Decision Making and e-Process Selection

Frina Albertyn
Information Technology Section, Faculty of Business and Computing,
Eastern Institute of Technology, Napier, New Zealand
Email: falbertyn@eit.ac.nz

Abstract

Multi-criteria decision making uses multiple criteria in order to identify a solution to a problem choosing one solution from a selection of possible solutions. This paper introduces and briefly discusses decision theory and how it is used in decision making. Three decision making processes have been researched to support selecting a suitable development process (e-Process) for the development of e-Commerce Information Systems (ECIS). These three are briefly summarised in this paper and include a selection method that uses Benefit Value Analysis (Weighted score selection processes), a selection method that uses Case-Based Reasoning and a selection method that uses the Analytic Hierarchy Process.

Keywords: Multi-Criteria, Decision models, e-Process.

1 Introduction

It is a complex process to make decisions in today's business environment. Technology, material, development and human resources are very expensive. The market today is also very competitive and in a very short time it is important to make well motivated effective decisions (Trueman, 1976). This is the reason why it is important to aid the decision maker as far as possible in this process. It is not an easy task to define what makes for good decision making. It is however important to take a number of different aspects into consideration when making decisions.

Decision making can be seen as **deciding on important matters:** the process of making choices or reaching conclusions (MSNEncarta, 2007). Decision theory as **study of outcomes for decisions:** the study of the best possible outcomes for decisions made under varying conditions (MSNEncarta, 2007).

According to (Oxford_English_Dictionary, 2007) the following definitions for decision is provided:

- The action of deciding (a contest, controversy, question, etc.); settlement, determination.
- The making up of one's mind on any point or on a course of action; a resolution, determination.

This Supplementary Proceedings paper appeared at the *21st Annual Conference of the National Advisory Committee on Computing Qualifications* (NACCQ 2008), Auckland, New Zealand. Samuel Mann and Mike Lopez (Eds).

Further introductory definitions can be found in Wikipedia (Wikipedia, 2007):

- A decision is the commitment to irrevocably allocate valuable resources.
- A choice or judgement.
- Firmness of conviction.
- A result arrived at by the judges.

In an article in (Wikipedia, 2007) decision making is decribed as the cognitive process leading to the selection of a course of action among variations. Every decision making process produces a final choice. It can be an action or an opinion. It begins when we need to do something but know not what. Therefore, decision making is a reasoning process which can be rational or irrational, can be based on explicit assumptions or tacit assumptions.

According to (Trueman, 1976) the different quantative methods along with the processing power of the computer represents the way that decisions can be aided. This thesis will be investigating different ways the decision makers can be supported in selecting suitable development processes for e-Commerce Information Systems.

The science of making decisions is complex and covers a wide range of topics. It involves theoretical as well as practical aspects. People are confronted with different problems on a daily basis – in their own lives as well as in the corporate environment. Choices need to be made between different alternatives based on a problem that was solved (Doumpos & Zopounidis, 2002; Pomerol & Barba-Romero, 2000).

There are a number of concepts that is important to decision making. These include ambiguity, the decision itself, action to be taken after a decision has been reached, available choices, the actual problem to be solved, the knowledge of the decision makers, the preferences of the people that will have to operated in the decision environment and consequences of the decision made – who/what will it effect and do one need to motivate why a specific decision was made – orally and documented (White, 2006 - Paperback Copyright 1969).

In (Doumpos & Zopounidis, 2002; Wikipedia, 2007) multi-Criteria Decision Analysis (MCDA), or Multi Criteria Decision Making (MCDM), is defined as being a discipline aimed at supporting decision makers who are faced with making numerous and conflicting evaluations. MCDA aims at highlighting these conflicts and deriving a way to come to a compromise in a transparent process.

There are different dimension in the decision making process. The individual (or group of) decision makers will investigate different choices of action in an uncertain environment in order to determine the best course of action. The theory of decision analysis can be seen as the support provided to the decision maker when making a choice between different pre-defined or pre-specified alternatives. The decision making process is reliant on a set of characteristics that provides information about the alternative options available (Unit-For-Sustainable-Development-and-Environment-Organization-of-American-States, 2007).

The decision making process can be made based on deterministic information (certainty about decision outcomes) or the decision making process can be based on information that has uncertain outcomes and is represented by fuzzy logic and probabilities. The quality of information can also enable the decision maker to make clear scientific determined decisions or decisions which are based on different interpretations(Unit-For-Sustainable-Development-and-Environment-Organization-of-American-States, 2007).

The wide variety of different types of decision problems and the quality of the information available at the time of making decisions identifies the need for methods and techniques that can assist the processing of the available information as well as ultimately lead to better decision making(Unit-For-Sustainable-Development-and-Environment-Organization-of-American-States, 2007).

Powerful mathematical programming tools has emerged using multiple objectives to assist in the process of searching for decisions which best satisfy a multitude of conflicting objectives, and there are a number of distinct methodologies for multicriteria decision-making problems available. These methodologies can be classified in a number of ways, such as form of model (e.g. linear, non-linear, stochastic), characteristics of the decision space (e.g. finite or infinite), or solution process (e.g. prior specification of preferences or interactive)(Pomerol & Barba-Romero, 2000; Unit-For-Sustainable-Development-and-Environment-Organization-of-American-States, 2007).

In 1896, Pareto first defined a basis for solving decision problems using multiple criteria (Doumpos & Zopounidis, 2002; Pareto, 1896). Multi-criterion analysis was defined and formulated more thoroughly in 1960 as: Multi-criteria analysis is the study of solving a problem by choosing one of the alternative solutions to a problem based on a set of multiple criteria associated with each solution.

At a mathematical conference in The Hague, Netherlands in 1970 a section was devoted to multi-criterion analysis. A number of publications resulted in the next 4 years on this topic (Pomerol & Barba-Romero, 2000). Since then multi criteria decision making has been the topic of much research. Multi criterion decision making can be seen as the field of activity in which practical applications as well as informatics are dominant. The focus in this process is in providing practical ideas and methods with depth that can solve the problems associated with decision making (Pomerol & Barba-Romero, 2000).

In every day life people are confronted with different decision that needs to be made. There are usually a number of options or probable solutions available say s_1, s_2, …s_n of a set of possible solutions S to the problem P. Each of these solutions can be evaluated using a set of criteria say c_1,c_2..c_n of a set of evaluating criteria which can help to differentiate between the different solutions. Each of these criteria can take specific values of a range say 0…n which will help to rate each of the solutions. Our task in this thesis is to identify a range of methods that can be applied to the solutions in order to allow a suitable development process to be identified for e-Process development.

2 Problem classification

It is not easy to classify problems, but as a main classification problems can be classified into two different groups. These are:

- **Discrete problems:** Discrete problems are problems that contain and involve the evaluation of a discrete set of alternatives. Each alternative problem solution is classified in terms of a set of aspects. These aspects form the basis for deciding between the different choices available.
- **Continuous problems:** Continuous problems are problems that contain and involve the evaluation of an infinite set of possible alternative solutions. When there are an infinite set of solutions the answer to problem solution will lie in a region rather than with a specific solution (Doumpos & Zopounidis, 2002).

The research done for this thesis all lie within the discrete problem solution domain.

2.1 Quantitive and qualitive decision rules

A decision can be seen in the following framework: If S is a set of states and X is a set of potential consequences of decisions made then the act can be seen as the mapping f from S to C, where for each s ε S the act f produces a result f(s) ε X. Sometimes the different acts are ranked in order to support the decision maker in the decision making process. This is often done using a specific, often numerical ranking method. (Riccia, Dubois, Kruse, & Lenz, 2003)

When considering discrete problems there is a number of different types of analysis that can be done on the problem according to (Doumpos & Zopounidis, 2002) in order to support decision makers. These include selecting a solution from the alternative solutions; sorting the solutions into order of preferences and then selecting a solution based on the raking done; and classifying solutions into predefined groups and then identifying which group would provide the best answer.

There are a number of factors that influence decision making and situations calling for these decisions to be made are never the same and differ from case to case. The factors that have an impact on the decision include:

- The context in which the problem is being defined which include the characteristics of the problem such as how well-defined it is; what we know about the problem domain and; how many solutions we need to choose from.

- The people involved with the decision making process which include their background and knowledge about the problem domain; how many people needs to be accommodated; who is responsible for which part of the problem solution; who is accountable for which decisions that needs to be made.
- Knowledge and characteristics of the decision-makers which include how intelligent and knowledgeable they are; their risk profile and whether they can accept a measure of uncertainty and their pre-knowledge and experience in the problem domain.

(French & Xu, 2005)

When decisions are made there are a number of factors that needs to be considered. These include:

- Firstly, the people involved with the decision making process – decision makers, experts, different problem domain stakeholders. Not all of these people will necessarily be involved from day one so it is necessary to determine when and on which level they will be involved. It is also necessary to define the roles that each of these people are going to play in the decision making process.
- The second factor to be considered is the methodologies that the people listed above prefers or believes in. If a specific method has always been preferred in the decision making process that is most probably going to be the preferred method in this process.
- The third aspect to be considered is how long this process can take – does the decision need to be made in one meeting or do the decision makers have months to consider their choice.
- The level of detail required in documenting the decision process can play a roles in the decision making process adopted.
- The level of involvement or impartiality of the decision makers needs to be factored into the decision making process.

(French & Xu, 2005)

Other factors that needs to be considered when deciding on how to approach the decision making process includes aspects such as:

- The structure of the problem which in this research includes the quality aspects to be included in the decision making process applied.
- How different weights and values used in the decision making process will be gathered and applied in the problem environment.
- How the data will be presented and whether sensitivity analysis will be applied.

(French & Xu, 2005)

Sometime the problem domain is clearly defined and in other cases the solution is difficult to identify. All of these aspects listed above will be further investigated in the application to the problem domain in the rest of this document.

2.2 Software options

There are a number of software packages available that supports decision analysis. In (French & Xu, 2005) it is stated that a survey in 2004 listed 45 different packages for decision analysis. All the packages are stated to be versatile, user-friendly and with the capability to accommodate fairly substantial problems (French & Xu, 2005). As each of these packages have their own view of the world and the problem to be solved the author of this thesis are focusing on the development of a prototype decision analysis software that focuses on e-Process Selection using different decision analysis methods.

There are a number of software packages available that supports the different schools of though. These include multi-value decision analysis (MAVT) (Keeney, 1992) and analytic hierarchy process (AHP) (Saaty, 1990) which will be covered in this thesis.

In the rest of the paper we will be presenting three different decision processes that have been applied to the problem of deciding on a suitable e-Process to use when developing eCIS.

3 Benefit Value Analysis (Weighted score selection processes)

As published in [Albertyn, 2005] the following selection methods is based on Benefit Value Analysis. Similar to [Fettke and Loos, 2003; Opdahl et. al., 2001, Wand et. al., 1999] as based on Tom Gruber, we consider an ontology as a specification of a conceptualisation that is shared by a number of people. For the current purpose, i.e., assessing e-Processes it is best that this specification takes the form of a list of concepts (including definitions of these) to characterise e-Processes. According to our discussion of related work above the quality aspects are listed in the appendix.

The proposed selection method does not focus on a specific heuristic. The focus is rather on organizing e-Process knowledge in a suitable way and storing it for later use. A meta-heuristic is introduced based on the idea of patterns. An e-Process pattern here is considered as a triplet (context, problem, e-Process) where it is assumed that the e-Process mentioned in such a triplet is an acceptable solution to the problem in the given context. For each e-Process, a score is introduced, to measure how well it suits as a problem solution in a given context. The meta-heuristic can be defined as (obtain a list of e-Process patterns and respective scores):

```
LOOP - CHOSE an e-Process from the list
      APPLY the e-Process
      ASSESS its success
      UPDATE the e-Process score
ENDLOOP
```

To formulate the initial heuristics some conventions are introduced. Let there be m characteristics c1,…,cm and n e-Processes P1, … , Pn. Then each e-Process can be represented by an m-tuple of numbers between 0 and 1. For each characteristic ci, i∈{1, … , m}, and each process Pj, j∈{1, … ,n}, we ask an enterprise staff who is an expert in the field to determine the weight w(1), …, w(m) of characteristic c(1), …, c(m) and the performance p(1,j), …, p(m,j) of process Pj with respect to characteristic c(1), …, c(m) respectively. We chose the numbers w(i) such that $\sum$ w(i) = 1 and 0 ≤ w(i) ≤.1, ∀ i∈{1, …,m}. Our initial heuristics is then:

1. C : = ∅
2. Chose J ⊆ {1, …, n} such that for j ∈ J the number S(j) = $\sum_{i\in\{1,\ldots,n\}}$ with w(i)*p(i,j) is maximal and define C : = J, observe to chose J maximal. Set C : = C ∪ J.
3. For all j ∈{1, …, n}\ J, k∈{1, … , m} perform a sensitivity analysis, i.e., Calculate S(k,j) = $\sum_{i\in\{1,\ldots,m\}\setminus\{k\}}$ w(i)*p(i,j). Chose sets W, P such that w(k)∈ W, p(k, j)∈ P and determine T(k,j) = max{w*p | w∈W, p∈P}. It will often be convenient to chose W and P such that W = {$c_W + h_W*\Delta_W$ | h_W ∈ {0, …, r_W}}, P = {$c_P + h_P*\Delta_P$ | h_P∈ {0, …, r_P}}. If S(k, j) + T(k, j) > S(j), then the values w, p for which the maximum T(k, j) was achieved need to be investigated. If they are reasonable, then redefine C : = C ∪{k}.
4. Do weak point analyses for each candidate in j ∈ C, i.e., determine those characteristics with high impact (weight higher than for, e.g. 70 % of the characteristics) and low performance (performance lower than for, e.g. 70 % of the characteristics).
   a. For each of the weak points consider the performance assessment and weight. If one of these should be corrected then do so.
   b. If weak points remain after a. then either C : = C \ {j} or replace P_j by an improved version Q scoring no less than P_j, and assess it.
5. If the weak point analysis in 4 does not change anything then chose among the candidates in C according to a predefined strategy. Otherwise go back to 3.

Note that the definition of weak point used here is somewhat arbitrary in that the thresholds of 70 % are not justified. One thus can attempt tuning the heuristics changing these values. In particular the threshold values could be chosen differently. It is proposed that a structured list of quality aspects (characteristics) is used to choose between different e-commerce ISD processes see the appendix. Each characteristic will be weighted to show its importance, with the weighted values adding to one.

4 Analytic hierarchy process

As published in [Kaschek, 2007] the following selection method is based on the Analytic Hierarchy Process. We consider three quality aspects of a method for e-P selection as essential, namely economically viable, justifiable, and provide insight into the selection made. We achieve viability of our approach by relying on the expertise of those who, in a case at hand, have to make a selection. This includes admitting qualified individuals to play both expert roles that we employ in our method. If the experts consider the selection task as clear-cut and one of the admissible e-Ps as the obvious one to select then we recommend the project to just go ahead with that e-P. However, if the experts are not sure about the e-P to select then we recommend applying the Analytical Hierarchy Process (AHP) that is described below. AHP (Saaty, 1990) is a very well-known and successfully applied method for multi-criteria decision making. Obviously our selection task is an instance of that decision making task and thus justified.

Selecting one from a number of admissible e-Ps and providing insight into this selection is an instance of decision making and Simon's model of decision making has three stages - intelligence, design, and choice (Ahituv and Neumann, 1990). We focus on choice because we consider existing software processes only and presuppose, simplifying significantly, that the decision maker knows all existing software processes and knows how to identify the admissible ones. We suggest quantitative assessment for identifying top-scorers and then choosing one of them. The overall quality of e-Ps that we aim at expressing quantitatively is their suitability for developing an anticipated B2C. The method thus is SUD specific and its outcome not necessarily extendible to other cases.

In our attempt to quantify e-P quality we opt for assessment rather than measurement because we feel that soft and contextual factors need to be included that due to the attempt to achieve repeatability and independence from the measuring agent are excluded from what can impact the obtained quality score. In our method these soft and contextual factors of e-P quality are taken into account by the unrestricted capacity of the involved experts to recommend what appears as sensible to them.

4.1 System quality aspects and expert selection

Quality of e-Ps is multi-faceted (as is the case for information systems in general (Ghezzi, et.al., 2004). Individual quality aspects are often considered as too broad and unspecific and are therefore decomposed into lists of lower level quality aspects. We exploit the two-level hierarchy in (Kaschek, Pavlov, et. al., 2006). For our method the employed quality aspects are irrelevant, i.e., any other set could be chosen as well. However, the outcome of using our method is likely to critically depend on the chosen quality aspects. Due to our focus on illustrating our method we do not argue in particular for the chosen quality aspects and do not list the second level quality aspects employed:

- *e-P aspects*, i.e. the e-P's modeling notions, abstraction concepts, etc.;
- *quality concepts* of the e-P, i.e., its reliability, robustness etc.;
- *cost* for using the e-P;
- *domain impact*, i.e. the impact of the project domain;

- *usability*, i.e. the e-P's aid in developing a high quality B2C;
- *compatibility* of the e-P with other methodologies, and
- *maturity*, i.e. e-P's stability, tool support, documentation etc.

We suggest that the expertise required for our selection task is provided by humans in the role of e-P expert or system quality expert. The AHP enables us blending the experts' expertise to select the e-Ps best suited for the problem at hand. Choosing the experts is not necessarily trivial or cheap. We thus recommend considering the qualification profile of the designated developers. If that profile indicates sufficient e-P and system quality expertise then we recommend using these developers as experts. We also recommend admitting any individual for playing both expert roles if the required expertise is available. A somewhat more complicated approach to expert selection is described in (Kaschek, Pavlov, et. al., 2006). Expert modeling for expert selection likely is to consider affiliation, area of competence, standing, availability, price, etc. The AHP can also be used for choosing the experts. However, experts with expertise in expert selection might be hard to identify. According to (NUREG, 1989) expert selection should consider demonstrated experience, expert versatility, expert group diversity, and expert cooperation.

AHP-based selection proceeds in steps that we call *e-Process ranking*, *system quality aspect ranking*, and *knowledge blending*. Our method, after scoring the admissible e-Ps, proceeds with *tradeoff analysis* and *sensitivity analysis* on the data provided by applying the AHP. These two techniques make our method providing insight into the selection suggested.

4.2 e-Process ranking

The e-P experts rank the admissible e-Ps by pair-wise comparison in terms of the second-level quality aspects.

Let x, y be e-P. For expressing a number of increasingly strong preferences for x over y we define the predicate m-better(x, y), for $m \in \{1, 3, 5, 7, 9\}$. The predicate 1-better(x, y), 3-better(x, y), 5-better(x, y), 7-better(x, y), 9-better(x, y) means that no preference, light preference, moderate preference, strong preference, and extreme preference respectively is given to x over y. Let C^1, C^2 the first- and second level quality aspects respectively, $C = C^1 \cup C^2$, and for $c \in C^1$ let J_c be the set of second-level quality aspects into which c has been decomposed, i.e. $C^2 = \bigcup_{c \in C^1} J_c$, the disjoint union. Let furthermore be E a set of e-Process experts, and X a set of e-Processes. We use the predicate β(e, m, x, y, c) to denote that expert e judges m-better(x, y) = TRUE with respect to quality aspect c and define a: $C \times E \times X \times X \rightarrow \{m, 1/m \mid m \in \{1, 3, 5, 7, 9\}\}$ as a(c,e,x,y) = m, if β(e,m,x,y,c) and 1/m, if β(e,m,y,x,c). The mapping a is called comparison mapping. Obviously, a(c, e, x, y) = 1/a(c, e, y, x), and in particular a(c, e, x, x) = 1, for all $c \in C$, $e \in E$, $x, y \in X$. Let be $c \in C$ a quality aspect, $e \in E$ an expert, and for brevity X = {1, …, n}, then for the restriction A(c,e) of a to X × X holds A(c, e) (i, j) = a(c, e, i, j), for all i, j. We represent it as the pair-wise comparison matrix (Saaty, 1990). Its elements are the results of all pair-wise e-Process comparisons: $A(c,e) = (a(c,e,i,j)_{1 \le i,j \le n})$. If this matrix is consistent (i.e. a(c,e,i,k) □a(c,e,k,m) □a(c,e,i,m) for any i, k, and m), the maximum eigenvalue $\lambda_{\max}(c,e)$ of A(c,e) and the corresponding eigenvector f(c,e) = $(y_1, \ldots, y_n)$ are known. This vector, after normalization, contains the relative weights y_i of all e-Processes $i \in \{1, \ldots, n\}$. The relative weight of $x \in X$ following (Saaty, 1996) is denoted as $f_{c,e,x}$. We define f: $C \times E \times X \rightarrow [0, 1]$, with f(c, e, x) = $f_{c,e,x}$ and use the geometric means for aggregating that various experts' findings into a score of x regarding c, i.e. $f_{c,x} = \sqrt[|E|]{\prod_{e \in E} f_{c,e,x}}$, with |E| being the cardinality of E. Software implemented by Roman Pavlov, Kharkiv enables us employing the simplified version of AHP described in (Noghin, 2004), where the pair-wise comparisons are reduced to the minimum needed for determining a consistent pair-wise comparison matrix. Defining a(c,e,i,j):= a(c,e,1,j) / a(c,e,1,i), for all $1 \le i, j \le n$ implies consistence. With $w_i = (1 / a(c,e,1,i)) / \Sigma_{j \in \{1, 2, \ldots, n\}} 1 / a(c,e,1,j)$ the relative weight for all i. Assigning the m-better predicate to pairs of e-Processes according to (Noghin, 2004) can be managed such that an arbitrary e-Process $x' \in X$ at first is chosen (we always assumed x'=1). Then $y \in X \setminus \{x'\}$ is chosen such that assigning the appropriate predicate m-better(x', y) becomes simplest. Then $z \in X \setminus \{x', y\}$ with the same property is chosen, and so forth until X is exhausted.

4.3 System quality aspect ranking

The system quality experts apply two times the pair-wise comparison technique for assessing the relative importance of the quality aspects in the case at hand. First the top-level aspects are ranked and then at the decompositions thereof. We thus obtain weights w^{**}_c, and $w^*_{c,j}$ for $c \in C^1$, and $j \in J_c$. For obtaining integrated weights for the second-level quality aspects we define w: $C^2 \rightarrow [0, 1]$, $j \mapsto w^{**}_c w^*_{c,j}$, for $c \in C^1$, $j \in J_c$.

4.4 Knowledge blending

For blending the available kinds of expertise we use the "ideal synthesis" AHP mode (Forman and Selly, 2001). For that we use the maximum value $f^*_c = \max \{f_{c,x} \mid x \in X\}$ for normalization, i.e., for $c \in C^2$, $x \in X$, we denote, as usual, the c-score $w_c f_{c,x} / f^*_c$ of x as $p_{c,x}$. The score p_x of $x \in X$ is then defined as $\sum_{c \in C^2} p_{c,x} \Big/ \sum_{x \in X} \sum_{c \in C^2} p_{c,x}$. The set of e-Ps best suited for the problem at hand is considered to be the set $\{x \in X \mid p_x = \max \{p_y \mid y \in X\}\}$. The restriction to a two level hierarchy of quality aspects is conventional. If necessary we could decompose further. For each additional nesting level we could aggregate a higher-level number out of the data available at that additional nesting level

4.5 Tradeoff analysis

After selecting the set of e-Ps best suited for the problem the next step is to provide insight into the selection made. First, to express preferences regarding quality aspect implied by employing a particular e-P, tradeoff analysis is performed. Consider a case of developing an ECIS and x, y∈X be e-Ps. Adapting the approach in (Zhu, et. al., 2005) we define the tradeoff entailed by favoring x over y as the relinquishment of the top-scoring quality aspects of y that are not top-scoring for x. The purpose of the tradeoff analysis is the creation of awareness of the tradeoffs implied by employing a particular e-P.

We follow Zhu, et. al. (2005) in using tradeoff diagrams as a tool for tradeoff analysis. For each pair (c, d) of e-P quality aspects a tradeoff diagram is created. Each tradeoff diagram represents a part of the plane with c, d being represented as the abscissa and ordinate respectively. As maximum value for abscissa and ordinate the maximum value of any quality aspect weight is used in all these diagrams. In any tradeoff diagram each x∈X is represented as a point $r(x) = (p_{c,x}, p_{d,x})$. In each of these diagrams the first quadrant is divided into four squares of equal size. For each pair $(c,d)\in C^2$ with $c\neq d$ this defines an equivalence relation e(c,d) on X with (x,y)∈e(c,d) if r(x), r(y) lie in the same square. Denote the upper-left, lower-left, lower-right, and upper-right square as UL, LL, LR, UR respectively. Then favoring x over y entails favoring c over d and favoring d over c if r(x) lies in LR and UL respectively. No tradeoff is considered to occur if r(x) lies in LL or UR. However, x is considered as favoring c and d if r(x) lies in UR, and disfavoring c and d if r(x) lies in LL. An example of tradeoff diagram for four e-Ps is shown on Figure 1. On this diagram, selecting e-P(1) and e-P(2) entails favoring d over c, selecting e-P(4) – favoring c over d, selecting e-P(3) disfavors both quality aspects.

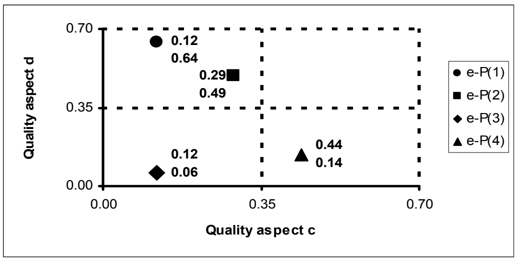

Figure 1: Tradeoff diagram example

We recommend calculating all tradeoff diagrams and considering those that indicate tradeoffs being entailed by selecting one of the top-scoring e-Ps. We recommend doing a more detailed analysis when any tradeoff diagram shows an odd tradeoff due to selecting a top-scoring e-P. A technique for a more detailed analysis is the following. Let x∈X be top-scoring and c, $d\in C^2$ such that r(x)∈UL, i.e. a tradeoff of d over c occurs.[1] Then $p_{c,x} < p_{d,x}$. This is obviously in line with the expertise provided if $f_{c,x} < f_{d,x}$. We thus recommend to consider a tradeoff as not significant with respect to the selection task if $f_{c,x} < f_{d,x}$. If, however, the opposite inequality holds, i.e. $f_{c,x} \geq f_{d,x}$, then $\frac{w_c}{f_c^*} < \frac{w_d}{f_d^*}$ .This inequality could result from two potential kinds of error. First, one could have wrongly ranked c or d too often second or first respectively in the pairwise comparison procedure. Second, one could have wrongly ranked x too often first with respect to c and too often second with respect to d. If checking these potential errors reproduces the result we consider the tradeoff as OK. Otherwise our method should be completed with the old data and revised new data obtained from repeated pairwise comparisons. If undesirable tradeoffs cannot be avoided and several top-scoring e-Ps exist we recommend using one that has the tradeoffs that seem to be most acceptable.

[1] We restrict to r(x)∈UL since our argument essentially extends to the case r(x)∈LR.

4.6 Sensitivity analysis

Sensitivity analysis indicates consequences of the potentially poor quality of the input data and identifies those quality aspects and e-Ps in terms of which the final ranking is most sensitive to changing data.

The input data, i.e., the pair-wise comparisons both of the e-Ps and the quality aspects are biased by the experts' subjective views and errors they might make. If one presupposes that the procedure, i.e., the AHP, itself is sound then the data quality is left over for assessment. We are suggesting conducting a respective sensitivity analysis and create a sensitivity diagram for each e-P quality aspect. The required calculations are performed by evaluating the formula for p_x by fixing all but one quality aspect and letting the latter vary in small steps in an interval I. We thus obtain the final score p_x for x∈X and c∈C as a function $p_{c,x}$: I → [0,1], with the restriction $p_{c,x}(v) = p_x$ for the score v of x regarding c. In our setting these functions $p_{c,x}$ are all linear. In the sensitivity diagram for quality aspect c we depict the function $p_{c,x}$ for all admissible x∈X. Additionally we indicate with a vertical line the c-score $p_{c,x}$ of a particular x∈X. An example sensitivity diagram for the quality aspect c∈C and four e-Ps is depicted in Figure 2.

Moving the vertical line horizontally signifies changing the c-scores of the admissible e-Ps. Approaching that way an intersection of any of the depicted functions' graphs means that a so-called rank reversal occurs if the quality aspect scores were changed accordingly. Such reversal can be unexpected because the AHP scoring procedure and the system of quality aspects both are complex.

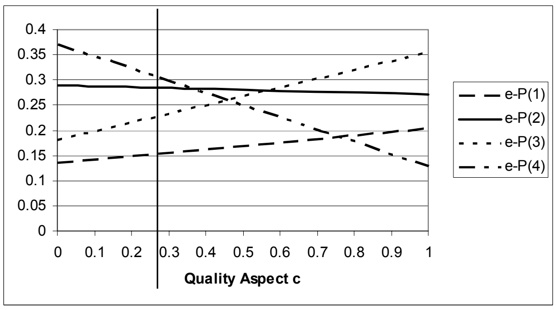

Figure 2: Sensitivity diagram example

As a result of the sensitivity analysis of the whole set of quality aspects and e-Ps (using the set of all possible sensitivity diagrams), the following sensitivity measures

(Triantaphyllou and Sanchez, 1997) can be obtained and presented to the method users:

1. *Quality aspect criticality degree*. This measure, for $i \in \{1,2\}$, is calculated for every quality aspect $c \in C^i$. c's criticality degree is the minimal percentage of change of w(c) for any rank reversal to occur. The criticality degree can be calculated using the distance between the vertical line in the sensitivity diagram and the closest intersection of any e-P functions' graphs. In Figure 2, it is the distance to the intersection of the functions of e-P(2) and e-P(4). If no functions' graphs intersect on the diagram, this measure cannot be calculated, it is the case of *criticality degree infeasibility* indicating that no variations of the quality aspect's weight affect the ranking of e-Ps. Ranking quality aspects according to this degree lets the user observe the differences in their sensitivity to the data variations.

2. *Percent-any critical quality aspect*. This, for $i \in \{1,2\}$, is the quality aspect $c \in C^i$ with the smallest criticality degree among all the quality aspects at level i. This quality aspect can be considered as the most sensitive to the data variations. If no such aspect is found (for all $c \in C$ the criticality degrees are infeasible), the e-Ps ranking can be considered stable.

3. *Percent-top critical quality aspect*. This, for $i \in \{1,2\}$, is the quality aspect $c \in C^i$ with the smallest relative change of w(c) for any rank reversal to occur among the top-scoring e-Ps. This quality aspect can be considered as the most important for our selection task, because the variations of its weight most likely affect the selection outcome. If no such aspect is found, the selection outcome can be considered stable.

4. *E-Process aspect-related criticality degree*. For $x \in X$ and $c \in C$ this is the smallest amount by which w(c) needs to be changed to alter the ranking of x. It measures the influence of w(c) variations on the ranking of x. This criticality degree can be calculated using the distance between the vertical line and the closest intersection of the x's function graph and any other e-P function graph. On Figure 2, the criticality degree of e-P(3) is based on the distance to the intersection of its graph and the graph of e-P(4). This degree can also be infeasible if no variations of w(c) affect the ranking of x.

5. *E-Process criticality degree*. This is the smallest aspect-related criticality degree for the e-P $x \in X$. Ranking e-Ps according to this degree lets the user observe the differences in strength of their AHP-produced ranking.

6. *Most critical e-P*. These are the $x \in X$ with the smallest criticality degree. Each most-critical e-P is among the e-Ps for which the ranking is most unclear and weak. In fact, the ranking of most critical e-Ps should be investigated further. If no most critical e-P can be found (for all $x \in X$ the criticality degrees are infeasible) it also indicates the stability of AHP-produced ranking. If the top-ranking e-Ps all are most critical the whole selection process has likely to be revised because this fact signifies that the outcome of this process is not reliable and sensitive to smallest errors in data.

5 Case-based reasoning

As published in [Albertyn, 2007] the following selection method is based on Case-Based Reasoning. Richter (Richter 1998) has identified four so-called knowledge containers. These are aspects of CBR approaches by means of which knowledge can be represented in the knowledge base. These are (1) the vocabulary used; (2) the similarity measure; (3) the case-base; and (4) the solution transformation. The CBR community has widely accepted these as a natural organisation of knowledge. Since we do not employ the solution transformation we are only left with the first three of these containers. For now we do not see how we additionally to aiming at standardizing the vocabulary used –at the method level- could exploit that vocabulary. In our method we thus focus on the case-base and the similarity measure. We adapt now Richter's knowledge containers to our problem, i.e. e-Process selection.

5.1 Similarity measure

We note that social choice approaches or the AHP could be used for obtaining similarity rankings after a heuristic pre-selection has been conducted. It is, however, common to obtain similarity measures automatically. Computing cost would tend to be too high if one would obtain the similarity to the case at hand of all cases in the case-base. Therefore heuristic rules for pre-selecting cases are in use. However, currently our case-base is still small and also due to the experimental character of our research all similarities will be obtained. According to (Coyle, Doyle et. al. 2004) the similarity between two cases Q and C is defined as the sum of the weighted similarity of the cases' constituent features:

$$Sim(Q,C) = \sum_{f \in F} w_f * \sigma_f(q_f, c_f).$$

In this equation w_f is the constituent feature weight, σ_f the similarity measure applied to feature *f* of Q and C, and *F* the set of all features. The weights are seen as feature attributes. The similarity measures obviously are more complex. (Coyle, Doyle et. al. 2004) use three different kinds of feature similarity measures. These are (1) the exact similarity measure, i.e., the similarity score is 1 if the feature values are equal and is 0 otherwise; (2) difference based similarity measure, i.e., the similarity score depends on the difference of the numerical feature values but not necessarily is 0 for non-equal feature vales; and (3) complex similarities, i.e., all other similarity measures. Using the difference based similarity measure essentially turns CBR into a version of BVA. In our method SCM also plays a role. We think the weaknesses of BVA (i.e. that it often is very hard to score items on a scale according to a number of features) can be overcome by the incorporation of SCM method parts that only rely on ranking items.

In selecting an E-Process both measure kinds (1) or (2) could be used. We are going to use (2), i.e., the difference based similarity measure. Currently we obtain values of weights and feature similarities from the developers who were asked to assess these variables quantitatively. In future we plan to account for the well-known critique of this approach by using SCM or AHP for obtaining scores based on rankings provided by developers. Prior to using these methods we need to finish implementing the software systems dedicated to aid humans in applying these methods.

5.2 Case-base

The quality of e-Processes can be seen as being multi-faceted, as is often experienced for other complex entities such as information systems (Ghezzi, C., Jazyeri, M., et al 2004). Individual quality aspects, however, are often considered as too broad and unspecific and are therefore decomposed into lists of lower level quality aspects. We exploit a two level system of e-Process quality aspects.

Using both the VBA (Albertyn 2005) and the AHP (Albertyn, Kaschek et al. 2007) a case-base has been developed. It has a number of levels that are used in the comparison. First there is information about each of the projects/cases being stored. Then for each of the cases we have group values (high-level quality aspects) that group a number of characteristics (Lower level quality aspects). Then we have weights assigned to each of the characteristics.

Since we focus on CBR as applied for e-Process selection and because similar systems of quality aspects have been used before for similar purposes we do not specifically argue for the quality aspects used in this study. They are, however, loosely based on (Kaschek, R., Pavlov, et. al. 2006) and have been used in (Albertyn 2005, 2007; Albertyn and Kaschek 2004, 2005). Our quality aspects are listed in the appendix and applied to all the presented decision processes.

5.3 Concise presentation of our method

The case-base was developed and implemented during our earlier research on e-Process selection but in order to keep the paper short this is ignored here. That implementation is part of a project to implement a tool-box that aids e-Process selection and which provides a number of respective methods.

To sum up, our method, represented in a step-wise fashion, consists in:

1. Use an SCM for selecting an e-Process most suitable for the case at hand. Verify that selection with inspecting some of the latest cases for which that selection was applied. Take into consideration the case-context information as well as the solution assessment.
2. If the selection made in step 1 can be verified, then go along with it. Otherwise:
   - Use an SCM approach to identify the feature set to be used.
   - Rank the features (according to whether no preference, light preference, moderate preference, strong preference, or extreme preference is given to one feature over another one) in a pair-wise fashion and use the AHP to obtain feature weights.
   - Score the case at hand with respect to the features selected. A light weight method such as an SCM should be used for this.
   - Choose a lower boundary b(Q) for limiting the similarity of case C in the case base to case at hand Q, i.e., b(Q) ≤ Sim(Q,C), for preselecting those cases C that are considered to match best the case at hand Q. Perform that preselection.
   - Use an SCM for identifying the one case in the pre-selection that is considered to be the one most similar to the case at hand and use the e-Process used in that case.
3. Store in the case-base the new case together with the solution as well as a brief solution assessment (obtained post project completion).

That procedure is efficient, as in case expertise emerges the more complex operations are not even executed. Rather the initial choice of the involved staff will be used. In case that expertise is not yet developed enough so that a selection cannot be made initially that would be consistent with past experiences the procedure continues with prescribing more advanced techniques. The procedure also helps acquiring expertise as it aids in recording, organizing and reusing cases of e-Process selection. We consider our method as natural since we pointed out evidence for case-based reasoning being applied in problem solving by many test persons. Varying the boundary b(Q) allows for tuning the cardinality of the preselected set of past cases. So our method is scalable.

6 Conclusions

It is important that good and defendable decisions are made about which e-Process to select when a new eCIS is developed. As seen in this paper there are a number of different decision methods that can be used for this selection process. Three of these decision models were presented in this paper. All three allow for the selection of e-Processes for the development of e-Commerce Information Systems. Future research involves comparing these decision models to clearly define a framework for users to support them when selecting an e-Process. Further research also involves the application of each of these decision models on a selection of case environments.

References

Albertyn, F. (2005). Ontology for the selection of e-Processes. Web Information Systems Engineering - WISE 2005 Workshops, New York, G. Y. Dean M., Jun W., Kaschek R., Krishnawasmy S., Pan Z., Sheng M.

Albertyn, F. and R. Kaschek (2005). eCIS development: refining e-Process selection. IADIS International Conference - e-Society 2005, Qawra, Malta,, P. K. a. M. McPherson 972-8939-03-5.

Albertyn F. (2007) E-Process Selection using Case-Based Reasoning Techniques. IADIS international conference E-Society 2007. Lisbon, Portugal. 3 to 6 July 2007. (http://www.esociety-conf.org/) part of the IADIS Multi Conference on Computer Science and Information Systems (MCCSIS 2007) Lisbon, Portugal, 3 - 8 July 2007 (http://www.mccsis.org) .

Albertyn, F. and R. Kaschek (2004). Ontology based e-Process selection. The Sixth International Conference on Information Integration and Web-based Applications & Services (iiWAS2004), Jakarta, Indonesia, G. Kotsis, S. Bressan, D. Taniar and I. K. Ibrahim3-85403-183-01.

Coyle, L., D. Doyle, et al. (2004). Representing similarity for CBR in XML. Advances in case-based reasoning ECCBR 2004 : European conference on case-based reasoning No7, Madrid, ESPAGNEISBN 3-540-22882-9

Doumpos, M., & Zopounidis, C. (2002). *Multicriteria Decision Aid Classification Methods*. Dordrecht: Kluwer Academic Publishers. 1-4020-0805-8.

Peter Fettke and Peter Loos. Ontological evalution of reference models using the Bunge-Wand-Weber model. in Ninth Americas Conference on Information Systems. 2003.

French, S., & Xu, D.-L. (2005). comparison Study of Multi-attribute Decision Analytic Software. *Journal of Multi-Criteria Decision Analysis, 13*, 65-80.

Ghezzi, C., Jazyeri, M., et al.: *Software qualities and principles*. In: Tucker, A.B.: Computer Science Handbook. Chapman & Hall/CRC (2004)

Kaschek, R., Albertyn, F., Shekhovtsov, V.A., Zlatkin, S. (2007) An e-Process Selection Model. Twenty-Sixth International Conference on Conceptual Modeling - ER 2007 - Tutorials, Posters, Panels and Industrial Contributions, Auckland, New Zealand. 5 - 7 November 2007. Conferences in Research and Practice in Information Technology, Vol. 83. John Grundy, Sven Hartmann, Alberto H. F. Laender, Leszek Maciaszek and John F. Roddick, Eds.

Keeney, R. L. (1992). *Value-Focused Thinking: A path to creative decision making.* Cambridge, M.A.: Harvard University Press

MSNEncarta (2007) from http://encarta.msn.com/encnet/features/dictionary/dictionaryhome.aspx

Oxford_English_Dictionary (2007) from http://dictionary.oed.com

Pareto, V. (1896). *Cours d'Economie Politique*. Lausanne

Opdahl A.L., Henderson-Sellers B., and. Barbier F, Ontological analysis of whole-part relationships in OO-models. Information and Software Technology, 2001. 43(2001): p. 387 - 399.

Pomerol, J.-C., & Barba-Romero, S. (2000). *Multicriterion Decision in Management: Principles and Practice*. MassaChusetts: Kluwer Academic Publishers. 0-7923-7756-7.

Riccia, G. D., Dubois, D., Kruse, R., & Lenz, H.-J. (2003). *Planning based on decision theory*. New York: Springer-Verlag Wien. 3-211-40756-1.

Richter, M. M. (1998). *Introduction - the basic concepts of CBR. Case-based reasoning technology: From foundations to applications.* M. Bartsch-Sporl and S. Wess, Springer-Verlag. LNCS vol. 1400.

Saaty, T. L. (1990). *Multicriteria Decision Making: The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation.* . Pittsburgh: RWS Publications

Trueman, R. E. (1976). *An Introduction to quantitive methods for decision making* (Second edition ed.): Library of Congress Cataloging in Publication Data. 0-03-018391-x.

Unit-For-Sustainable-Development-and-Environment-Organization-of-American-States (2007) Multi-Criteria / Dimensions in Decision-Making from http://www.oas.org/dsd/idsd/methodologies/criteria.htm

Wand Y., Storey V.C., and Weber R., An ontological analysis of the relationship contruct in conceptual modeling. ACM Transactions on Database Systems, 1999. 24(4): p. 494 - 528.

White, D. J. (2006 - Paperback Copyright 1969). *Decision Theory*: Aldine Transaction - Division of Transaction Publishers. 0-202-30898-7.

Wikipedia (2007) from http://en.wiktionary.org/wiki

Appendix 1

- **e-Process characteristics**, i.e. the modelling notions, abstraction concepts, and other aid suggested by the e-Process
  - *Completeness*, i.e. the degree to which the e-Process provides means of expression (such as modelling notions, abstraction concepts, patterns, and anti-patterns) that enable the ECIS developer to effectively and efficiently solve development tasks within the domain of application of the e-Process.
  - *Understandability*, i.e. the degree to which one easily understands the e-Process.
  - *Visibility, i.e. the degree to which the defined activities of the e-Process result in* clear outcomes and effective and efficient project management is enabled.
  - *Supportability*, i.e. the degree to which CASE tools are accessible that aid in using the e-Process.
  - *Maintainability*, i.e. the degree to which the e-Process enables managing requirements in

particular with respect to change requests after project commencement.

- **Specific Quality concepts of the e-Process** (its reliability, robustness etc)
  - *Readability*, i.e. the degree to which the notation prescribed for use in the e-Process is easy to read.
  - *Reliability*, i.e. the degree to which the e-Process is designed in such a way that errors in the development process are avoided or identified and fixed prior to system deployment.
  - *Robustness*, i.e. the degree to which the e-Process continues to aid developers in case of unexpected events occurring.
- **Cost** for using the e-Process
  - *Development Budget*, i.e. the degree to which a tight development budget does not tend to reduce the applicability of the e-Process;
  - *Running Costs*, i.e. the degree to which the ECIS developed with the e-Process will run according to allowed budget.
- **Domain impact**, i.e. the impact of the project domain
  - *Infrastructure*, i.e. the degree to which the technical environment of the enterprise affects the applicability of the e-Process.
  - *Enterprise Culture*, i.e. the degree to which matching the e-Process to the enterprise culture does not reduce the applicability of this e-Process.
  - *Technology*, i.e. the degree to which matching the e-Process to all other technology being used in the enterprise does not reduce the applicability of this e-Process.
  - *Geographic Interaction*, i.e. the degree to which the globalization of the enterprise is necessary for the e-Process being applied.
  - *IT Strategy*, i.e. the degree to which the enterprise's system development standards (specifically ECIS development) do not reduce the applicability of this e-Process.
  - *Business Strategy*, i.e. the degree to which the enterprise's business strategy does not reduce the applicability of this e-Process.
  - *Team Experience*, i.e. the degree to which the involvement of the team with previous ECIS development affects the applicability of this e-Process.
  - *Domain Knowledge*, i.e. the degree to which the domain knowledge of the development team affects the applicability of this e-Process.
  - *E-Process Knowledge*, i.e. the degree to which the knowledge of the e-Process of the development team affects the applicability of this e-Process.
  - *Development Time*, i.e. the degree to which the development time influences the applicability of this e-Process.
- **Usability**, i.e. the e-Process aid in developing a high quality ECIS
  - *Functionality*, i.e. the degree to which the e-Process specifies the required development artefacts (such as requirements, design, implementation, testing, etc.) and instructs how to create and use them.
  - *Manageability*, i.e. the degree to which the e-Process aids managing projects effectively (including planning, tracking, and risk management etc.).
  - *Quality assurance*, i.e. the degree to which the e-Process aids developers in following the principles, requirements, and recommendations of the Total Quality Management announced at ISO 9000:2000, ISO 9001 and ISO 9004.
  - *Adjustability*, i.e. the degree to which the e-Process can be adjusted to meet the specific needs of the ECIS project in the most effective way possible.
- **Compatibility** of the e-Process with other methodologies
  - *Exchangeability*, i.e. the degree to which the artefacts of the given e-Process can be exchanged between the tools implementing this e-Process.
  - *Map ability*, i.e. the degree to which the artefacts of the given e-Process can be mapped into the artefacts of other e-Processes.
- **Maturity**, i.e. e-Process stability, tool support, documentation etc.
  - *Stability*, i.e. the degree to which the e-Process has been proven, i.e., its standing.
  - *Tool support*, i.e. the degree of availability of the tools supporting this e-Process (such as version control and document/workflow management systems) and the quality of these tools.
  - *Documentation*, i.e. the degree to which adequate documentation is enabled.