**. Initial *quantitative* judgments derived from Table 6.8 by using defaults values from Column IV of Table 5.1.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	1.0	1.6	1.6	0.63	1.6
Efficiency	C_2	0.63	1.0	0.38	0.63	0.63
Satisfaction	C_3	0.63	2.6	1.0	2.6	2.6
Freedom From Risk	C_4	1.6	1.6	0.38	1.0	2.6
Context Coverage	C_5	0.63	1.6	0.38	0.38	1.0
inconsistency coefficient $cm_A = 0.85 > 0.3$.						

Table 6.13: **Software Quality In Use**. Consistent pairwise comparisons matrix derived from Table 6.12. This table is almost identical as Table 6.11. The different cells are shaded.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	1.0	1.9	0.80	0.9	1.4
Efficiency	C_2	0.52	1.0	0.43	0.5	0.8
Satisfaction	C_3	1.15	2.3	1.0	1.03	1.64
Freedom From Risk	C_4	1.1	1.8	0.97	1.0	1.64
Context Coverage	C_5	0.67	1.14	0.43	0.6	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5
values		**21%**	**12%**	**26%**	**25%**	**16%**
inconsistency coefficient $cm_A = 0.12 < 0.3$.						

Table 6.14: **Software Quality In Use**. Final *qualitative* judgments of key attributes derived from Tables 6.11 and 6.13 by using the intervals from Column III of Table 5.1. Corrected cells are shaded. Also final weights.

Relations	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	$\approx$	$\sqsupset$	$\approx$	$\approx$	$\sqsupset$
Efficiency	C_2	$\sqsupset$	$\approx$	$\subset$	$\supset$	$\approx$
Satisfaction	C_3	$\approx$	$\supset$	$\approx$	$\approx$	$\sqsupset$
Freedom From Risk	C_4	$\approx$	$\subset$	$\approx$	$\approx$	$\sqsupset$
Context Coverage	C_5	$\sqsubset$	$\approx$	$\sqsubset$	$\sqsubset$	$\approx$
Final weights		**21%**	**12%**	**26%**	**25%**	**16%**

6.3 Application 3: Quality in Use in Video Games

Insuring the quality of software in general is usually a crucial part of its development and for video games, it is even more important. The quality in use in video games has been defined by researchers in a very specific objective. However, unlike traditional software, measuring quality in use in video games is rather subjective and personal.

Some researchers have highlighted and analyzed not only functional values of video games but also a set of specific non-functional value. It has been referred to by some researchers as *Quality in Use Model Based on Playability* [16].

These studies have analyzed and divided the *Quality in Use Model Based on Playability* into five main characteristics: *effectiveness* (i.e. completeness in how players can achieve the proposed goals); *efficiency* (i.e. determine the ease of learning and immersion); *safety* of the game (i.e. not putting the player's health at any risk); *flexibility* of video game and allowing its use in different contexts, by different players, or game profiles; and finally degree of *satisfaction* that players get from the content and context of use [55, 40].

Applying the pairwise comparisons based techniques can provide rankings and weights assignments that are more trustworthy.The initial judgment of criteria (step 1 of Procedure 2) has been assigned by two experts from Computing and Software Department at McMaster University .

After analyzing and studying the obtained results we have found that the pairwise comparisons matrix, presented in Table 6.17, has the inconsistency coefficient $cm_A = 0.75$ which is considered unacceptable for the distance-based inconsistency [35]. After reducing the inconsistency index to $cm_A = 0.13$ as presented in Table 6.18 we found that the difference of appropriate weights from tables 6.18 and 6.20 is rather small which could also be seen as validation of our method.

Although the differences between the initial and final judgments were not sharp, the final judgments are more responsible due to the acceptable level of consistency and this can be seen in Table 6.21.

Additionally, by looking at the attribute *Satisfaction* in Table 6.21, we see that it represents more than a half of the total weights. This can interprets the importance of this factor as it is the only factor that deals directly with players. Moreover, *Satisfaction* contains many sub-factors such as *fun, attractiveness, emotion, socialization, comfort and trust*. Given its importance, if producers concentrate more on *Satisfaction*, their overall success will be bigger. This is in contrast with *Safety* which represents the lowest priority for the players. This ,in our experts opinion, is due to players believe that their actions are their responsibilities themselves.

The experts have approved our final results, both qualitative and quantitative. The final qualitative results and its bottom row in tabe 6.21 contains final averages of the calculated weights (from tables 6.18 and 6.20).

The weights obtained by our analysis (Table 6.21) have the following values (scaled to 100%):

Effectiveness =22.5% , Efficiency = 8% , Safety = 7% , Flexibility = 8.5% , Satisfaction = 55% .

Table 6.15: **Quality in Use in Video Games**. Initial *qualitative* judgments of experts about mutual relationship of key attributes.

Relations	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	$\approx$	$\sqsupset$	$>$	$>$	$\sqsubset$
Efficiency	C_2	$\sqsubset$	$\approx$	$\sqsubset$	$\sqsupset$	$\prec$
Safety	C_3	$\supset$	$\sqsupset$	$\approx$	$\sqsubset$	$\prec$
Flexibility	C_4	$<$	$\sqsubset$	$\sqsupset$	$\approx$	$<$
Satisfaction	C_5	$\sqsupset$	$\succ$	$\succ$	$>$	$\approx$

Table 6.16: **Quality in Use in Video Games**. Initial *quantitative* judgments of the same experts when using values of b_{ij} from Column I of Table 5.1 (*additive* model) to represent appropriate relations from Column VII. The experts decided to use the default values from Column II of Table 5.1.

b_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	0.5	0.6	0.8	0.8	0.4
Efficiency	C_2	0.4	0.5	0.4	0.6	0.1
Safety	C_3	0.2	0.6	0.5	0.4	0.1
Flexibility	C_4	0.2	0.4	0.6	0.5	0.2
Satisfaction	C_5	0.6	0.9	0.9	0.8	0.5

Table 6.17: **Quality in Use in Video Games**. The values of a_{ij} obtained from Table 6.16.

$a_{ij} = \frac{b_{ij}}{1-b_{ij}}$	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	1.0	1.5	4.0	4.0	0.67
Efficiency	C_2	0.67	1.0	0.67	1.5	0.11
Safety	C_3	0.25	1.5	1.0	0.67	0.11
Flexibility	C_4	0.25	0.67	1.5	1.0	0.25
Satisfaction	C_5	1.5	9.0	9.0	4.0	1.0
inconsistency coefficient $cm_A = 0.75 > 0.3$.						

Table 6.18: **Quality in Use in Video Games**. Consistent (i.e. with acceptable inconsistency) pairwise comparisons matrix derived from Table 6.17, using distance-based consistency. The weights were calculated using the geometric means.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	1.0	2.68	3.16	2.90	0.40
Efficiency	C_2	0.37	1.0	1.14	0.99	0.14
Safety	C_3	0.32	0.9	1.0	0.80	0.12
Flexibility	C_4	0.34	1.0	1.25	1.0	0.14
Satisfaction	C_5	2.5	7.1	8.3	7.1	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5
values		**23%**	**8%**	**7%**	**8%**	**55%**
inconsistency coefficient $cm_A = 0.13 < 0.3$.						

Table 6.19: **Quality in Use in Video Games**. Initial *quantitative* judgments of the same experts when using the default values from Column IV of Table 5.1 (multiplicative model), to represent appropriate relations from Column VII of Table 5.1.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	1.0	1.6	4.7	4.7	0.63
Efficiency	C_2	0.63	1.0	0.63	1.6	0.14
Safety	C_3	0.21	1.6	1.0	0.63	0.14
Flexibility	C_4	0.21	0.63	1.6	1.0	0.21
Satisfaction	C_5	1.6	7.0	7.0	4.7	1.0
inconsistency coefficient $cm_A = 0.79 < 0.3$.						

Table 6.20: **Quality in Use in Video Games**. Consistent (i.e. with acceptable inconsistency) pairwise comparisons matrix derived from Table 6.19, using distance-based consistency. The weights were calculated using the geometric means. This table is almost identical as Table 6.18. The different cells are shaded.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	1.0	2.81	3.16	2.90	0.40
Efficiency	C_2	0.36	1.0	1.14	1.0	0.14
Safety	C_3	0.32	0.9	1.0	0.80	0.12
Flexibility	C_4	0.34	0.9	1.25	1.0	0.14
Satisfaction	C_5	2.5	7.1	8.3	7.1	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5
values		**22%**	**8%**	**7%**	**9%**	**55%**
inconsistency coefficient $cm_A = 0.13 < 0.3$.						

Table 6.21: **Quality in Use in Video Games**. Final *qualitative* judgments of key attributes derived from Table 6.18 by using the intervals from Column III of Table 5.1 for a_{ij} .Corrected cells are shaded.

Relations	Name	C_1	C_2	C_3	C_4	C_5
Effectiveness	C_1	$\approx$	$\supset$	$\supset$	$\supset$	$\subset$
Efficiency	C_2	$\subset$	$\approx$	$\sqsupset$	$\approx$	$\prec$
Safety	C_3	$\subset$	$\sqsubset$	$\approx$	$\approx$	$\prec$
Flexibility	C_4	$\subset$	$\approx$	$\approx$	$\approx$	$\prec$
Satisfaction	C_5	$\supset$	$\succ$	$\succ$	$\succ$	$\approx$
Final weights		**22.5%**	**8%**	**7%**	**8.5%**	**55%**

6.4 Application 4: Reusing Software

Software reuse is defined as the process of creating software systems from predefined software components [13]. Its importance has caused it to be extensively researched and studied and therefore there have been a number of studies that develop the most important attributes needed for it. Those are: *flexibility*(effort required modifying an operational program), *reusability* concerning the degree of ease of how an existing application or component can be reused, *scalability* of an application or component to be modified to expend its existing capacities, *robustness* and the ability of an algorithm to continue operating and to cope with errors during execution, and *usability* of the software that to be reused [34, 51].

Similarly to the previous applications, we have followed and used Procedure 2 and the results are presented in Table 6.22 to Table 6.28.

After analyzing and discussing the obtained results we have found that the pairwise comparisons matrix, presented in Table 6.24, has the inconsistency coefficient cm_A = 0.85 which is considered unacceptable for the distance-based inconsistency [35]. After reducing the inconsistency index to cm_A = 0.11 as presented in Table 6.25 we found that the difference of appropriate weights from tables 6.25 and 6.27 is rather small which could also be seen as validation of our method.

Notes that in this case the weights from Table 6.25 and Table 6.27 are identical. The difference of appropriate weights from Table 6.25 and Table 6.27 is rather small and the experts have approved our final results, both qualitative and quantitative.

With the opinions that our experts have, attribute *Reusability* has scored the highest weight with more than 55% of the total weights. This interprets that the *Reusability* (i.e the ease of extending software) plays a key role in reuse quality by having strategy for increasing productivity and improving quality in the software industry.

Although it is simple in concept, successful software reuse implementation is rather difficult in practice which require a code that is focused, composable components with a high cohesion and loose coupling. In addition, *Robustness* represents the second highest weight and this interprets how important it is for a reused software to continue operating and to handle unexpected conditions

.

The experts have approved our final results, both qualitative and quantitative. In Table 6.7 and its bottom row, the final qualitative results are given with the final averages calculated weights (from Tables 6.4 and 6.6).

The weights obtained by our analysis (Table 6.28) have the following values (scaled to 100%):

Flexibility =8% , Reusability = 56% , Scalability = 3% , Robustness = 27% , Usability = 6% .

Table 6.22: **Reusing Software**. Initial *qualitative* judgments of experts about mutual relationship of key attributes.

Relations	Name	C_1	C_2	C_3	C_4	C_5
Flexibility	C_1	$\approx$	$\prec$	$>$	$<$	$\supset$
Reusability	C_2	$\succ$	$\approx$	$\succ$	$\succ$	$\succ$
Scalability	C_3	$<$	$\prec$	$\approx$	$\subset$	$\approx$
Robustness	C_4	$>$	$\prec$	$\supset$	$\approx$	$\supset$
Usability	C_5	$\subset$	$\prec$	$\approx$	$\subset$	$\approx$

Table 6.23: **Reusing Software**. Initial *quantitative* judgments of the same experts when using values of b_{ij} from Column I of Table 5.1 (*additive* model) to represent appropriate relations from Column VII. The experts decided to use the default values from Column II of Table 5.1

b_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Flexibility	C_1	0.5	0.1	0.8	0.2	0.7
Reusability	C_2	0.9	0.5	0.9	0.9	0.9
Scalability	C_3	0.2	0.1	0.5	0.3	0.5
Robustness	C_4	0.8	0.1	0.7	0.5	0.7
Usability	C_5	0.3	0.1	0.5	0.3	0.5

Table 6.24: **Reusing Software**. The values of a_{ij} obtained from Table 6.23.

$a_{ij} = \frac{b_{ij}}{1-b_{ij}}$	Name	C_1	C_2	C_3	C_4	C_5
Flexibility	C_1	1.0	0.11	4.0	0.25	2.3
Reusability	C_2	9.0	1.0	9.0	9.0	9.0
Scalability	C_3	0.25	0.11	1.0	0.43	1.0
Robustness	C_4	4.0	0.11	2.3	1.0	2.3
Usability	C_5	0.43	0.11	1.0	0.43	1.0
inconsistency coefficient $cm_A = 0.85 > 0.3$.						

Table 6.25: **Reusing Software**. Consistent (i.e. with acceptable inconsistency) pairwise comparisons matrix derived from Table 6.24, using distance-based consistency. The weights were calculated using the geometric means.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Flexibility	C_1	1.0	0.14	2.97	0.30	1.24
Reusability	C_2	7.1	1.0	23.22	2.1	8.58
Scalability	C_3	0.34	0.04	1.0	0.09	0.40
Robustness	C_4	3.3	0.48	11.1	1.0	4.08
Usability	C_5	0.81	0.12	2.5	0.25	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5
values		**8%**	**56%**	**3%**	**27%**	**6%**
inconsistency coefficient $cm_A = 0.11 < 0.3$.						

Table 6.26: **Reusing Software**. Initial *quantitative* judgments of the same experts when using the default values from Column IV of Table 5.1 (multiplicative model), to represent appropriate relations from Column VII of Table 5.1.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Flexibility	C_1	1.0	0.14	4.7	0.21	2.6
Reusability	C_2	7.0	1.0	7.0	7.0	7.0
Scalability	C_3	0.21	0.14	1.0	0.38	1.0
Robustness	C_4	4.7	0.14	2.6	1.0	2.6
Usability	C_5	0.38	0.14	1.0	0.38	1.0
inconsistency coefficient $cm_A = 0.88 < 0.3$.						

Table 6.27: **Reusing Software**. Consistent (i.e. with acceptable inconsistency) pairwise comparisons matrix derived from Table 6.26, using distance-based consistency. The weights were calculated using the geometric means. This table is almost identical as Table 6.25. The different cells are shaded.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Flexibility	C_1	1.0	0.14	2.97	0.30	1.24
Reusability	C_2	7.1	1.0	23.0	2.1	8.58
Scalability	C_3	0.34	0.14	1.0	0.09	0.40
Robustness	C_4	3.3	0.48	11.1	1.0	3.95
Usability	C_5	0.81	0.12	2.5	0.38	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5
values		**8%**	**56%**	**3%**	**27%**	**6%**
inconsistency coefficient $cm_A = 0.11 < 0.3$.						

Table 6.28: **Reusing Software**. Final *qualitative* judgments of key attributes derived from Table 6.25 by using the intervals from Column III of Table 5.1 for a_{ij} .Corrected cells are shaded.

Relations	Name	C_1	C_2	C_3	C_4	C_5
Flexibility	C_1	$\approx$	$\approx$	$\supset$	$<$	$\approx$
Reusability	C_2	$\approx$	$\approx$	$\succ$	$\supset$	$\succ$
Scalability	C_3	$\subset$	$\prec$	$\approx$	$\approx$	$\approx$
Robustness	C_4	$>$	$\subset$	$\approx$	$\approx$	$>$
Usability	C_5	$\approx$	$\prec$	$\approx$	$<$	$\approx$
Final weights		**8%**	**56%**	**3%**	**27%**	**6%**

6.5 Application 5: Smart Grid

ISO's vision is to allow each network service to operate individually but share information with other services all through a base system. From this point they started to introduce features for smart grid such as: reliability, availability and efficiency etc. [7].

A smart grid is a modernized electrical grid that uses analog or digital information and communications technology to gather and act on information - such as information about the behaviours of suppliers and consumers - in an automated fashion to improve the efficiency, reliability, economics, and sustainability of the production and distribution of electricity [58].

The recommended assessment attributes are usually: *Reliability* to ensure more reliable supply of electricity, and reduced vulnerability to natural disasters or attack, *Flexibility* in network topology, *Efficiency* to overall improvement of the efficiency of energy infrastructure is anticipated from the deployment of smart grid technology, *Sustainability* significant challenges to power engineers who need to ensure stable power levels through varying the output of the more controllable generators such as gas turbines and hydroelectric generators and *Market-Enabling* to allow systematic communication between suppliers and consumers. These five attributes have given to three experts from Electrical Engineering at McMaster University to start their judgments. [21, 57]

Using the Procedure 2(proposed in the beginning of this chapter) has given us a significant results which have presented from Table 6.29 to Table 6.35. By analyzing the obtained results we could see in this case the weights from Tables 6.32 and 6.34 are slightly different and the pairwise comparisons matrix, presented in Table 6.31, has the inconsistency coefficient $cm_A = 0.75$ which is considered unacceptable for the distance-based inconsistency [35]. After reducing the inconsistency index to $cm_A = 0.12$ as presented in Table 6.32 we found that the difference of appropriate weights from tables 6.4 and 6.6 is rather small which could also be seen as validation of our method. The final weights presented in Table 6.35 are just averages of these from Tables 6.32 and 6.34. However, qualitative Tables 6.32 and 6.34 are identical and equal to Table 6.35. As in the previous cases, the experts approved final results of our analysis.

We can see in this case how *Efficiency* and *Market-Enabling* are exactly have identical relations and weights (see Table 6.35). This results inspired and interpreted of the importance of these two factors in *smart grid's* composition and structure since their weights represented more than 60% in our experts opinion.
Increasing *Efficiency* and conservation would be achieved several advantages including the effect is less redundancy in transmission and distribution lines, and greater utilization of generators, leading to lower power prices. In addition, the positive impact for the factor *Market-Enabling* can be interpreted on the following: demand response support, platform for advanced services and provision megabits, control power, sell the rest [7].

The scaled weights have the following values:
Reliability = 18%, Flexibility = 8%, Efficiency = 31.5%, Sustainability = 8%, Market-enabling = 31.5%.

Table 6.29: **Smart Grid** Initial *qualitative* judgments.

Relations	Name	C_1	C_2	C_3	C_4	C_5
Reliability	C_1	$\approx$	$\supset$	$\sqsupset$	$\sqsupset$	$\approx$
Flexibility	C_2	$\subset$	$\approx$	$<$	$\sqsubset$	$<$
Efficiency	C_3	$\sqsubset$	$>$	$\approx$	$>$	$\supset$
Sustainability	C_4	$\sqsubset$	$\sqsupset$	$<$	$\approx$	$<$
Market-enabling	C_5	$\approx$	$>$	$\sqsubset$	$>$	$\approx$

Table 6.30: **Smart Grid**. Initial *quantitative* judgments derived from Table 6.29 by using default values from Column II of Table 5.1.

b_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Reliability	C_1	0.5	0.7	0.6	0.6	0.5
Flexibility	C_2	0.3	0.5	0.2	0.4	0.2
Efficiency	C_3	0.4	0.8	0.5	0.8	0.7
Sustainability	C_4	0.4	0.6	0.2	0.5	0.2
Market-enabling	C_5	0.5	0.8	0.4	0.8	0.5

Table 6.31: **Smart Grid**. The values of a_{ij} obtained from Table 6.30.

$a_{ij} = \frac{b_{ij}}{1-b_{ij}}$	Name	C_1	C_2	C_3	C_4	C_5
Reliability	C_1	1.0	2.3	1.5	1.5	1.0
Flexibility	C_2	0.43	1.0	0.25	0.67	0.25
Efficiency	C_3	0.67	4.0	1.0	4.0	2.3
Sustainability	C_4	0.67	1.5	0.25	1.0	0.25
Market-enabling	C_5	1.0	4.0	0.67	4.0	1.0
inconsistency coefficient $cm_A = 0.75 > 0.3$.						

Table 6.32: **Smart Grid**. Consistent pairwise comparisons matrix derived from Table 6.31.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Reliability	C_1	1.0	2.3	0.57	2.3	0.57
Flexibility	C_2	0.43	1.0	0.25	1.0	0.22
Efficiency	C_3	1.7	4.0	1.0	4.0	1.0
Sustainability	C_4	0.43	1.0	0.25	1.0	0.25
Market-enabling	C_5	1.7	4.5	1.0	4.0	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5
values		**19%**	**8%**	**32%**	**8%**	**33%**
inconsistency coefficient $cm_A = 0.12 < 0.3$.						

Table 6.33: **Smart Grid**. Initial *quantitative* judgments derived from Table 6.29 by using the default values from Column IV of Table 5.1.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Reliability	C_1	1.0	2.6	1.6	1.6	1.0
Flexibility	C_2	0.38	1.0	0.21	0.63	0.21
Efficiency	C_3	0.63	4.7	1.0	4.7	2.6
Sustainability	C_4	0.63	1.6	0.21	1.0	0.21
Market-enabling	C_5	1.0	4.7	0.63	4.7	1.0
inconsistency coefficient $cm_A = 0.79 > 0.3$.						

Table 6.34: **Smart Grid**. Consistent pairwise comparisons matrix derived from Table 6.33. This table is almost identical as Table 6.32. The different cells are shaded.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5
Reliability	C_1	1.0	2.3	0.57	2.3	0.54
Flexibility	C_2	0.43	1.0	0.25	1.0	0.22
Efficiency	C_3	1.7	4.0	1.0	4.0	0.9
Sustainability	C_4	0.43	1.0	0.25	1.0	0.25
Market-enabling	C_5	1.8	4.5	1.1	4.0	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5
values		**18%**	**8%**	**31%**	**8%**	**30%**
inconsistency coefficient $cm_A = 0.12 < 0.3$.						

Table 6.35: **Smart Grid**. Final *qualitative* judgments of key attributes derived from Tables 6.32 and 6.34 by using the intervals from Column III of Table 5.1. Corrected cells are shaded. Also final weights in bottom row.

Relations	Name	C_1	C_2	C_3	C_4	C_5
Reliability	C_1	$\approx$	$\supset$	$\sqsubset$	$\supset$	$\sqsubset$
Flexibility	C_2	$\subset$	$\approx$	$<$	$\approx$	$<$
Efficiency	C_3	$\approx$	$>$	$\approx$	$>$	$\approx$
Sustainability	C_4	$\subset$	$\approx$	$<$	$\approx$	$<$
Market-enabling	C_5	$\sqsupset$	$>$	$\approx$	$>$	$\approx$
Final weights		**18%**	**8%**	**31.5%**	**8%**	**31.5%**

6.6 Application 6: Quality in Healthcare in Canada

The quality of Healthcare in Canada has been analyzed and assessed in *six* key domains [42, 59]: the *Effectiveness of* the healthcare sector in improving health outcomes; *Access* to healthcare services; the *Capacity* of systems to deliver appropriate services; the *Safety* of care delivered; the degree to which healthcare in Canada is *Patient-Centredness*; and *Equity* in healthcare outcomes and delivery [33]. However the importance of particular domains/attributes is far from obvious and open for a discussion. Using pairwise comparisons based techniques can provide ranking and weights assignments that are more trustworthy than these derived from informal or other semi-formal derivations.

Procedure 2 has been used by four experts from McMaster Children's Hospital and the results of applying this procedure have presented from table 6.36 to table 6.42.

After analyzing and studying the obtained results we have found that the pairwise comparisons matrix, presented in Table 6.38, has the inconsistency coefficient $cm_A = 0.57$ which is considered unacceptable for the distance-based inconsistency [35]. After reducing the inconsistency index to $cm_A = 0.15$ as presented in Table 6.39 we found that the difference of appropriate weights from tables 6.39 and 6.41 is rather small which could also be seen as validation of our method.
With different views and opinions that our experts have, *Effectiveness* has scored the highest weight and it is about 30% of the total weights. *Effectiveness* states the relationship between the level of resources invested and the level of results, or improvements in health. Assessing effectiveness consists of measuring the effects of medical practices and techniques therapeutic, diagnostic, surgical and pharmacological on individuals' health and

wellbeing., these are the reasons why *Effectiveness* was given the highest priority by our experts. In contrast *Equity* have scored as the lowest weight of quality in healthcare in Canada .

Table 6.42 contains the final qualitative results and its bottom row contains final calculated weights. The weights in Table 6.42 are just averages of the appropriate weights from Tables 6.39 and 6.41. Note that the difference of appropriate weights from Table 6.39 and Table 6.41 is rather small, which could also be seen as validation of our method. The experts have approved our final results, both qualitative and quantitative.

The difference between initial and final judgments were not severe, but the initial judgments were not consistent, while the final judgments were consistent (i.e. with acceptable level of inconsistency), hence they were much more trustworthy.

The weights obtained by our analysis (see Table 6.42) have the following values (scaled to 100%):

Effectiveness = 29.5%, Access = 12%, Capacity = 5.5%, Safety = 23.5%, Patient-Centredness = 24%, Equity = 5.5% .

Table 6.36: **Healthcare in Canada**. Initial *qualitative* judgments of experts.

Relations	Name	C_1	C_2	C_3	C_4	C_5	C_6
Effectiveness	C_1	$\approx$	$\supset$	$>$	$\approx$	$\approx$	$\supset$
Access	C_2	$\subset$	$\approx$	$\supset$	$\subset$	$<$	$\supset$
Capacity	C_3	$<$	$\subset$	$\approx$	$<$	$<$	$\approx$
Safety	C_4	$\approx$	$\supset$	$>$	$\approx$	$\approx$	$>$
Patient-Centredness	C_5	$\approx$	$>$	$>$	$\approx$	$\approx$	$>$
Equity	C_6	$\subset$	$\subset$	$\approx$	$<$	$<$	$\approx$

Table 6.37: **Healthcare in Canada**. Initial *quantitative* judgments derived from Table 6.36 by using defaults values from Column II of Table 5.1.

b_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6
Effectiveness	C_1	0.5	0.7	0.8	0.5	0.5	0.7
Access	C_2	0.3	0.5	0.7	0.3	0.2	0.7
Capacity	C_3	0.2	0.3	0.5	0.2	0.2	0.5
Safety	C_4	0.5	0.7	0.8	0.5	0.5	0.8
Patient-Centredness	C_5	0.5	0.8	0.8	0.5	0.5	0.8
Equity	C_6	0.3	0.3	0.5	0.2	0.2	0.5

Table 6.38: **Healthcare in Canada**. The values of a_{ij} obtained from Table 6.37.

$a_{ij} = \frac{b_{ij}}{1-b_{ij}}$	Name	C_1	C_2	C_3	C_4	C_5	C_6
Effectiveness	C_1	1.0	2.3	4.0	1.0	1.0	2.3
Access	C_2	0.43	1.0	2.3	0.43	0.25	2.3
Capacity	C_3	0.25	0.43	1.0	0.25	0.25	1.0
Safety	C_4	1.0	2.3	4.0	1.0	1.0	4.0
Patient-Centredness	C_5	1.0	4.0	4.0	1.0	1.0	4.0
Equity	C_6	0.43	0.43	1.0	0.25	0.25	1.0
inconsistency coefficient $cm_A = 0.57 > 0.3$.							

Table 6.39: **Healthcare in Canada**. Consistent (i.e. with acceptable inconsistency) matrix derived from Table 6.38, using distance-based consistency. The weights were calculated using the geometric means.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6
Effectiveness	C_1	1.0	2.3	5.2	1.3	1.1	4.6
Access	C_2	0.43	1.0	2.3	0.5	0.5	2.0
Capacity	C_3	0.2	0.43	1.0	0.25	0.25	1.0
Safety	C_4	0.77	1.8	4.0	1.0	1.0	4.0
Patient-Centredness	C_5	0.9	2.0	4.0	1.0	1.0	4.0
Equity	C_6	0.22	0.5	1.0	0.25	0.25	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5	w_6
values		**29%**	**12%**	**6%**	**23%**	**24%**	**6%**
inconsistency coefficient $cm_A = 0.15 < 0.3$.							

Table 6.40: **Healthcare in Canada**. Initial *quantitative* judgments derived from Table 6.36 by using defaults values from Column IV of Table 5.1.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6
Effectiveness	C_1	1.0	2.6	4.7	1.0	1.0	2.6
Access	C_2	0.38	1.0	2.6	0.38	0.21	2.6
Capacity	C_3	0.21	0.38	1.0	0.21	0.21	1.0
Safety	C_4	1.0	2.6	4.7	1.0	1.0	4.7
Patient-Centredness	C_5	1.0	4.7	4.7	1.0	1.0	4.7
Equity	C_6	0.38	0.38	1.0	0.21	0.21	1.0
inconsistency coefficient $cm_A = 0.62 > 0.3$.							

Table 6.41: **Healthcare in Canada**. Consistent pairwise comparisons matrix derived from Table 6.40. This table is almost identical as Table 6.39. The different cells are shaded.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6
Effectiveness	C_1	1.0	2.6	5.7	1.3	1.1	5.7
Access	C_2	0.38	1.0	2.6	0.5	0.5	2.3
Capacity	C_3	0.17	0.38	1	0.21	0.21	1.0
Safety	C_4	0.77	2	4.7	1.0	1.0	4.7
Patient-Centredness	C_5	0.9	2.0	4.7	1.0	1.0	4.7
Equity	C_6	0.17	0.43	1.0	0.21	0.21	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5	w_6
values		**30%**	**12%**	**5%**	**24%**	**24%**	**5%**
inconsistency coefficient $cm_A = 0.15 < 0.3$.							

Table 6.42: **Healthcare in Canada**. Final *qualitative* judgments of key attributes derived from Tables 6.39 and 6.41 by using the intervals from Column III of Table 5.1. Corrected cells are shaded. Also final weights.

Relations	Name	C_1	C_2	C_3	C_4	C_5	C_6
Effectiveness	C_1	$\approx$	$\supset$	$>$	$\sqsupset$	$\approx$	$>$
Access	C_2	$\subset$	$\approx$	$\supset$	$\subset$	$\subset$	$\supset$
Capacity	C_3	$<$	$\subset$	$\approx$	$<$	$<$	$\approx$
Safety	C_4	$\sqsubset$	$\supset$	$>$	$\approx$	$\approx$	$>$
Patient-Centr.	C_5	$\approx$	$\supset$	$>$	$\approx$	$\approx$	$>$
Equity	C_6	$<$	$\subset$	$\approx$	$<$	$<$	$\approx$
Final weights		**29.5%**	**12%**	**5.5%**	**23.5%**	**24%**	**5.5%**

6.7 Application 7: Quality System in Medical Devices

The primary goal in the design and manufacture of a medical device is to produce a quality product that meets the applicable requirements and specifications for its intended use. Such a product provides assurance that the medical device can be consistently manufactured and will perform as planned, safely and effectively.

This can be achieved by paying more attention and giving priority to some the characteristics and reviewing the ISO 13485 Medical Devices Quality management systems requirements for regulatory purposes [12].

Analysis the following characteristics of medical devices quality took us to get better results in the weights, so that would help experts and manufactures designing an efficient medical devices depending on the area to meet the needs of their operations and the needs of their customers. The ISO 13485 [12, 46] standard proposes the following attributes (called 'characteristics') for the assessment of Medical Devices Quality Managements Systems [22]:

- *Developing an effective*
- *Conducting risk analyses*
- *Following adequate and appropriate standard operating procedures and protocols for testing*
- *Using validated methods and procedures*
- *Monitoring and auditing*
- *Establishing and implementing a company quality policy*

- *Ensuring adequate staff training*
- *Conducting quality audits*
- *Implementing a quality control program*
- *Maintaining, analyzing and following up on complaint files*
- *Adopting and implementing appropriate corrective and preventive action plans*

Procedure 2 has been used by two experts from Computing and Software Department at McMaster University to provide appropriate weights for each of the above attributes..
The results are presented in Tables 6.43–6.49. While the Tables 6.46 and 6.48 are slightly different, the weights they generate are identical (as in the case of Software Quality In Use). Similarly as in the previous three cases, the qualitative Tables 6.46 and 6.48 are identical and equal to Table 6.49, and the experts approved final results of our analysis. The final results are consistent, so they are more trustworthy than the initial ones, that were based only on the experts opinions.

The scaled weights have the following values in this case:
Effective Design = 7%, Conducting Risk Analyses = 5%, Following Standard = 12%, Using Validated Methods = 29%, Monitoring and Auditing= 15%, Staff Training = 22%, Adopting and Implementing = 10%.

Table 6.43: **Medical Devices**. Initial *qualitative* judgments.

Relations	Name	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Effective Design	C_1	$\approx$	$\sqsupset$	$\subset$	$<$	$\subset$	$<$	$\sqsupset$
Conducting Risk Analyses	C_2	$\sqsubset$	$\approx$	$\subset$	$\prec$	$\subset$	$<$	$\sqsubset$
Following Standard	C_3	$\supset$	$\supset$	$\approx$	$\subset$	$\approx$	$\sqsubset$	$\sqsubset$
Using Validated Methods	C_4	$>$	$\succ$	$\supset$	$\approx$	$\sqsupset$	$\approx$	$>$
Monitoring and Auditing	C_5	$\supset$	$\supset$	$\approx$	$\sqsubset$	$\approx$	$\approx$	$\sqsupset$
Staff Training	C_6	$>$	$>$	$\sqsupset$	$\approx$	$\approx$	$\approx$	$\supset$
Adopting and Implementing.	C_7	$\sqsubset$	$\sqsupset$	$\sqsupset$	$<$	$\sqsubset$	$\subset$	$\approx$

Table 6.44: **Medical Devices**. Initial *quantitative* judgments derived from Table 6.43 by using the default values from Column II of Table 5.1.

b_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Effective Design	C_1	0.5	0.6	0.3	0.2	0.3	0.2	0.6
Conducting Risk Analyses	C_2	0.4	0.5	0.3	0.1	0.3	0.2	0.4
Following Standard	C_3	0.7	0.7	0.5	0.3	0.5	0.4	0.4
Using Validated Methods	C_4	0.8	0.9	0.7	0.5	0.6	0.5	0.8
Monitoring and Auditing	C_5	0.7	0.7	0.5	0.4	0.5	0.5	0.6
Staff Training	C_6	0.8	0.8	0.6	0.5	0.5	0.5	0.7
Adopting and Implementing.	C_7	0.4	0.6	0.6	0.2	0.4	0.3	0.5

Table 6.45: **Medical Devices**. The values of a_{ij} obtained from Table 6.43.

$a_{ij} = \frac{b_{ij}}{1-b_{ij}}$	Name	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Effective Design	C_1	1.0	1.5	0.43	0.25	0.43	0.25	1.5
Conducting Risk Analyses	C_2	0.67	1.0	0.43	0.11	0.43	0.25	0.67
Following Standard	C_3	2.3	2.3	1.0	0.43	1.0	0.67	0.67
Using Validated Methods	C_4	4.0	9.0	2.3	1.0	1.5	1.0	4.0
Monitoring and Auditing	C_5	2.3	2.3	1.0	0.67	1.0	1.0	1.5
Staff Training	C_6	4.0	4.0	1.5	1.0	1.0	1.0	2.3
Adopting and Implementing.	C_7	0.67	1.5	1.5	0.25	0.67	0.43	1.0
inconsistency coefficient $cm_A = 0.81 > 0.3$.								

Table 6.46: **Medical Devices**. Consistent pairwise comparisons matrix derived from Table 6.45.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Effective Design	C_1	1.0	1.5	0.54	0.25	0.43	0.3	0.67
Conducting Risk Analyses	C_2	0.7	1.0	0.43	0.18	0.32	0.21	0.52
Following Standard	C_3	1.8	2.3	1.0	0.43	0.82	0.54	1.1
Using Validated Methods	C_4	4.0	5.5	2.3	1.0	1.85	1.16	3.29
Monitoring and Auditing	C_5	2.3	3.1	1.1	0.5	1.0	0.76	1.5
Staff Training	C_6	3.3	4.7	1.8	0.8	1.3	1.0	2.3
Adopting and Implementing.	C_7	1.5	1.9	0.8	0.3	0.6	0.43	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5	w_6	w_7
values		**7%**	**5%**	**12%**	**29%**	**15%**	**22%**	**10%**
inconsistency coefficient $cm_A = 0.19 < 0.3$.								

Table 6.47: **Medical Devices**. Initial *quantitative* judgments derived from Table 6.43 by using the default values from Column IV of Table 5.1.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Effective Design	C_1	1.0	1.6	0.38	0.21	0.38	0.21	1.6
Conducting Risk Analyses	C_2	0.63	1.0	0.38	0.14	0.38	0.21	0.63
Following Standard	C_3	2.6	2.6	1.0	0.38	1.0	0.63	0.63
Using Validated Methods	C_4	4.7	7.0	2.6	1.0	1.6	1.0	4.7
Monitoring and Auditing	C_5	2.6	2.6	1.0	0.63	1.0	1.0	1.6
Staff Training	C_6	4.7	4.7	1.6	1.0	1.0	1.0	2.6
Adopting and Implementing.	C_7	0.63	1.6	1.6	0.21	0.63	0.38	1.0
inconsistency coefficient $cm_A = 0.99 > 0.3$.								

Table 6.48: **Medical Devices**. Consistent pairwise comparisons matrix derived from Table 6.47. This table is almost identical as Table 6.45. The different cells are shaded.

a_{ij}	Name	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Effective Design	C_1	1.0	1.15	0.54	0.25	0.43	0.3	0.80
Conducting Risk Analyses	C_2	0.8	1.0	0.43	0.18	0.32	0.21	0.52
Following Standard	C_3	1.8	2.3	1.0	0.43	0.82	0.54	1.1
Using Validated Methods	C_4	4.0	5.5	2.3	1.0	1.85	1.16	3.29
Monitoring and Auditing	C_5	2.3	3.1	1.1	0.5	1.0	0.76	1.5
Staff Training	C_6	3.3	4.7	1.8	0.8	1.3	1.0	2.3
Adopting and Implementing.	C_7	1.15	1.9	0.8	0.3	0.6	0.43	1.0
weights of criteria		w_1	w_2	w_3	w_4	w_5	w_6	w_7
values		**7%**	**5%**	**12%**	**29%**	**15%**	**22%**	**10%**
inconsistency coefficient $cm_A = 0.19 < 0.3$.								

Table 6.49: **Medical Devices**. Final *qualitative* judgments of key attributes derived from Tables 6.46 and 6.48 by using the intervals from Column III of Table 5.1. Corrected cells are shaded. Final weights are also presented.

Relations	Name	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Effective Design	C_1	$\approx$	$\approx$	$\sqsubset$	$<$	$\subset$	$<$	$\sqsupset$
Conducting Risk Analyses	C_2	$\approx$	$\approx$	$\subset$	$<$	$\subset$	$<$	$\sqsubset$
Following Standard	C_3	$\sqsupset$	$\supset$	$\approx$	$\subset$	$\approx$	$\sqsubset$	$\sqsubset$
Using Validated Methods	C_4	$>$	$>$	$\supset$	$\approx$	$\sqsupset$	$\approx$	$>$
Monitoring and Auditing	C_5	$\supset$	$\supset$	$\approx$	$\sqsubset$	$\approx$	$\sqsubset$	$\sqsupset$
Staff Training	C_6	$>$	$>$	$\sqsupset$	$\approx$	$\sqsupset$	$\approx$	$\supset$
Adopting and Implementing.	C_7	$\sqsubset$	$\sqsupset$	$\sqsupset$	$<$	$\sqsubset$	$\subset$	$\approx$
Final weights		**7%**	**5%**	**12%**	**29%**	**15%**	**22%**	**10%**

Chapter 7

Conclusion

7.1 Final comments

The idea and the work are an outcome of the analysis of the human mind. Humans are without doubt the most powerful learners in most aspects. Their ability to handle situations where they have never seen or been trained for, is one of their exclusive capabilities. Perhaps the human mind takes advantage of irrelevant learningsfrom other domains to find similarities with the current situation to handle it. This complementary information will help other to find more applications and study cases not only in software engineering area but also in different area in order to make it more accurate.

Overall, it appears that applying the *Mixed Model* methodology to the previous applications has improved weights assignment and has resulted for finding consistent weights and providing consistant ranking.

In this research, We have used three versions of pairwise comparisons, two of these are quantitative scales which are additive scale and multiplicative scale and the third version is a qualitative scale. This makes a one strong universal method that helps the weakest part of quantitative pairwise comparisons more systematic and efficient.

Two observations that highlighted and taken into account, represent a part of our contributions of this thesis:

1 Using the qualitative judgments scale at the early stage of the process and using the non linear scale proposed in [31] instead of any other scales Saaty's scale and Koczkodaj's Scale.

2 Taking into account that experts were feeling more comfortable and prefer dividing 100 points when they were asked to set the relationships between two entities C_i and C_j rather than providing multiplicative relationship

The second part was applying and implementing this methodology to a collection of applications and I have done eight applications in different fields like: software engineering area, electrical engineering, and medical domain. All these applications went throw the same procedures but they differ in processing time depending on the experts and their effectiveness and satisfaction on the initial results.

A few of these applications like Quality in Healthcare in Canada, and Quality System in Medical Devices Product needed to re-apply the proposed procedure for getting the best and acceptable results.

In some cases, working on only one application took too long due to the number of the attributes that the application has.

7.2 Future work

The most important problem in this thesis is how to assign appropriate numbers to the qualitative relations. According to Non-linear comparison scale proposed by Janicki and Zhai that presented in Table 5.1, The numbers were derived assuming that any transformation in either way should preserve consistency. From this point and as transformation of qualitative relations into quantitative values is still the most controversial problem, we suggest as a future work related to this thesis that Non-linear scales (as presented in Table 5.1) need more research. We also need more results similar to Proposition 6, but in other direction.

Bibliography

[1] K. Arrow. 1963. social choice and individual values. *Wiley, New York.*, 1951.

[2] R. Baggen, J. P. Correia, K. Schill, and J. Visser. Standardized code quality benchmarking for improving software maintainability. *Software Quality Journal*, 20(2):287–307, 2012.

[3] J. Barzilai. Deriving weights from pairwise comparison matrices. *Journal of the operational research society*, 48(12):1226–1232, 1997.

[4] N. Bevan. Quality in use: Meeting user needs for quality. *Journal of Systems and Software*, 49(1):89–96, 1999.

[5] N. Bevan. Extending quality in use to provide a framework for usability measurement. In *Human Centered Design*, pages 13–22. Springer, 2009.

[6] N. Bevan and M. Azuma. Quality in use: Incorporating human factors into the software engineering lifecycle. In *Software Engineering Standards Symposium and Forum, 1997. Emerging International Standards. ISESS 97., Third IEEE International*, pages 169–179. IEEE, 1997.

[7] D. Bhor and R. Mapari. Renewable energy based the smart grid technology.

[8] S. Bozóki and T. Rapcsák. On saaty's and koczkodaj's inconsistencies of pairwise comparison matrices. *Journal of Global Optimization*, 42(2):157–175, 2008.

[9] R. Bradley and M. Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, pages 324–345, 1952.

[10] M. Condorcet. Essai sur l'application de l'analyse a la probabilite des decisions rendue a la pluralite des voix. Paris, 1785.

[11] N. Cowan. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *BEHAVIORAL AND BRAIN SCIENCES*, 24:87–185, 2001.

[12] C. DeMarco. *Medical device design and regulation*. ASQ Quality Press, 2011.

[13] J. do Prado Leite, Y. Yu, L. Liu, S. Eric, and J. Mylopoulos. Quality-based software reuse. In *Advanced Information Systems Engineering*, pages 535–550. Springer, 2005.

[14] J. Dyer. Remarks on the analytic hierarchy process. *Management science*, 36(3):249–258, 1990.

[15] G. Fechner. *Elemente der Psychophysik*. Breitkopf und Härtel, 1860.

[16] A. Fernandez, E. Insfran, S. Abrahão, J. Carsí, and E. Montero. Integrating usability evaluation into model-driven video game development. In *Human-Centered Software Engineering*, pages 307–314. Springer, 2012.

[17] P. Fishburn. *Interval orders and interval graphs: A study of partially ordered sets*. Wiley New York, 1985.

[18] P. Fishburn and W. Gehrlein. A comparative analysis of methods for constructing weak orders from partial orders. *Journal of Mathematical Sociology*, 4(1):93–102, 1975.

[19] S. French. *Decision Theory*. Ellis Horwood, New York 1986.

[20] J. Fülöp, W. Koczkodaj, and S. Szarek. A different perspective on a scale for pairwise comparisons. *Transactions on computational collective intelligence I.(Lecture notes in computer science 6220.)*, pages 71–84, 2010.

[21] V. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. Hancke. Smart grid technologies: communication technologies and standards. *Industrial informatics, IEEE transactions on*, 7(4):529–539, 2011.

[22] Industry Canada. *Quality system requirements for medical devices*. http://publications.gc.ca/collections/collection2011.

[23] ISO/IEC 25022. System and software engineering – system and software quality requirements and evaluation (square) – system and software quality models. 2012.

[24] R. Janicki. Pairwise comparisons, incomparability and partial orders. In *ICEIS (2)*, pages 297–302, 2007.

[25] R. Janicki. Some remarks on approximations of arbitrary binary relations by partial orders. In *Rough Sets and Current Trends in Computing*, pages 81–91. Springer, 2008.

[26] R. Janicki. Pairwise comparisons based non-numerical ranking. *Fundamenta Informaticae*, 94(2):197–217, 2009.

[27] R. Janicki. Approximations of arbitrary binary relations by partial orders: classical and rough set models. In *Transactions on rough sets XIII*, pages 17–38. Springer, 2011.

[28] R. Janicki. On qualitative and quantitative pairwise comparisons. *Prof. of ICAT'2015*, to appear.

[29] R. Janicki and W. Koczkodaj. A weak order approach to group ranking. *Computers & Mathematics with Applications*, 32(2):51–59, 1996.

[30] R. Janicki and M. Soudkhah. On classification with pairwise comparisons, support vector machines and feature domain overlapping. *The Computer Journal*, 2014, doi:10.1093/comjnl/bxu085.

[31] R. Janicki and Y. Zhai. Remarks on pairwise comparison numerical and non-numerical rankings. In *Rough Sets and Knowledge Technology*, pages 290–300. Springer, 2011.

[32] R. Janicki and Y. Zhai. On a pairwise comparison-based consistent non-numerical ranking. *Logic Journal of IGPL*, 20(4):667–676, 2012.

[33] B. Kelowna. Measuring the performance of primary healthcare: existing capacity and potential information to support population-based analyses. 2009.

[34] K. Khosravi and Y. Guéhéneuc. On issues with software quality models. In *The Proceedings of the 11th Working Conference on Reverse Engineering. GEODES*, pages 172–181, 2004.

[35] W. Koczkodaj. A new definition of consistency of pairwise comparisons. *Mathematical and computer modelling*, 18(7):79–84, 1993.

[36] W. Koczkodaj. Classifying visual objects with the consistency-driven pairwise comparisons method. *Machine Graphics and Vision*, (3):143–153, 2009.

[37] W. Koczkodaj, M. Herman, and M. Orłowski. Using consistency-driven pairwise comparisons in knowledge-based systems. In *Proceedings of the sixth international conference on Information and knowledge management*, pages 91–96. ACM, 1997.

[38] W. Koczkodaj, M. Kosiek, J. Szybowski, and D. Xu. Fast convergence of distance-based inconsistency in pairwise comparisons. *Fundamenta Informaticae*, 2015.

[39] W. Koczkodaj and S. Szarek. On distance-based inconsistency reduction algorithms for pairwise comparisons. *Logic Journal of IGPL*, 18(6):859–869, 2010.

[40] M. Komal, Z. Khalil, and M. Sirshar. A comprehensive study on quality assurance in game development. 2014.

[41] F. Kong and H. Liu. Applying fuzzy analytic hierarchy process to evaluate success factors of e-commerce. *International Journal of Information and Systems Sciences*, 3–4(1):406–412, 2005.

[42] S. Leatherman, K. Sutherland, C. H. S. R. Foundation, et al. *Quality of healthcare in Canada: a chartbook*. Canadian Health Services Research Foundation, 2010.

[43] J. Letouzey and T. Coq. The sqale models for assessing the quality of real time source code. *Second International Conference on Advances in System Testing and Validation Lifecycle*, pages 43–48, 2010.

[44] T. L.L. The analytic hierarchy process. *Psychological Reviews*, (34):273–286, 1927.

[45] R. Llull. Artifitium electionis personarum, manuscript written between 1274-83. *Studia Lulliana*, (41):3–38, 2001.

[46] Medical devices. *Quality management systems. Requirements for regulatory purposes*, 2003 Standard. http://www.iso.org/iso/catalogue detail?csnumber=36786.

[47] B. Merkin. Group choice. *NY: John Wiley & Sons*, pages 69–88, 1979.

[48] G. Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[49] A. Mirdad and R. Janicki. Applications of mixed pairwise comparisons. *Prof. of ICAI'2015*, to appear.

[50] J. Oh, D. Park, B. Lee, J. Lee, E. Hong, and C. Wu. Certification of software packages using hierarchical classification. In *Software Engineering Research and Applications*, pages 209–224. Springer, 2004.

[51] M. Rothenberger, K. J. Dooley, U. R. Kulkarni, N. Nada, et al. Strategies for software reuse: A principal component analysis of reuse practices. *Software Engineering, IEEE Transactions on*, 29(9):825–837, 2003.

[52] T. Saaty. A scaling method for priorities in hierarchical structures. *Journal of mathematical psychology*, 15(3):234–281, 1977.

[53] T. Saaty. The analytic hierarchy process. *McGraw-Hill, New York.*, 1980.

[54] T. Saaty. *Theory and applications of the analytic network process: decision making with benefits, opportunities, costs, and risks*. RWS publications, 2005.

[55] J. Sánchez, F. Simarro, N. Zea, and F. Vela. Playability as extension of quality in use in video games. In *2nd International Workshop on the Interplay between Usability Evaluation and Software Development (I-USED)*, 2009.

[56] E. Schröder. *Vorlesungen über die Algebra der Logik (Exakte Logik). Dritter Band: Algebra und Logik der Relative*. B. G. Teubner, 1895.

[57] F. Sioshansi. *Smart grid: integrating renewable, distributed & efficient energy*. Academic Press, 2011.

[58] Smart Grids and Renewables: A Guide for Effective Deployment. International renewable energy agency (irena). 2013.

[59] J. Williams, W. Young, et al. A summary of studies on the quality of health care administrative databases in canada. *Patterns of health care in Ontario: the ICES practice atlas. 2nd ed. Ottawa: Canadian Medical Association*, 339:45, 1996.

[60] D. Xu. *Concluder*. http://sourceforge.net/projects/concluder/?source=directory.

[61] Y. Zhai. *Non-numerical ranking based on pairwise comparisons*. PhD thesis, McMaster University, 2010.