Chapter 6
A Decision Analytical Perspective on Public Procurement Processes

Mats Danielson, Love Ekenberg, Mattias Göthe, and Aron Larsson

Abstract If procurement processes are to be taken seriously, purchase managers need decision support tools beyond those that only ascertain that the formal requirements are met. This chapter demonstrates some fundamental flaws with common models used in procurement situations, flaws that are so serious that the evaluations of tenders often become meaningless and may lead to large and costly miscalculations. We demonstrate how the equitability of the tender evaluations can be significantly improved through the use of multi-criteria decision analysis with numerically imprecise input information. Due to this, the computational part of the evaluation step becomes more complex, and algorithms targeted for decision evaluation with imprecise data are used. We therefore present a procurement decision tool, *Decide*IT, implementing such algorithms that can be used as an instrument for a more meaningful procurement process. Of importance is to allow for a more realistic degree of precision in the valuation and ranking of tenders under each evaluation criterion, as well as the associated weighting of the criteria, since the criteria are often of a more qualitative nature. Through this, both quantitative and qualitative statements could be easily managed within the same framework

M. Danielson
Department of Computer and Systems Sciences, Stockholm University, Kista, Sweden

International Institute for Applied Systems Analysis, Laxenburg, Austria
e-mail: mats.danielson@su.se

L. Ekenberg (✉)
International Institute for Applied Systems Analysis, Laxenburg, Austria

Department of Computer and Systems Sciences, Stockholm University, Kista, Sweden
e-mail: ekenberg@iiasa.ac.at

M. Göthe
Department of Computer and Systems Sciences, Stockholm University, Kista, Sweden
e-mail: mattias@tanto-system.se

A. Larsson
Department of Computer and Systems Sciences, Stockholm University, Kista, Sweden

Department of Information and Communications Systems, Mid Sweden University, Sundsvall, Sweden
e-mail: aron@dsv.su.se

J. Papathanasiou et al. (eds.), *Real-World Decision Support Systems*,
Integrated Series in Information Systems 37,
DOI 10.1007/978-3-319-43916-7_6

and without the need to introduce ad-hoc and often arbitrary conversion formulas supposed to capture the trade-off between criteria.

6.1 Introduction

The values at stake in public sector procurement are very high, and therefore public procurement is an important issue for policy makers, suppliers and the general public. For instance, public procurement in Sweden alone has an annual turnover of between 16 and 18 % of the gross domestic product and similar figures are found in other of the countries within the OECD.[1] In the European Union (EU), this is in part a consequence of a pro forma comprehensive EU regulatory framework, with the double ambition of both increasing the efficiency and dynamics of the free European market, increasing transparency and predictability, and to avoid corruption. However, procurement processes as they are conducted in practice are often far from satisfactory from the perspectives of the buyers, the suppliers, and that of the general public.

As a consequence, in Sweden the confidence in the effectiveness of public procurement is relatively low within the suppliers,[2] and there have been some calls for dismantling of the entire legal framework for procurement in political circles.[3] It is commonplace that suppliers challenge the processes in costly and lengthy appeals, but also that the buyers are forced to start the entire process over when the effects of an accepted tender become apparent. These problems have of course a bearing on unnecessary financial costs, but also on the efficiency of policy implementation at large, and the public's confidence in the government in general.

The underlying rationale behind the current chapter is that we believe that the reasons for the shortcomings can, at least in part, be attributed to the methods employed to evaluate the bids when awarding a contract to a winning supplier since from a decision analytic perspective the procurement methods are unsatisfactory at best. The problems emanate to a large extent from three key issues, namely that those responsible for procurement processes often

(1) require unrealistic precision
(2) deal with qualitative values in an erroneous way, and
(3) manage value scales incorrectly.

Procurement in general is a large and growing field of research, but the vast majority of studies has been on private sector procurement, see e.g. [15]. Since the 1960s, the structured use of multiple criteria assessments when selecting a supplier has become more and more common, with a growing number of evaluation methods

[1] http://www.oecd.org/gov/ethics/public-procurement.htm.

[2] http://handelskammaren.net/sv/Nyheter/Nyhetsarkiv/Pressmeddelanden/2015/mars/lagt-fortroende-for-offentlig-upphandling-bland-vastsvenska-foretag/ (in Swedish).

[3] http://www.dagenssamhalle.se/debatt/skrota-lou-i-sjukvarden-134 (in Swedish).

proposed. An extensive review of methods can be found in [2]. In the public sector as compared to procurement in the private sector, there are special requirements and circumstances that make the process even more complex. One key difference is that in the public sector, both stakeholders and objectives are more diverse and possibly conflicting. Another difference is that public procurement is not merely a way to acquire goods or services, but also an important tool for policy makers, for example as an instrument for driving innovation from the demand-side or promoting environmental or social values [21].

However, the main formal difference between private and public sector procurement is that a buying entity in the public sector is typically regulated by a more extensive legal framework than a corresponding entity in the private sector. This paper deals to some extent with the limitations of the current European legislation on public sector procurement from a decision analysis perspective, and that subject has been the focus of studies like [15, 25], and also in [1], where a majority of 189 public procurement processes was found to be using deficient award mechanisms.

A procuring entity within the public sector "must provide the best value for money in public procurement while respecting the principles of transparency and competition".[4] Award decisions can be made using the lowest price criterion or a combination of qualitative and quantitative aspects (most economically advantageous tender—MEAT).

In the literature, the evaluation of suppliers in procurement is commonly referred to as "tender evaluation" or "supplier selection". A review of tender evaluation and selection methods suggested in the OR literature is provided in Ref. [2], however not restricting the scope to public procurement and its regulations and limitations. Most of the proposed methods published in the OR literature instead focus on supplier selection in for-profit manufacturing companies supply chain management, such as the reviews in [2, 16, 26]. When delimiting the context to EU public procurement, less studies have been made. Typically in such contexts, the buyer can base the selection based upon highest quality (aggregating the quality aspects), lowest price (of those qualifying with respect to quality), or an aggregation of price and quality searching for the economically most advantageous tender (MEAT).

There is a plethora of evaluation models available if several criteria are to be evaluated simultaneously. Some rely on a formal quantification of the quality dimension of the bid, such as using well-defined proxy measures, for example number of employees with a university degree or number of previous relevant contracts won. Other models define a minimum level of quality, and select a winner from the bidders passing the minimum level using price as the single criterion. Depending upon the procurement context and the information available, i.e. to what extent the cost of quality is known, different tender evaluation approaches are proposed. If the cost of quality is well known and many competitors can offer optimal quality, it is claimed that lowest price is the proper approach and not MEAT [1].

[4] http://ec.europa.eu/growth/single-market/public-procurement/rules-implementation.

In MEAT situations, there exists a large variety of scoring formulas aimed to aggregate price and quality aspects. Some methods evaluate each offer within a criterion using a point scale, weighing the points using the weights of each criterion to reach a score. The different models can give different outcomes when applied to the same bid, but can still meet the authorities requirements for transparency and predictability.[5] In [23] 38 different scoring formulas and their usage for MEAT tender evaluations is analysed, indicating that the priority weights put on quality and price is dependent on the selection of scoring formula and that many widespread scoring formulas are overly sensitive to price. A fundamental property of the scoring formula used is however buyer preference consistency, meaning that the top-ranked supplier as advocated by the scoring formula also shall meet the preferences of the buyer better than the other suppliers. In order to ensure consistency, Ref. [20] advocates that instead of setting a score on price, setting a price on quality. The problem with this approach is that it is cognitively demanding, and in more complex procurement situations it becomes an utterly delicate activity that the buyer must be capable of.

From the viewpoint of decision science, the award stage in a procurement process is a multi-criteria decision problem. There are several approaches to multi-criteria decision making, the key characteristic being that there are more than one perspective (criterion, aspect) to view the alternatives and their consequences from. We have during the latter decades developed various computational methods for formal decision analysis with imprecise information also implemented in toolkits for multi-criteria decision making (see, e.g., [4, 7, 8, 10]) and will below discuss how to apply some aspects of these on the problems addressed above. The system described in this chapter is essentially a model driven DSS. It supports procurement decisions by making a model of the alternative actions available in an actual procurement decision situation and modelling the possible selection of each supplier as one distinguishable course of action in a decision tree.

6.2 Decision Analysis for Procurement

As previously mentioned in the former section, there may be several reasons why a certain procurement process fails and ends up in an unsatisfactory evaluation result, but in many cases, failure can be attributed to the methods employed to evaluate the bids when selecting a winning supplier. The point-of-departure of this chapter is that the difficulties emanate to a large extent from three key issues in the procurement process.

The first issue has to do with *unrealistic precision* when stating the weights of different criteria in tender documents. The second issue is about assigning and comparing *qualitative values* in a somewhat naïve way, without ample regard to the

[5] Att utvärdera anbud - utvärderingsmodeller i teori och praktik, Rapport 2009–10, Konkurrensverket (in Swedish).

profound difficulties of, say, comparing monetary, ethical and aesthetic values. The third and final issue is the problem of managing *value scales* without the relevant technical understanding. Below, we explain each of these in some detail.

Assume that we intend to negotiate a contract for a consultancy service in the form of an interior designer. The procurement process is typically divided in three stages. The first stage is creating *specifications* about what the service should consist in, what specific tasks that should be performed, what evaluation criteria we will use and the relative weights of these. The second stage is the *selection* of a set of suppliers meeting the specifications. The third and final stage is deciding which submitted tender is the most preferred one, and *awarding* a contract to the winning tender. When awarding the contract, we will look at several criteria as defined in the specification.

Monetary cost is one of these, but not the only one. Using multiple criteria in this way is common practice in public procurement, and is for example in accordance with the international GPA (Agreement on Government Procurement) treaty, where the European Union is a party. Employed criteria could, apart from price/cost, be, e.g., quality, technical merit, aesthetic and functional characteristics, environmental characteristics, operational costs, cost effectiveness, service costs, technical assistance, delivery date and delivery periods, period of completion, social considerations, sustainability and level of innovation.

We will in this example be using four different criteria, price being one of them. The criteria that have been established are cost, competence, responsiveness, and design concepts, further specified like this:

- Cost—the full monetary cost of the service, divided into hourly rates.
- Competence—mainly how well experienced and/or educated in the field the contractor is.
- Responsiveness—we will conduct interviews with the potential contractors, and make an assessment of the suppliers responsiveness to the relevant demands.
- Design concept—each supplier is supposed to describe how the task could be carried out, and the description will be evaluated in terms of creativity, style, level of innovation, etc.

We have four bids from suppliers *A*, *B*, *C* and *D*, all of which have submitted well-prepared tenders. When using several criteria in the award process, we need some way of expressing the importance of a certain criterion compared to another. The GPA treaty and several corresponding national regulations state that this “relative importance” of each criteria also should be presented to the potential suppliers, for transparency reasons.[6] According to EU regulations, if the authorities use different criteria when evaluating tenders, “each applicant should be informed of the different weighting given to the different criteria (for example price, technical characteristics and environmental aspects)”.[7] There are several ways of stating

[6] https://www.wto.org/english/docs_e/legal_e/rev-gpr-94_01_e.htm.

[7] http://europa.eu/youreurope/business/public-tenders/rules-procedures/index_en.htm.

the relative importance. The perhaps most straightforward method is to give each criterion a numerical weight. Further, weights may be assigned a range instead of a fixed number, where the application of such criteria weight intervals is up to the contracting authority.

A common method is to first split the weights between price and quality, and then to further specify quality using sub-criteria. When it is not possible to provide weights that are based on objective measures or proxy values, the criteria can be listed in descending order of importance, a mere ordinal ranking. In our case, we assign each criterion a percentage. We assume the following weights:

- Cost is 40 %
- Competence is 30 %
- Responsiveness is 20 %
- Design concept is 10 %

In our example, we will classify the bids using a five level scale within each criterion. In the tender documents, the scale is described as follows:

5 Much better than the criterion base level
4 Better than the criterion base level
3 Meets criterion base level
2 Somewhat worse than the criterion base level
1 Not corresponding to the criterion base level

When the tenders from the four suppliers are evaluated using this scale, we get the matrix in Table 6.1.

A direct observation is that no supplier strictly dominates another in the sense that it has a higher value than any other for some criterion, and is equally good for all the remaining ones. We therefore cannot directly select a winner without a continued analysis. The next step is to multiply the values for each criterion with the corresponding weight, and add the result to get a weighted sum total. In our case, this becomes:

$$V(A) = 0.4 \cdot 5 + 0.3 \cdot 2 + 0.2 \cdot 2 + 0.1 \cdot 4 = 3.40$$

$$V(B) = 0.4 \cdot 4 + 0.3 \cdot 4 + 0.2 \cdot 3 + 0.1 \cdot 3 = 3.70$$

$$V(C) = 0.4 \cdot 2 + 0.3 \cdot 3 + 0.2 \cdot 5 + 0.1 \cdot 1 = 2.80$$

$$V(D) = 0.4 \cdot 1 + 0.3 \cdot 5 + 0.2 \cdot 2 + 0.1 \cdot 5 = 2.80$$

Table 6.1 Evaluation of tenders

Supplier	Cost	Competence	Responsiveness	Design concept
A	5	2	2	4
B	4	4	3	3
C	2	3	5	1
D	1	5	2	5

Using this way of evaluating, we should award the contract to supplier B. Now, the obvious question is whether we made the right choice or not. The answer is simply that we cannot know for sure. The model has too many short-comings to be of any substantial guidance in this respect. In the introduction of the paper, we indicated that there are three fundamental problems pertaining to the model as it is defined above.

6.2.1 Unreasonable Precision

It is not always (indeed, rarely) possible to specify weights with any higher degree of precision. When making non-formalized, everyday choices, we routinely employ weighing of different criteria, but almost never with a precision close to a fixed percentage. When, for example, buying a new car, we typically use several criteria for selecting a make and model, such as price, comfortability, design, social value, etc. We typically rank one criterion equal to, lower than or higher than another criterion, perhaps with a qualifying "much higher/lower" than, but that is about the level of precision we can expect in an everyday situation. In the present context, where stakes could be substantially higher than in the everyday scenario, we need a more formalized and transparent decision procedure relying less on intuition. However, to avoid unrealistic precision, the procedure should preserve some level of imprecision. Using fixed percentages is not a good representation of how we normally deal with several criteria simultaneously, and therefore the model should allow for some kind of fuzziness, such as ranges.

The EU legislation[8] also recognizes the difficulty of assigning precise numerical weights and allows for ordinal rankings of the criteria. However, in real-life procurements, weights are commonly treated as precise statements, even though this is not a legal requirement. This is, in our view, the first of the three systematic mistakes presented above. One way of presenting a mere ranking of the criteria involved is to assign ranges instead of precise percentages. The relative importance of the criterion "Competence" could be presented as "25–35 %", etc. In an "improved" version of the model, we will therefore use ranges instead of fixed percentages. In terms of calculations, it becomes a bit harder to evaluate options with imprecise statements, but there are methods for this as we will see below.

Now assume that the weights of each criterion are the following percentage ranges:

- Cost is 35–45 %
- Competence is 25–35 %
- Responsiveness is 15–25 %
- Design concept is 5–15 %

[8]Article 90 of the EU directive on public procurement, available at http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0024&from=EN.

Table 6.2 Evaluation of tenders using intervals

	Cost	Competence	Responsiveness	Design concept
A	4–5	1–3	1–3	3–4
B	3–4	3–4	3–4	3–4
C	1–3	3–4	4–5	0–2
D	0–2	4–5	1–3	4–5

The difficulties encountered with precise values apply even more to the evaluation of the alternatives within each criterion. In the same way as with the criteria weights, we can preserve a certain level of imprecision and use less precise values by assigning intervals to the alternatives within each criterion (Table 6.2).

As before, we multiply the assigned values with the percentages, but in the improved model we use the sum of the lower-end values multiplied with the lower-end percentages as the lower-end of the sum total range, and vice versa for the higher-end values and percentages. The minimum and maximum aggregated values for A are

$$\min(V(A)) = 0.35 \cdot 4 + 0.35 \cdot 1 + 0.25 \cdot 1 + 0.05 \cdot 3 = 2.15$$
$$\max(V(A)) = 0.45 \cdot 5 + 0.25 \cdot 3 + 0.15 \cdot 3 + 0.15 \cdot 4 = 4.05$$

So for all alternatives we get the following value ranges

$$V(A) \text{ between } 2.15 \text{ and } 4.05$$
$$V(B) \text{ between } 3.00 \text{ and } 4.00$$
$$V(C) \text{ between } 1.80 \text{ and } 3.80$$
$$V(D) \text{ between } 1.45 \text{ and } 3.65$$

We now see that the situation, not surprisingly, is not as clear-cut anymore. In fact, the lack of precision in the outcome of the improved model only reflects the corresponding property of uncertainty in everyday, intuition-based evaluation using several criteria. It demonstrates that allowing precise values in most cases is an over-simplification that generates unwarranted precision in the result. Intervals are a bit more difficult to assess, in particular regarding the assessments of the qualitative criteria, but at least we can represent and preserve some aspects of the inevitable imprecision in the background information. This is a commendable feature of the model compared to the original version, not a drawback.

6.2.2 *Handling Value Scales Over Qualitative Estimates*

Using point scales for handling qualitative values can be problematic. If we for some reason cannot use previously agreed-upon measurements or proxies, the evaluation of alternatives must rely on some kind of intuition or subjective sentiment. When subjectively comparing two alternatives with regard to the same qualitative criteria, we can usually without great difficulty state that the one is better than the other. We have a good understanding of what it means to prefer one alternative to the other. But if we also are required to state that preference or "betterness" using, for example, the five-point scale used in the example above, we run into problems. What information, exactly, is a "5" supposed to carry, except that it is better or more preferable than a "4"?

Even if we have some objective measure available, it is of great importance how this measure translates to a point scale. It is important to realize that the ranking order of the alternatives is not the only thing that matters when evaluating alternatives according to a point scale. The fact that our evaluation uses a cardinal scale means that the numerical values used are significant. The actual point values assigned to the different alternatives can be a factor that decides what alternative is ultimately the winner, even if the order between them is kept invariant. Suppose that the suppliers in the example above have presented tenders that rank as follows according to the criterion Cost ("5" being the best, "1" the worst) in Table 6.3.

It would now be reasonable to ask whether we could have constructed the scale a bit differently. For example, it might be more reasonable to give 1 point instead of 2 to the alternative evaluated as "Somewhat worse than the base level", and to adjust the rest of the scale accordingly. With another numerical scale, we might end up with the points in Table 6.4 instead, maintaining the order of the alternatives.

Note that the alternatives are ranked in the same order as before. The only difference is the point value assigned to alternatives *B*, *C* and *D*. If the evaluations

Table 6.3 Ranking of suppliers according to Cost

	Cost	Value
A	Much better than the base level	5
B	Better than the base level	4
C	Somewhat worse than the base level	2
D	Much worse than the base level	1

Table 6.4 Revised ranking of suppliers according to Cost

	Cost	Value
A	Much better than the base level	5
B	Better than the base level	2
C	Somewhat worse than the base level	1
D	Much worse than the base level	0

in the other criteria stay the same as in Table 6.1, we get the following revised point totals:

$$V(A) = 3.40$$
$$V(B) = 2.90$$
$$V(C) = 2.40$$
$$V(D) = 2.40$$

Now supplier A becomes the preferred one. Not paying enough attention to how the cardinality of a point scale can impact the end result will lead to arbitrariness and transparency problems. This is an obvious problem when we are handling more qualitative aspects of the criteria, but in general, the problem is the same for any criteria where there is no linear relationship between the points assigned and an objectively measurable value.

6.2.3 Deficiencies in the Handling of Value Scales

In a situation where different criteria are to be valued and weighed against each other using a common measurement such as a ten-point scale, it is important to be aware of the fundamental difficulties of measuring completely different things using the same scale. Even if legal frameworks require weighing of criteria according to a certain method, the question is whether the framework is comprehensive enough to take account for the how applying the value scale may impact the outcome, and in some cases even completely offset the initial weights. By themselves, the weights are without semantic content. Assume a hypothetical procurement where we only have two criteria to take into account. According to the EU directive 2004/17/EG on public procurement, we have to specify how the different criteria will be weighted when assessing the tenders. Assuming the weights would be

- Cost is 50 %
- Quality is 50 %

We receive two bids from suppliers, A and B. We create a score table as follows on a ten-point scale that we have defined in the specifications (Table 6.5)

Table 6.5 Evaluation of suppliers

	Cost	Quality
A	6	4
B	4	6

and obtain:

$$V(A) = 0.5 \cdot 6 + 0.5 \cdot 4 = 5$$
$$V(B) = 0.5 \cdot 4 + 0.5 \cdot 6 = 5$$

Now instead assume that when receiving the bids, we realize that we actually wanted another weighing, where quality should be more important than cost. In hindsight, we understand that what we really wanted was these weights:

- Cost is 25 %
- Quality is 75 %

This would give the following result:

$$V(A) = 0.25 \cdot 6 + 0.75 \cdot 4 = 4.5$$
$$V(B) = 0.25 \cdot 4 + 0.75 \cdot 6 = 5.5$$

However, we have already specified the weights in the tender documents for the legal requirement to be met, so we cannot alter those. But we can instead redefine the scales by calculating scaling factors. Assume that we have the weights w_i originally provided. Let v_i be our new weights and calculate $z_i = w_i/v_i$ (z_i are thus scaling factors for v_i). The scaling factors in our example are $25/50 = 0.5$ and $75/50 = 1.5$. Multiply the values with these and recalculate the mean values and keep the former weights (the law requirement is by this still fulfilled). And now we obtain (Table 6.6):

$$V(A) = 0.5 \cdot 3 + 0.5 \cdot 6 = 4.5$$
$$V(B) = 0.5 \cdot 2 + 0.5 \cdot 9 = 5.5$$

We simply adjusted the scale so that we obtain the desired result anyway, without changing the weights. The weights that we initially stated are preserved, but we shifted the scales so that they fully meet our new, revised preferences. Similarly, the order of the alternatives within each criterion is preserved. Naturally, this leaves a window for arbitrariness which by a skilled official can be used to circumvent the legal requirements of fairness, to award any winner he or she prefers. At first glance, the possibility to offset the weights completely by adjusting the value scales makes the entire weighing and assessing process meaningless.

Now, perhaps it can be argued that fairness and transparency should require us not to reset values in this way, and that the legal framework therefore should be modified

Table 6.6 Revised evaluation of suppliers

	Cost	Quality
A	3	6
B	2	9

as to prohibit this technique. One idea might be to require that the scales should be defined initially along with the weights, and that the procuring body should be required to stick to the initial scales. But the difficulty of attaching precise points to different options remains, particularly when we are dealing with qualitative values without an objective proxy measurement. And even if the procuring entity has the best possible intentions in terms of fairness and transparency, this type of problem can arise unnoticed when you simply use any kind of unreflected intuition.

The legal requirements are therefore insufficient in terms of how the valuation should be conducted. Many organizations have made extensive procurements according to the rules and in all respects in accordance with the EU directive on public procurement, or corresponding national legislation. The scales here therefore make no real sense as they are presently being used. But, as we will see below, this can be handled by an approach handling the imprecision in the situation in a proper way.

6.3 System Requirement Analysis

When devising an effective decision and evaluation process for public procurement, the three short-comings described above should be taken into account. First, when defining criteria importance, we should only state a ranking with some low level of qualification, to avoid unwarranted precision. In the example above, we could simply state that Cost is "considerably more important" than all other criteria, followed by Competence, which is "somewhat more important" than Responsiveness which in turn is "clearly more important" than the Design Concept. Secondly, when evaluating alternatives in different criteria, we should avoid assigning precise values, at least when there is no objective measure or proxy value at hand. Thirdly, the system should take the interdependence between weights and value scales into account.

To formalize the intuitive evaluation process, we should require a thorough needs analysis and extraction procedure. Such a procedure should additionally document what is in the decision-makers' heads. If the weights have been made public in the procurement documents, we need a framework for analysing each tender with respect to how they meet the criteria stated in the specifications. The proposed idea here is to start by ranking the alternatives in relation to each other under each criterion. Conceptually, this is simple. Either an alternative is as good as another, or it is better/worse. After having considered all the alternatives, we obtain a ranking from best to worst, possibly with several options as equally good at one or more places in the ranking. We now have a simple (ordinal) ranking.

The next step is to define the distances between the different alternatives, with respect to each criterion. Start with the best alternative and compare it with second best. Enter for each such pair whether the difference between them is small, medium or large. When this is done, we have a more qualified (cardinal) ranking. Continuing the example above, assume that our rankings look like the situation in Table 6.7.

Table 6.7 Evaluation of suppliers using an imprecise cardinal scale

Cost[a]	Competence	Responsiveness	Design concept
$A \succ\succ B$	$D \succ\succ B$	$C \succ B$	$D \succ A$
$B \succ\succ\succ C$	$B \succ C$	$B \succ\succ\succ D$	$A \succ\succ B$
$C \succ D$	$C \succ\succ\succ A$	$D \succ\succ A$	$B \succ\succ\succ C$

[a] $\succ$ represents 'a little better than', $\succ\succ$ 'clearly better than', and $\succ\succ\succ$ 'considerably better than'

Table 6.8 Criteria hierarchy

Cost $\succ\succ\succ$ Competence
Competence $\succ$ Responsiveness
Responsiveness $\succ\succ$ Design concept

Next, we indicate how the different criteria relate to each other in order to calibrate the scales. The approach is similar to what we did in order to rank the alternatives under each criterion above. The procedure is as follows:

1. Compare the criteria regarding their importance. Either a criterion has an equal importance as another one, or it is more or less important. After considering all the criteria, we get a ranking from most to least important, possibly with several criteria having equal importance in one or more places in the ranking. Thereafter we have a simple (ordinal) ranking of importance.
2. Enter the distances in the ranking. Start with the most important criterion and compare it with the second criterion in the ranking. Enter for each such pair of importance difference whether it is small, clear or significant. When this is done, we have a qualified (cardinal) hierarchy of importance.

Suppose we went through the above procedure with the ranking describe in Table 6.7 and obtained the qualified hierarchy in Table 6.8 between the potentials of the criteria in the example:

The notation used is similar to the notation for ranking alternatives, $\succ$ represents a 'small difference', $\succ\succ$ 'clear difference', and $\succ\succ\succ$ 'significant difference'.

From these comparisons, the cardinal number k_i for each criterion i can be calculated. A higher cardinal number indicates a higher potential of the criterion. The calculation of the cardinal number is quite straight-forward, and does not involve manually assigning any numerical values to the criteria potentials. Thereafter the scales are calibrated to correspond to what came up through the comparisons above, i.e., let k_i be the cardinal numbers and calculate $z_i = w_i/k_i$, where z_i are the calibration factors transforming the scale potentials to the predetermined weights w_i (e.g. stated in a procurement document). Thereafter, z_i are applied to the original values.

6.4 System Design

For a decision support system in the context outlined above we need calculations and a formal representation of the decision problem to be able to compute the best option and provide means for sensitivity analysis. Note that from the requirements, no numerical values are needed to be assigned any of the options under any criterion, however interval-values should be supported. In order to enable for decision evaluation with such prerequisites, a flora of different approaches have been suggested in the literature. Some approaches stems from the idea of representing imprecision in the form of intervals, or in more general terms, set-valued statements. In interval decision analysis, numerically imprecise information is modelled by means of constraints on variables for criteria weights and alternative values, see, e.g., [4, 19, 22, 27] for approaches relying on this interpretation. Other approaches aims to capture a set of value functions consistent with provided rankings [12, 14], models that exploit fuzzy numbers to represent numerical imprecision in rank statements [17] or representing linguistic statements such as "good" and "very good" using fuzzy numbers [13].

The approach presented below, called the Delta method from its primary decision evaluation rule [see Eq. (6.3)], is based upon constraint sets that complements range constraints (interval statements) with comparative statements. An important feature of the Delta method is the embedded sensitivity analysis and that rank statements can be mixed with interval statements, see [5, 8]. The computations become rather complex, involving maximization of non-linear objective functions such as the one in Eq. (6.1), but the computational mathematics is handled by platforms described in e.g. [3, 8, 9, 11, 18].

As mentioned above, at this stage in the procurement process, the problem to solve can be viewed as a multi-criteria decision problem. One large category of approaches to multi-criteria decision problems is where the decision criteria can be arranged in hierarchies, see Fig. 6.1.

For a criteria hierarchy, on each level the criteria are assigned weights and the alternatives are valued with respect to each sub-criterion. Flat criteria weight approaches can be seen as a special case—a one-level hierarchy. The minimum (or sometimes maximum) of the weighted value in Eq. (6.1) is usually employed as an evaluation rule.

For instance, in Fig. 6.1 above, the value of alternative A_i under sub-criterion jk is denoted by v_{ijk}. The weight of criteria j and sub-criteria jk are denoted by w_j and w_{jk} respectively, and the product term $w_{jk} \cdot v_{ijk}$ is referred to as the "part worth" value that A_i gets from criterion jk. The weighted, or aggregated, value of alternative A_i is

$$V(A_i) = \sum_{j=1}^{2} w_j \sum_{k=1}^{2} w_{jk} v_{ijk} \tag{6.1}$$

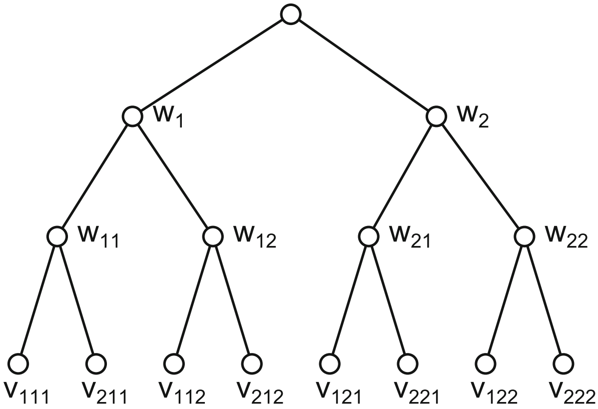

Fig. 6.1 A criteria hierarchy

Thus, if using a minimax approach, the alternative with the greatest minimum weighted value is suggested to be chosen. In this case of two alternatives evaluated under two criteria each having two sub-criteria, each weighted total value is a sum of four terms with each term containing three factors. A criteria tree is a symmetric tree. Given consequences c_i and c_j, denote their values v_i and v_j respectively. Then the user statements will be of the following kind for real numbers d_1, and d_2:

- Comparisons: v_i is from d_1 to d_2 larger than v_j, denoted $v_i - v_j \in (d_1, d_2)$ and translated into $v_i - v_j > d_1$ and $v_i - v_j < d_2$.

All the value statements in a decision problem share a common structure. Imprecise statements such as those in Table 6.7 are translated into comparisons which are constraints on the variables, and they are in turn translated into inequalities and collected together in a value constraint set. For weight statements, the same is done into a *weight constraint set*.

6.4.1 Node Constraint Set

The collection of value and weight statements in a decision situation is called the node constraint set. A constraint set is said to be consistent if it can be assigned at least one real number to each variable so that all inequalities are simultaneously satisfied.

The primary evaluation rule of the criteria tree is based on the hierarchical additive value function. Since neither weights nor values are fixed numbers, the evaluation of the hierarchical additive value yields multi-linear objective functions. Thus, given a criteria tree the hierarchical additive value of an alternative A_i, $HV(A_i)$, is given by

$$HV(A_i) = \sum_{i_1=1}^{n_{i_0}} w_{ii_1} \sum_{i_2=1}^{n_{i_1}} w_{ii_1 i_2 \ldots} \sum_{i_{m-1}=1}^{n_{i_{m-2}}} w_{ii_1 i_2 \ldots i_{m-2} i_{m-1}} \sum_{i_m=1}^{n_{i_{m-1}}} w_{ii_1 i_2 \ldots i_{m-2} i_{m-1} i_m} v_{ii_1 i_2 \ldots i_{m-2} i_{m-1} i_m} \tag{6.2}$$

where m is the depth of the hierarchy/tree corresponding to A_i, n_{i_k} is the number of child-criteria to a criterion with weight w_{i_k}, $w_{\ldots ij \ldots}$, $j \in [1, \ldots, m]$, denote weight variables and $v_{\ldots ij \ldots}$ denote value variables as above.

Optimisation with such non-linear expressions subject to linear constraints (the node and value constraint sets) are computationally demanding problems to solve for an interactive tool in the general case, using techniques from the area of non-linear programming. In the literature there are discussions about computational procedures to reduce non-linear problems to systems with linear objective functions, solvable with ordinary linear programming methods.

The area of linear programming (LP) deals with the maximizing (or minimizing) of a linear function with a large number of likewise linear constraints in the form of weak inequalities. Research efforts in the field are mainly focused on developing efficient representations and algorithms for finding local and global optima. The LP problem is the following optimising problem:

$$\begin{aligned} &\text{maximize}\, f(x) \\ &\text{when}\, Ax \geq b \\ &\text{and}\, x \geq 0 \end{aligned}$$

where $f(x)$ is a linear expression of the type $c_1x_1 + c_2x_2 + \ldots + c_nx_n$. $Ax \geq b$ is a matrix equation with rows $a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \geq b_1$ through $a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \geq bm$, and $x \geq 0$ are the non-negativity constraints $x_i \geq 0$ for each variable. Amongst all feasible points, the solution to $f(x)$ is sought that has the highest numerical value, i.e. the best solution vector x the components of which are all non-negative and satisfy all constraints. In the same way, a minimum can be searched for by negating all terms in the $f(x)$ expression. For the purposes of this system, $f(x)$ are expressions involving $HV(A_i)$ and $HV(A_j)$ for different alternatives A_i and A_j.

6.4.2 Comparing Alternatives

Alternatives are compared according to their (hierarchical) additive value. However, for evaluation purposes the notion of strength is introduced, where the strength of A_i to A_j is simply the difference

$$\delta_{ij} = HV(A_i) - HV(A_j) \tag{6.3}$$

where $\delta_{ij} > 0$ would mean that A_i is preferred to A_j. However, if $HV(A_i)$ and $HV(A_j)$ are interval-valued, δ_{ij} is interval-valued as well which may lead to overlapping value intervals meaning that preference is not that straightforward to conclude. To handle this, the concept of contraction has been proposed as an embedded form of sensitivity analysis when dealing with overlapping value intervals.

The contraction analysis consists of (proportionally) shrinking the range of each weight and value interval while studying $\max(\delta_{ij})$ and $\min(\delta_{ij})$ at different contraction levels. This means that the optimisation problem becomes obtaining $\max(\delta_{ij})$. The level of contraction is indicated as a percentage, so that for a 50 % level of contraction the range of each variable (weight, value) interval has been reduced to half their initial length and for a 100 % level of contraction to the marginal centroid value. Hence, contraction is done towards each polytope centroid given an (implicitly stated) uniform second-order distribution over each solution set, see, e.g., [8]. This means that a traditional point-wise result (coinciding with the 100 % contraction level) can be analysed in all dimensions simultaneously by widening the intervals from a point (interval with zero width) to for example 50 % contraction (halfway between the point and the maximum intervals consistent with the input information). The underlying assumption behind the contraction analysis is that the there is less belief in the outer endpoints of the intervals than in points closer to the most-likely point, which also has been verified in, e.g., [24]. The level of contraction required in order to get δ_{ij} being always positive or always negative is then used as an embedded form of sensitivity analysis, used to inform on the confidence in the resulting ranking of alternatives.

6.5 System Development

The starting point of the system development was the identification of procurement vendor selection processes with general MCDM structures and solution methods. The selection of one or more vendors from a set of offers in a procurement process can be seen as equivalent to the problem of selection one or more alternative courses of action from a predetermined set. The latter situation is a traditional multi-criteria problem, with each vendor mapped onto one course of action and the criteria that determines the best offer are of the same nature as ordinary criteria in MCDM. The importance of such criteria can, in procurement as well as in general MCDM, be expressed as (explicit or implicit) weights. By this identification, procurement evaluation processes can be mapped onto MCDM structures and that was the realisation that led to the development of the system in this paper.

The task of designing and developing software containing complex algorithms that are not easy to imagine and completely specify beforehand requires some specific approach regarding choice of design methods. In this article, we discuss the development of algorithms originally intended for decision analysis [6] and now adapted for procurement evaluation. The moving target nature of developing this system required different development techniques and approaches than more ordinary software development efforts would require.

Developing software containing complex algorithms differs from everyday software development in some respects. In most software development, the design can be planned in an orderly fashion using experience or extrapolation from previous development projects. In many cases, parts of the code can even be reused or at least patterns of design can be reused. But in designing algorithm centred software, containing new algorithms or new requirements unknown at specification time, what is normally good software design practices cannot always be applied or would not lead to effective development work.

For example, while object-oriented design and coding is often good practice, it might become a hindrance when there are no natural objects to discover or structures cannot be manipulated in detail in an implementation independent way. This article describes the use in procurement evaluation of a software library for decision analysis that was developed for maximal efficiency and minimal footprint. At its core is a linear optimisation algorithm but not of the traditional kind trying to solve a very large problem using minutes of CPU time or more. This library is the basis for an interactive procurement tool and the response needs to be experienced as immediate. Also, there is not one but a set of LP problems to be solved in sequence in order to obtain the sensitivity analysis (contraction). The sequence of LP problems solved are correlated and most of the effort expended is used to find, keep track of, and exploit the similarities of the problems in order to minimise the solution times.

6.6 User Interface Design

The user interface should enable for a user with little or no previous experience of multi-criteria decision analysis to provide input statements conforming to the Delta method. Two different user interfaces exists, one consisting of the *Decide*IT decision tool implemented in the Java programming language, and one implemented as a web based tool in Node.js called the "Preference Decision Wizard". Both tools employs a step wise process for modelling and evaluation of multi-criteria decision problems, and designed in collaboration with the Stockholm University procurement unit. The step wise process follows the typical decision analysis process. The process holds with the following steps:

Step 1—Set suppliers
The set of suppliers whose tenders fulfil the basic requirements is added. These are the alternatives of the decision model.
Step 2—Set main criteria and optionally also sub-criteria
The set of evaluation criteria is added. Some criteria may be further divided into sub-criteria. See Fig. 6.2.

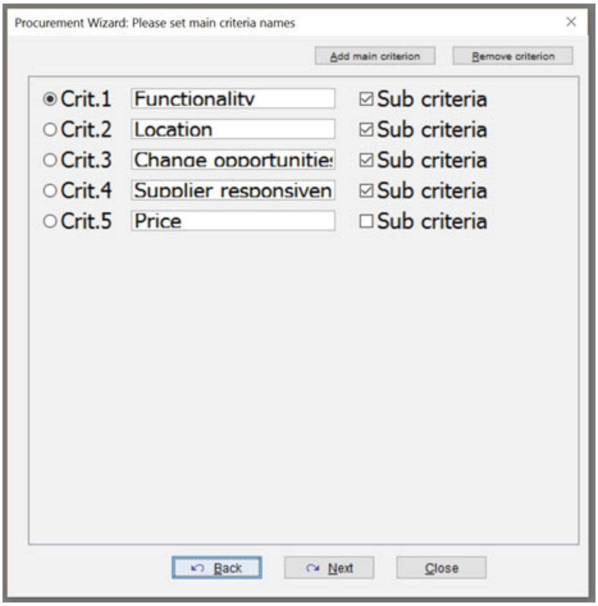

Fig. 6.2 Entering main criteria

Step 3—Value/rank suppliers
A cardinal ranking of the suppliers is done for each criterion. In the case of sub-criteria, this is done for each sub-criterion. See Fig. 6.3.
Step 4—Weigh/rank criteria
A cardinal ranking of the criteria is done, both for the main criteria and the sub-criteria belonging to the same main criterion.
Step 5—Evaluate suppliers
Decision evaluation of the suppliers are done. Essentially, all evaluation methods of the *Decide*IT tool is available, although the procurement module propose a default evaluation presentation format in the form of part-worth bar charts, showing the contribution from each criterion in a bar chart where the contribution is obtained from the product of the weight and value centroids.

6.7 System Implementation

The computational library package was initially designed using a contract based specification, using an object-based approach (where object-based refers to object-oriented minus inheritance). The use of non-inheriting objects led to a design that could survive changing requirements over time, while at the same time not enough natural code objects were found to allow an efficient implementation using

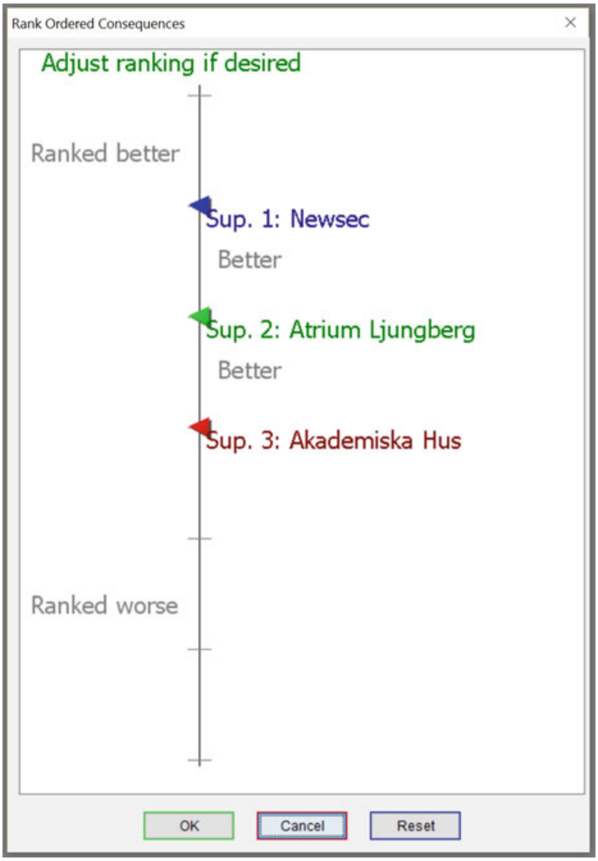

Fig. 6.3 Ranking of suppliers

object-oriented programming. Issues of code optimisation and footprint minimization were handled by using a pure imperative language without object extensions, in this case C.

Using conditional compilation and macros, techniques akin to aspect-orientation was used in coding parts particularly involving memory management, logging, and exception handling. The main challenge has been algorithmic complexity and changing requirements and specifications. One solution to the changing requirements problem is a configuration program on a meta level. The source code is then automatically recompiled prior to execution on a new platform. In this way, the source code becomes optimised for the actual target machine. Such inclusions are simpler to manage using aspect-structured code than object-oriented code since different hardware influenced aspects are more important than an object hierarchy would have been in determining the success of the software.

This is not to argue that object-orientation is less usable in general for developing calculation intense algorithms. But the point here is that code objects are not always the primary code structure choice, especially not when the outcome of the coding effort is partly unknown at the outset. The concept of aspect-oriented software design deals with architectural crosscutting concerns such as, i.a., memory management, code optimisation, and real-time behaviour. The software package was designed using an object-based overall design approach but was implemented using coding techniques more closely related to aspect-oriented techniques.

For integration purposes with the user interfaces, for *Decide*IT, the computational C library is wrapped in a C++ layer using the Java Native Interface, for the Preference Decision Wizard a C++ wrapping layer together with a C# API is implemented, using DLLImport interoperability.

6.8 System User Experience

We describe a simple example of how a more adequate support can easily manage a procurement situation and obtain a significant increase in quality in the assessment stage. The example is from Stockholm University's procurement of new facilities at Campus Kista. It was decided, in brief, to obtain, for one of the university's departments, new premises since the existing ones had become inadequate. This was a 90 million EUR investment. The criteria emphasized the premises *functionality*. Furthermore they emphasized *localization* (which implicitly was within Kista Science City) and opportunities for *interaction with the surrounding society* and the *possibilities for change*, *flexibility* and the supplier *responsiveness and innovation levels*. To these main criteria, sub-criteria were added under the relevant criteria. Finally the price parameter was asserted. Note in particular that the price criterion was in no way perceived as decisive. The main and sub-criteria were obviously mainly of a qualitative nature.

Following the announcement, it was found that there were three facility suppliers who met the criteria: Newsec, Atrium Ljungberg, and Akademiska hus. Thereafter, the analysis began and preferences could be established within the evaluation team. The notation in the tables below expresses the evaluation team preferences. For example, the ordering for criterion 1 (Symbolic value) is $a2 \succ a1 \succ a3$, meaning that it was perceived by the group that Atrium Ljungberg was the vendor that best met the criterion, followed by Newsec and then Akademiska hus. The notation $\{a1, a2, a3\}$ means that the suppliers perceived to be equivalent under the criterion. Analogously, $4 \succ 1 \succ \{2, 3, 5, 6\}$ means that criterion 4 (Realization of office space in the building) were perceived as more important than symbolic value, which in turn was seen as more important than the criteria 2, 3, 5 and 6. These latter were perceived as equivalent. The categories were prioritized in the order they are presented in Table 6.9. The most important group of criteria was the ones listed under Functionality, the second most important group was the ones listed under Location, etc. Note that the evaluation team chose to work exclusively qualitatively,

Table 6.9 Evaluation team preferences

Category	Criterion	Ranking[a]
Functionality	1. Symbolic value	$a2 \succ a1 \succ a3$
	2. Contributions to social contacts	$\{a1, a2\} \succ a3$
	3. Access to public facilities	$a2 \succ \{a1, a3\}$
	4. Realization of office space	$\{a1, a2, a3\}$
	5. Technical standard	$\{a1, a2, a3\}$
	6. Environmental requirements	$\{a1, a2, a3\}$
		$4 \succ 1 \succ \{2, 3, 5, 6\}$
Location	7. Interaction with the environment	$a2 \succ a3 \succ a1$
	8. Access to common facilities	$a2 \succ \{a1, a3\}$
		$7 \succ 8$
Change opportunities	9. Interaction during planning phase	$\{a1, a2, a3\}$
	10. Interaction during contract phase	$\{a2, a3\} \succ a1$
	11. Flexibility during planning phase	$\{a1, a2, a3\}$
	12. Flexibility during contract phase	$\{a1, a2, a3\}$
		$\{9, 10\} \succ \{11, 12\}$
Supplier responsiveness	13. Responsiveness	$\{a1, a2, a3\}$
	14. Innovation	$a2 \succ \{a1, a3\}$
		$14 \succ 13$
Price	15. Price	$\{a3, a2, a1\}$

[a] a1 = Newsec, a2 = Atrium Ljungberg, a3 = Akademiska hus

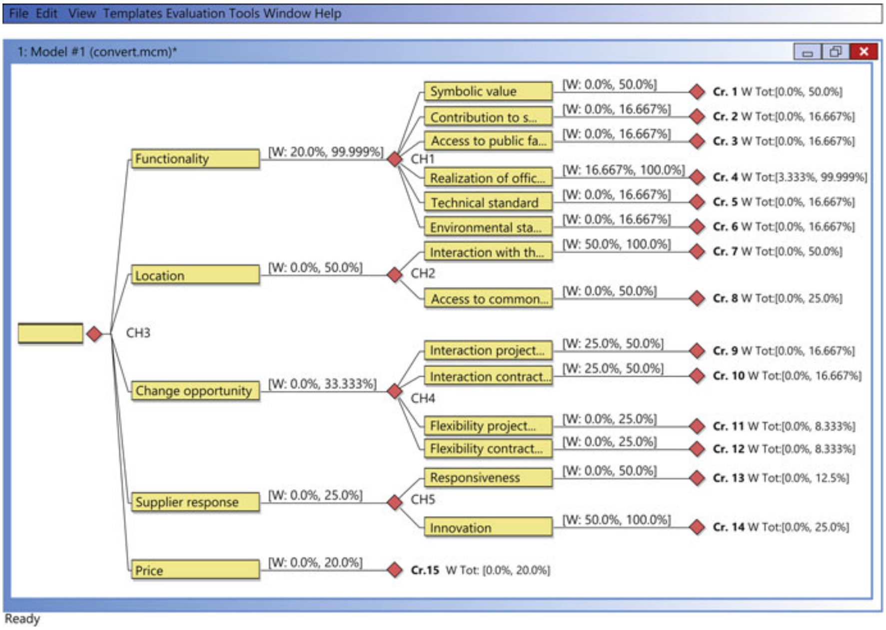

Fig. 6.4 Multi-criteria decision tree generated

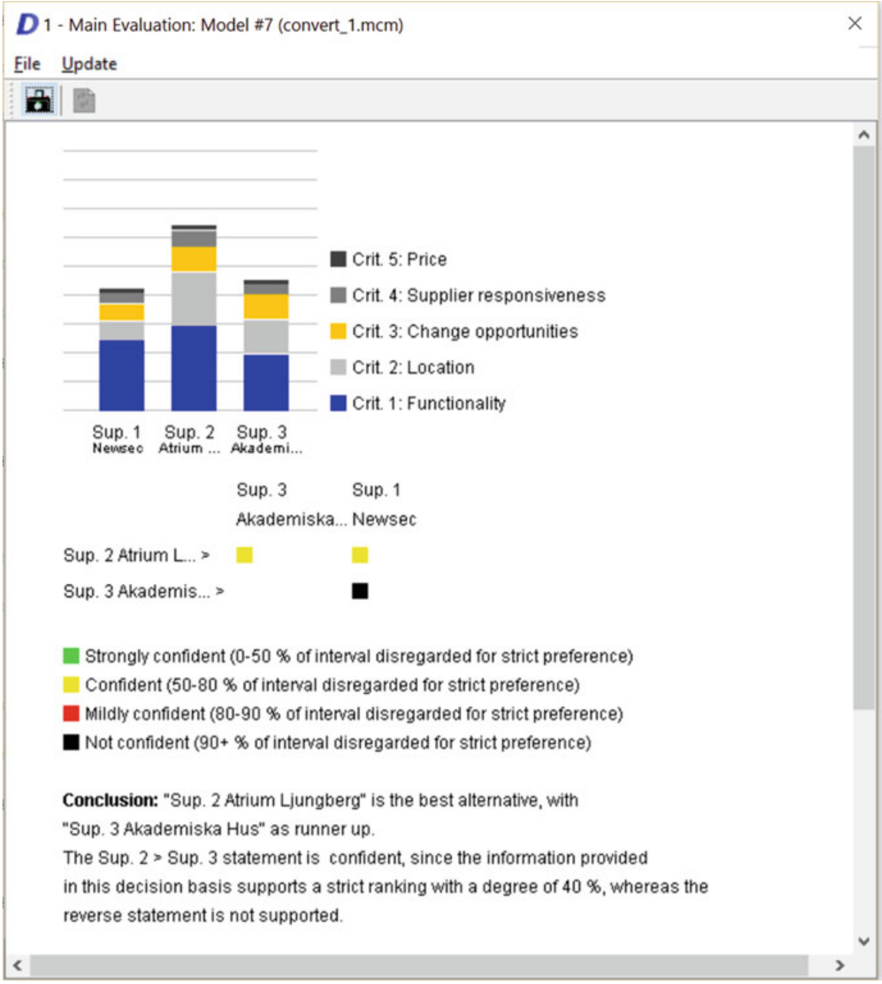

Fig. 6.5 Main evaluation window with main criteria part worth bar charts for each supplier. The level of contraction level needed in order for δ_{23} to be always positive is 40 %, here calibrated to indicate a "confident" ranking

but there are no computational obstacles to introduce range or precise figures in this analysis and mix with qualitative statements.

Using this information as input to the decision support system entered through the step wise process, a multi-criteria decision can be generated if using the *Decide*IT tool, as shown in Fig. 6.4.

Then the decision problem was evaluated, below the final result of the main evaluation window is shown, see Fig. 6.5.

The final result is not a numerical value, but something richer. The methods used managed to preserve some level of imprecision, reflecting the qualitative nature of the input, without making sacrifices in terms of transparency or deterministic output. In short, the figure above says that Atrium Ljungberg is the supplier that best met the criteria, followed by Newsec and Akademiska hus. The latter were reasonably equivalent alternatives.

6.9 Concluding Remarks

We have demonstrated three fundamental problems with the models normally used in procurement situations. These are so serious so that procurements often lose their meaning and lead to large and costly miscalculations. We have shown how by managing them systematically the quality of analysis can significantly improve and, despite the use intuitively more natural assessments of suppliers' proposals, we can get a result that gives a fuller analysis. The method not only points out the supplier who should be awarded the contract. If there is a candidate that is better than the other as it gives a much clearer picture of the situation and pointing out where analysis critical points. Note here that this cannot be done without the rather elaborate nonlinear optimisation algorithms, but a number of specially developed those found in tool *Decide*IT used in the analysis. By utilizing this tool qualitative statements could therefore be easily managed without the need to introduce artificial and somewhat arbitrary conversion formulas.

In conclusion we note that purchasers clearly need support tools beyond those that only ascertain that the formal requirements are met, i.e. there is a need for tools that are using appropriate calculation functions that support scalable management and allow imprecision in value and weight statements. These features can then be easily integrated into a complete support tools to support both the formal process calculation steps. *Decide*IT is one such tool that fairly easily and properly handle mixtures of qualitative and quantitative criteria.

The algorithms for solving the evaluations of cardinal rankings are optimisation algorithms, but of a slightly different nature than ordinary optimisation problems. In ordinary optimisation, the task is often to find a local optimum (sometimes a global) for a problem with many variables, possibly millions. This is often done in batch mode, i.e. the real-time (or interactive) requirements are low. But in this case, the design is required to solve many (hundreds) of optimisation problems in fractions of a second, the speed requirement being that the user should not experience any delay in response. For this to be possible, a network of result caches had to be devised. While the exact design of the caches are not important, it is interesting to note that these kinds of requirements are not easily anticipated before the specific procedure was produced. Thus, the overall software design depends on algorithms whose specifications are not known from the outset and whose development cannot be foreseen since there are no originals or templates to start with.

In summary, the library package is still alive and continues to evolve at the research frontier, more than 15 years after its first release without requiring a rewrite or architectural redesign. Part of its longevity, despite complexity and changing requirements and specifications, is due to the following set of principles:

- an object-based approach
- a contract based specification
- aspect-orientation-like management of key code features
- a pure imperative programming language

resulting in reasonable development control without introducing overhead in the form of over-specification, slow execution, or a too large footprint.

References

1. Bergman, M.A., Lundberg, S.: Tender evaluation and supplier selection methods in public procurement. J. Purch. Supply Manag. **19**(2), 73–83 (2013)
2. de Boer, L., Labro, E., Morlacchi, P.: A review of methods supporting supplier selection. Eur. J. Purch. Supply Manag. **7**(2), 75–89 (2001)
3. Caster, O., Ekenberg, L.: Combining Second-Order Belief Distributions with Qualitative Statements in Decision Analysis, pp. 67–87. Springer, Berlin (2012)
4. Danielson, M.: Handling imperfect user statements in real-life decision analysis. Int. J. Inf. Technol. Decis. Mak. **3**, 513–534 (2004)
5. Danielson, M.: Sensitivity analysis in interval decision modelling. Eng. Lett. **17**(1), 1–7 (2009)
6. Danielson, M., Ekenberg, L.: A method for decision and risk analysis in probabilistic and multiple criteria situations. Patent (2007). U.S. Patent No. 11/145229
7. Danielson, M., Ekenberg, L., Johansson, J., Larsson, A.: The decideit decision tool. In: ISIPTA 03, pp. 204–217 (2003)
8. Danielson, M., Ekenberg, L., Idefeldt, J., Larsson, A.: Using a software tool for public decision analysis: the case of Nacka municipality. Decis. Anal. **4**, 76–90 (2007)
9. Danielson, M., Ekenberg, L., Ekenberg, A., Torsten Hökby, Lidén, J.: Decision process support for participatory democracy. J. Multi-Criteria Decis. Anal. **15**, 15–30 (2008)
10. Danielson, M., Ekenberg, L., Larsson, A., Riabacke, M.: Transparent Public Decision Making: Discussion and Case Study in Sweden, pp. 263–281. Springer, Berlin (2010)
11. Ding, X.S., Danielson, M., Ekenberg, L.: Non-linear programming solvers for decision analysis support systems. In: International Conference on Operations Research 2003, pp. 475–482 (2004)
12. Figueira, J.R., Greco, S., Slowinski, R.: Building a set of additive value functions representing a reference preorder and intensities of preference: grip method. Eur. J. Oper. Res. **195**(2), 460–486 (2009)
13. García-Cascales, M.S., Lamata, M.T.: Solving a decision problem with linguistic information. Pattern Recogn. Lett. **28**, 2284–2294 (2007)
14. Greco, S., Mousseau, V., Slowinski, R.: Ordinal regression revisited: multiple criteria ranking using a set of additive value functions. Eur. J. Oper. Res. **191**(2), 416–436 (2008)
15. Harland, C., Telgen, J., Callender, G.: International Research Study of Public Procurement, pp. 372–399. SAGE, London (2013)
16. Ho, W., Xu, X., Dey, P.K.: Multi-criteria decision making approaches for supplier evaluation and selection: a literature review. Eur. J. Oper. Res. **202**(1), 16–24 (2010)
17. Jiménez, A., Mateos, A., Sabio, P.: Dominance intensity measure within fuzzy weight oriented MAUT: an application. Omega **41**(2), 397–405 (2013)
18. Larsson, A., Johansson, J., Ekenberg, L., Danielson, M.: Decision analysis with multiple objectives in a framework for evaluating imprecision. Int. J. Uncertain. Fuzz. Knowl. Based Syst. **13**(5), 495–509 (2005)

19. Lee, K.S., Park, K.S., Kim, S.H.: Dominance, potential optimality, imprecise information, and hierarchical structure in multi-criteria analysis. Comput. Oper. Res. **29**(9), 1267–1281 (2002)
20. Lunander, A., Arne, A.: Metoder vid utvärdering av pris och kvalitet i offentlig upphandling, 11 (2004)
21. McCrudden, C.: Using public procurement to achieve social outcomes. Nat. Res. Forum **28**(4), 257–267 (2004)
22. Salo, A.A., Hämäläinen, R.P.: Preference ratios in multiattribute evaluation (prime) – elicitation and decision procedures under incomplete information. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **31**(6), 533–545 (2001)
23. Stilger, P.S., Siderius, J., van Raaij, E.M.: A Comparative Study of Formulas for Choosing the Economically Most Advantageous Tender (2015). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2626934
24. Sundgren, D., Danielson, M., Ekenberg, L.: Warp effects on calculating interval probabilities. Int. J. Approx. Reason. **50**, 1360–1368 (2009)
25. Telgen, J., Uenk, N., Lohmann, W., Manunza, E.: Conditions on lots: practical desirability versus legal acceptability. In: Piga, G., Tatrai, T., (eds.) Public Procurement Policy. The Economics of Legal Relationships, pp. 90–98. Routledge, London (2015)
26. Weber, C.A., Current, J.R., Benton, W.C.: Vendor selection criteria and methods. Eur. J. Oper. Res. **50**, 2–18 (1991)
27. Weber, M.: A method of multi attribute decision making with incomplete information. Manag. Sci. **31**(11), 1365–1371 (1985)