Handling Uncertain and Qualitative Information in Impact Assessment – Applications of IDS in Policy Making Support

Dong-Ling Xu[1], Jian-Bo Yang[1,2], and Xinbao Liu[2]

[1] Manchester Business School, The University of Manchester, Manchester M15 6PB, UK
L.Xu@mbs.ac.uk, Jian-Bo.Yang@mbs.ac.uk

[2] School of Management, Hefei University of Technology, Hefei, Anhui 230009, China
lxinbao@mail.hf.ah.cn

Summary. Impact assessment (IA) in policy making processes has received increasing attention in recent years. One of the major challenges in IA is how to rationally handle and make maximum use of information in uncertain and qualitative data so that the best course of action can be reliably identified. It is discussed in this chapter how the Evidential Reasoning (ER) approach for multiple criteria decision analysis (MCDA) can be used to take the challenge. The ER approach and its software implementation, called the Intelligent Decision System (IDS), are developed with a focus on rationally handling a large amount of information of both a qualitative and quantitative nature and possibly with different degrees of uncertainties in assessment problems. It applies belief decision matrices for problem modelling so that different formats of available data and uncertain knowledge can be incorporated into assessment processes. It uses an evidential reasoning process on the data to generate assessment outcomes that are informative, rational and reliable. Several examples are examined to demonstrate how IDS can be used to support activities in different stages of an IA process, namely (a) problem structuring, (b) assessment model building, including value elicitation, (c) data collection, management, and aggregation, and (d) data presentation and sensitivity analysis. This investigation shows that IDS is not only a versatile assessment supporting tool, but also a knowledge management tool which helps to organise assessment knowledge and data systematically for better traceability, consistency and efficiency in assessment.

1 Introduction

Policies and regulations affect many people or businesses in many ways. To enable better policy making, impact assessment (IA), a process of identifying the future consequences of a current or proposed action, has received increasing attention in recent years among OECD countries (Organisation for Economic Co-operation and Development) [10, 17].

D.-L. Xu et al.: *Handling Uncertain and Qualitative Information in Impact Assessment – Applications of IDS in Policy Making Support*, Studies in Computational Intelligence (SCI) **117**, 159–185 (2008)
www.springerlink.com

Initially IA was focused on whether regulations would impose an unnecessary burden on the private, public or third sectors. It was essentially an economic cost benefit analysis tool. Realising that an assessment may not be complete without properly taking into account all factors in question, over recent years, a number of countries have begun to establish new forms of integrated IA that include the assessment of unintended, long-term or non-market effects and inter linkages between different issues of concern. For example in the UK, IA has been expanded to include the consideration of social, environmental and economic impacts [4, 10] and is becoming more complicated.

To further add to the complication, various types of uncertainty may exist in data collected for IA, such as probability due to random events and factors, imprecise estimates for long term effects, vagueness in subjective judgements, and incomplete data sets due to unknown or missing parts of facts. How to rationally incorporate qualitative criteria and uncertain knowledge in an assessment poses a major challenge to both IA practitioners and researchers in the field of multiple criteria decision analysis (MCDA).

To cater for the needs of handling the increasing complexity and difficulty in IA, MCDA approaches have been introduced and applied in IA as reported in numerous literatures. [16, 25, 47]. In this chapter, it is illustrated by examples how the recently developed approach, the Evidential Reasoning approach in MCDA, and its software implementation, Intelligent Decision System (IDS) [40,43,44] can be applied to support IA and what are its advantages and limitations.

Generally, there are four stages in an IA process. The first two stages are concerned with the modelling process of an assessment problem, which are relatively independent of individual policy options to be assessed. The other two stages are mainly specific to individual policies. In practice, it may be necessary to go through some of the stages a number of times in order to refine the assessment model and clarify some of the uncertainties in the assessments of alternative options. The four stages are summarised as follows.

The first stage is to define and construct an assessment problem. At this stage, the following questions need to be addressed. What are the scopes of the assessment? What are the alternative options? In what areas or on which criteria the performances of the options need to be assessed?

The second stage is to establish an assessment framework or model by asking the following questions. How should the performance of each option in each area be measured? Are better performances in some areas more important than in others? If so, how to elicit the relative importance of each area or criterion? How uncertainties in assessments can be clarified and recorded for further analysis?

The third stage is data collection and handling. At this stage data from different sources are collected in order to rate the performances of each option in the concerned areas. The data may be of heterogeneous nature, and their quality may vary. Hence potential problems in this stage are how to manage data from different sources and extract quality information from the data, how to handle uncertainties in the data, and how to aggregate the data to arrive at reliable and rational assessment outcomes.

The fourth stage is the interpretation of the assessment outcomes and the following questions may be asked. Are the outcomes convincing? Have they included all aspects and taken into account all opinions of different stakeholders? Are the outcomes explainable and can they be traced back to their sources? What are the effects of any uncertainty in data and subjective judgements? How can the outcomes, the effects of any uncertainties, and their traceability be clearly presented to stakeholders?

In this chapter, it is described how IDS can support IA in each of the four stages. It is arranged as follows. In the next section, the ER approach and the IDS software are briefly outlined. The processes of using IDS to support IA in its four stages are then illustrated using examples. The features and advantages of the ER approach are discussed in the concluding remarks.

A few points should be noted while reading this chapter.

- In MCDA, attribute and criterion are often used interchangeably. It is also the case in this chapter.
- The following section on the ER approach may be skipped for readers who are not interested in the technical details of the approach. To apply the approach, the IDS software provides friendly interfaces for users to construct assessment models, record assessment data and carry out necessary calculations.

2 The ER Approach and IDS

MCDA is a branch of operational research concerned with making assessments and choices when there are several alternatives, and when each alternative has merits as well as drawbacks. Over its short history of over 30 years, along with the advancement of computer technology, many approaches have been developed to support systematic analysis of complex MCDA problems [2]. One of the major challenges in the MCDA is how to rationally handle uncertain knowledge including qualitative factors [5,31,38,39]. Without properly taking all relevant attributes or criteria into account, an assessment is incomplete and the outcome may be biased [12, 21, 22, 29, 42].

Over the past two decades, considerable research has been conducted on integrating techniques from artificial intelligence and operational research for handling uncertain information [1,3,6,45]. Along this line of research, the ER approach and IDS software are developed in response to the growing needs to develop scientific methods and tools for dealing with MCDA problems under uncertainty in a way that is rational, reliable, repeatable, and transparent. The ER approach uses concepts from several disciplines, including decision sciences in particular utility theory [14], artificial intelligence in particular the theory of evidence [26] statistical analysis and computer technology [41,

Table 1. Decision matrix

	Attribute 1	...	Attribute l	...	Attribute L
Alternative 1	$S(A_1(O_1)) = H_3$		$S(A_l(O_1))$		$S(A_L(O_1))$
...					
Alternative m	$S(A_1(O_m))$		$S(A_l(O_m))$		$S(A_m(O_L))$
...					
Alternative M	$S(A_1(O_M))$		$S(A_l(O_M))$		$S(A_M(O_L))$

Table 2. Belief decision matrix

	Attribute 1	...	Attribute l	...	Attribute L
Alternative 1	$S(A_1(O_1))$ $=\{(H_1, \beta_{l,1}), \ldots$ $(H_N, \beta_{l,N})\}$		$S(A_l(O_1))$		$S(A_L(O_1))$
...					
Alternative m	$S(A_1(O_m))$		$S(A_l(O_m))$		$S(A_m(O_L))$
...					
Alternative M	$S(A_1(O_M))$		$S(A_l(O_M))$		$S(A_M(O_L))$

42, 46]. Compared with conventional MCDA methods, in the ER approach a MCDA problem is modelled using a belief decision matrix [11, 43], of which the conventional decision matrix [9], as indicatively shown in Table 1, is a special case.

2.1 MCDA Problem Modelling Using Belief Decision Matrix

In a belief decision matrix, the performance of an assessed option on a criterion is represented by a distribution instead of a single value, as indicated in Table 2. For example, some business people believe that if UK joins the Euro, there will be less uncertainty in their business planning because the uncertainty associated with the fluctuation of exchange rates between pound sterling and the Euro is no longer an issue. However, for businesses whose customers and suppliers are either in the UK or other countries outside Euro zone, there will be no differences. If people are asked to rate the impact of UK Euro membership on "Stability for business planning", it is unlikely to get a unanimous answer. Suppose we use the following five grades to rate the impact

- H_1: Very negative
- H_2: Negative
- H_3: Neutral
- H_4: Positive
- H_5: Very positive

and 70% of the responses are Positive and 30% Neutral, then the assessment (or a piece of performance evidence) should be expressed as

$$S(A_1(O_1)) = \{(\beta_{1,1}, H_1), (\beta_{2,1}, H_2), (\beta_{3,1}, H_3), (\beta_{4,1}, H_4), (\beta_{5,1}, H_5)\} \\ = \{(0, H_1), (0, H_2), (0.7, H_3), (0.3, H_4), (0, H_5)\} \quad (1)$$

Equation (1) is referred to as a distributed assessment or simply a distribution where O_1 denotes option 1 (UK to join the Euro), A_1 criterion 1 (Stability for business planning), and $S(A_1(O_1))$ the performance of O_1 on A_1. $0 \leq \beta_{n,1} \leq 1$ ($n = 1, \ldots, 5$) denotes the degree of belief that the alternative O_1 is assessed on A_1 to the grade H_n. $S(A_1(O_1))$ reads that O_1 is assessed to the grade H_n to a degree of $\beta_{n,1}$ on the criterion A_1 ($n = 1, \ldots, 5$), or the option "Joining the Euro" is assessed to be Positive on "Stability for business planning" to degree of 30% and neutral 70%.

Using decision matrix, the performance information shown in (1) needs to be approximated by a single value, such as "Positive", while in belief decision matrix, each element can be a distribution and it accepts the distributed performance information as it is without approximation.

Further more, if there is missing information in data, it can be represented by a distribution without either adding new or taking away existing information from the data. For example, suppose the responses in the above example are 25% Positive, 60% Neutral and 15% no answers given. Normally either the missing answers need to be replaced by some estimates or the responses with missing answers are discarded, including the answers to other questions. Either way, information in data may have been distorted. Using a distribution, the information in data can be maintained by expressing the assessment as

$$S(A_1(O_1)) = \{(\beta_{1,1}, H_1), (\beta_{2,1}, H_2), (\beta_{3,1}, H_3), (\beta_{4,1}, H_4), (\beta_{5,1}, H_5)\} \\ = \{(0, H_1), (0, H_2), (0.6, H_3), (0.25, H_4), (0, H_5)\}$$

Note that in the above equation, $\sum_{n=1}^{5} \beta_{n,1} = 0.85 \leq 1$. Generally, there must be $\sum_{n=1}^{5} \beta_{n,1} \leq 1$ and if $\sum_{n=1}^{5} \beta_{n,1} < 1$, then the assessment $S(A_1(O_1))$ is considered to be incomplete. Obviously, if $\sum_{n=1}^{5} \beta_{n,1} = 1$ then the assessment is complete. In the ER framework, both complete and incomplete assessments can be accommodated [40].

More generally, if an assessment problem has L attributes A_i ($i = 1, \ldots, L$), M options O_j ($j = 1, \ldots, M$) and using N evaluation grades H_n ($n = 1, \ldots, N$) to assess the options on each attribute, then a matrix can be constructed with $S(A_i(O_j))$ as its element in the ith row and jth column where $S(A_i(O_j))$ is given as follows:

$$S(A_i(O_j)) = \{(H_n, \beta_{n,i}(O_j)), \quad n = 1, \ldots, N\} \\ i = 1, \ldots, L, \quad j = 1, \ldots, M \quad (2)$$

This matrix is called belief decision matrix (Table 2), in contrast to the normal decision matrix (Table 1). It should be noted that a performance on a criterion can be measured using numerical values or a set of evaluation grades. It should also be noted that different grade sets, possibly with different number of grades in them, may be used for assessing different attributes [40].

It is commonly known that different attributes may play different roles in an assessment and their importance is represented by attribute weights. Suppose ω_i is the weight of attribute A_i $(i = 1, \ldots, L)$. Because weights represent the relative importance of attributes, they can be scaled (or normalised). In the ER approach, the normalisation is such that $0 \leq \omega_i \leq 1$ and $\sum_{i=1}^{L} \omega_i = 1$.

2.2 ER Approach for Information Aggregation

Instead of aggregating average scores, the ER approach employs an evidential reasoning algorithm [40–42] developed on the basis of the evidence combination rule of the Dempster–Shafer theory [26] to aggregate belief degrees in performance distributions. The outcome of the aggregation is also a distribution, not a single score.

Without loss of generality and for illustration purpose, the ER algorithm is presented below by assuming that the performance of an alternative option is decided by its performances on two criteria A_1 and A_2. Detailed descriptions and the properties of the aggregation process can be found in [43, 44].

Suppose the performance on criterion A_1 is given by (1) and on A_2 by

$$\begin{aligned} S(A_2(O_1)) &= \{(\beta_{1,2}, H_1), (\beta_{2,2}, H_2), (\beta_{3,2}, H_3), (\beta_{4,2}, H_4), (\beta_{5,2}, H_5)\} \\ &= \{(0, H_1), (0.5, H_2), (0.5, H_3), (0, H_4), (0, H_5)\} \end{aligned} \tag{3}$$

Further suppose the normalised weights of A_1 and A_2 are $\omega_1 = 0.4$ and $\omega_2 = 0.6$ respectively. The problem is to aggregate the two assessments $S(A_1(O_1))$ and $S(A_2(O_1))$ to generate a combined assessment $S(A_1(O_1)) \oplus S(A_2(O_1))$. In the example $S(A_1(O_1))$ and $S(A_2(O_1))$ are both complete. If not, the rational handling of the unknown portion of its performances is to assume that the missing portion of the performance can be rated to any grade from H_1 to H_5. The details of the ER algorithm for the example is given below and its more generic format capable of aggregating both complete and incomplete assessments is described in [41] and [40]. Let

$$p_n = \omega_1 \beta_{n,1} \ (n = 1, \ldots, 5) \text{ and } p_H = 1 - \omega_1 \sum_{n=1}^{5} \beta_{n,1} = 1 - \omega_1 = 0.6 \tag{4}$$

$$q_n = \omega_2 \beta_{n,2} \ (n = 1, \ldots, 5) \text{ and } q_H = 1 - \omega_2 \sum_{n=1}^{5} \beta_{n,2} = 1 - \omega_2 = 0.4 \tag{5}$$

where each p_n or q_n $(n = 1, \ldots, 5)$ is referred to as basic probability mass, and p_H and q_H are the remaining probability mass unassigned to any of the grade H_n $(n = 1,\ 2,\ 3,\ 4,\ 5)$. Their values are given in the 1st row and 1st column of Table 3.

The ER algorithm is used to aggregate the basic probability masses to generate combined probability masses, denoted by r_n $(n = 1, \ldots, 5)$ and r_H using the following equations:

Table 3. Probability masses

	$S(A_1(O_1))\oplus S(A_2(O_1))$	$S(A_1(O_1))$: $p_1 = 0$ $\{H_1\}$	$p_2 = 0$ $\{H_2\}$	$p_3 = 0.28$ $\{H_3\}$	$p_4 = 0.12$ $\{H_4\}$	$p_5 = 0$ $\{H_5\}$	$p_H = 0.6$ $\{H\}$
$S(A_2(O_1))$	$q_1 = 0$ $\{H_1\}$	$p_1q_1 = 0$ $\{H_1\}$	$p_2m_1 = 0$ $\{\Phi\}$	$p_3q_1 = 0$ $\{\Phi\}$	$p_4q_1 = 0$ $\{\Phi\}$	$p_5q_1 = 0$ $\{\Phi\}$	$p_Hq_1 = 0$ $\{H_1\}$
	$q_2 = 0.3$ $\{H_2\}$	$p_1q_2 = 0$ $\{\Phi\}$	$p_2q_2 = 0$ $\{H_2\}$	$p_3q_2 = 0.084$ $\{\Phi\}$	$p_4q_2 = 0.036$ $\{\Phi\}$	$p_5q_2 = 0$ $\{\Phi\}$	$p_Hq_2 = 0.18$ $\{H_2\}$
	$q_3 = 0.3$ $\{H_3\}$	$p_1q_3 = 0$ $\{\Phi\}$	$p_2q_3 = 0$ $\{\Phi\}$	$p_3q_3 = 0.084$ $\{H_3\}$	$p_4q_3 = 0.036$ $\{\Phi\}$	$p_5q_3 = 0$ $\{\Phi\}$	$p_Hq_3 = 0.18$ $\{H_3\}$
	$q_4 = 0$ $\{H_4\}$	$p_1q_4 = 0$ $\{\Phi\}$	$p_2q_4 = 0$ $\{\Phi\}$	$p_3q_4 = 0$ $\{\Phi\}$	$p_4q_4 = 0$ $\{H_4\}$	$p_5q_4 = 0$ $\{\Phi\}$	$p_Hq_4 = 0$ $\{H_4\}$
	$q_5 = 0$ $\{H_5\}$	$p_1q_5 = 0$ $\{\Phi\}$	$p_2q_5 = 0$ $\{\Phi\}$	$p_3q_5 = 0$ $\{\Phi\}$	$p_4q_5 = 0$ $\{\Phi\}$	$p_5q_5 = 0$ $\{H_5\}$	$p_Hq_5 = 0$ $\{H_5\}$
	$q_H = 0.4$ $\{H\}$	$p_1q_\mathrm{H} = 0$ $\{H_1\}$	$p_2q_\mathrm{H} = 0$ $\{H_2\}$	$p_3q_\mathrm{H} = 0.112$ $\{H_3\}$	$p_4q_\mathrm{H} = 0.048$ $\{H_4\}$	$p_5q_\mathrm{H} = 0$ $\{H_5\}$	$p_Hq_\mathrm{H} = 0.24$ $\{H\}$

$$r_n = k(p_nq_n + p_Hq_n + p_nq_H),\ (n = 1, \dots, 5) \tag{6}$$

$$r_H = k(p_Hq_H) \tag{7}$$

where

$$k = \left(1 - \sum_{t=1}^{5}\sum_{\substack{n=1\\n\neq t}}^{5} p_tq_n\right)^{-1} \tag{8}$$

From Table 3, we have

$$k = (1 - (0.084 + 0.036 + 0.036))^{-1} = 0.844^{-1} = 1.185,$$
$$r_1 = 0,\ r_2 = k \times p_Hq_2 = 1.185 \times 0.18 = 0.213,$$
$$r_3 = k \times (p_3q_3 + p_Hq_3 + p_3q_H) = 1.185 \times (0.084 + 0.18 + 0.112) = 0.446$$
$$r_4 = k(p_4q_H) = 1.185 \times 0.048 = 0.057,\ r_H = k(p_Hq_H) = 1.185 \times 0.24 = 0.284$$

If there are more than two criteria, the combined probability masses can then be aggregated with the third assessment in the same way. The process is repeated until all assessments are aggregated. The combined probability masses are independent of the order in which individual assessments are aggregated. If there are several levels of criteria in a hierarchy, the aggregation process is carried out from the bottom level criteria one branch at a time until the top of the hierarchy is reached. The belief degrees in the aggregated performance distribution are calculated from the combined probability masses. Suppose the final combined assessment for the option O_1 is represented as follows:

$$S(O_1) = \{(H_1, \beta_1), (H_2, \beta_2), (H_3, \beta_3), (H_4, \beta_4), (H_5, \beta_5)\} \tag{9}$$

where β_n $(n = 1, \ldots, 5)$ are the combined belief degrees generated by:

$$\beta_n = \frac{r_n}{1 - r_H}(n = 1, \ldots, 5) \tag{10}$$

For the example, we have $\beta_1 = 0$, $\beta_2 = \frac{r_2}{1-r_H} = \frac{0.213}{1-0.284} = 0.297$, $\beta_3 = 0.623$, $\beta_4 = 0.080$, and $\beta_5 = 0$.

2.3 Expected Utility Scores

If necessary a score can be calculated from the distribution. Before the calculation, a utility value needs to be assigned to each grade to represent the preference of policy makers towards the grade [14]. For example, suppose the utilities for the five grades in (1) are as follows:

$$u(H_1) = 0,\ u(H_2) = 0.25,\ u(H_3) = 0.5,\ u(H_4) = 0.75,\ \text{and } u(H_5) = 1$$

An expected utility score for O_1, denoted by $u(O_1)$, can be calculated as follows with the belief degrees as weights,

$$u(O_1) = \sum_{i=1}^{5} u(H_i)\beta_i = 0.45 \tag{11}$$

It should be noted that the ER aggregation is in essence a statistical and non-linear approach, which reinforces harmonic judgements and weaken conflict ones [44].

2.4 Applying the ER Approach through IDS

As we can see from the example, ER approach involves handling data in a structured way and without computer support it is difficult to be applied manually. To facilitate its easy application, IDS[1] is developed to transform the model building and result analysis processes into an easy window-based click and design activity. It aims to provide not only technical supports in data processing including data collection, storing, retrieving and presentation, but also cognitive supports in problem structuring and assessment process. The rest of the chapter is devoted to demonstrating the application of IDS in each of the four stages of an IA process.

3 IDS and Its Applications in Impact Assessment Support

IDS is a Windows based software tool based on the ER approach. During the past few years, it has been applied to support assessment activities in different areas. Example of such applications include supplier assessment in

[1] A free demo version of IDS can be obtained from Prof J B Yang via email: jian-bo.yang@mbs.ac.uk or www.e-ids.co.uk

procurement [28,37], market performance assessment and consumer preference identification in new product design [7], business performance assessment and organisational self-assessment in total quality management [8, 19, 27, 35], customer satisfaction survey [37] in customer relationship management, impact assessment in policy making [33, 34], and risk assessment in engineering design [15]. The results show that the ER approach, supported by IDS, has significantly helped to improve consistency, transparency and objectivity in the assessments.

In the following discussion, the impact assessment of UK Euro membership is used as an example to illustrate the application of the IDS in each of the four stages of an IA as outlined in Sect. 1, namely problem structuring, establishing an assessment model, data collection and handling, and interpretation of outcomes.

3.1 Problem Structuring

In the problem structuring phase, stakeholders, an initial set of alternatives, key issues, constraints, and uncertainties need to be identified.

There are many qualitative frameworks for problem structuring. Many soft operational research techniques can be used. The value focused thinking [13] is also an excellent and well accepted approach for generating new alternatives creatively. Post-It is often used for capturing and organising ideas. Belton and Stewart [2] provide a comprehensive summary on approaches for problem structuring. The CAUSE framework is one of them. The acronym CAUSE stands for

- C – identifying Criteria. Criteria should be measurable and understandable, cover all aspects of concern to decision makers, and should not have redundancy
- A – identifying Alternatives
- U – identifying Uncertainties
- S – identifying Stakeholders
- E – identifying Environmental factors and constraints

In the UK Euro membership problem, there are two natural alternatives: either join or not join. It is important that opinions from both pro- and anti-Euro sides are taken into account so that a balanced assessment can be made. A quick search of the Internet can lead to many sites discussing the gains and losses of UK joining the Euro in various aspects. Those aspects form the basic sets of assessment criteria for the problem.

Generally, in an assessment problem, alternative options are assessed by many criteria and sub criteria. If the sub criteria are still too general and abstract to be measured, they should be broken down further until they are measurable. The process leads to the formation of a criterion hierarchy. IDS provides user friendly interfaces to document the alternatives and construct the criterion hierarchy.

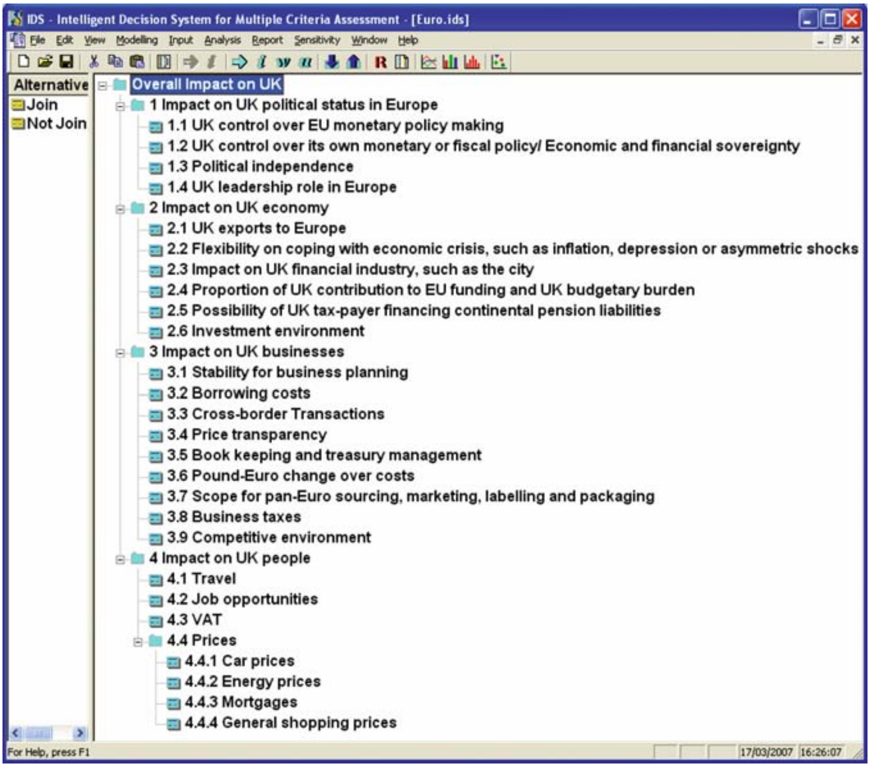

Fig. 1. Support of problem structuring: assessing UK Euro membership

In its main window (Fig. 1), there are two panes, the left is for listing the alternative options (or simply alternatives), and the right for listing a criterion hierarchy. New alternatives can be added by clicking on the left pane once and then the yellow arrow button  on the Toolbar of the main window. The alternatives can be renamed, and described with more details if necessary by right clicking on it once. New criteria can be added at any position by clicking at the desired position and then the blue arrow button  (Fig. 1). The newly added criteria can also be renamed and defined with more details. For example, from searching the Internet, the impacts of UK Euro membership are mainly on the following four areas: Political, Economy, Business and People. Under each category, there are more detailed sub areas which are treated as sub criteria and the criterion hierarchy can be built using the IDS as shown in Fig. 1. IDS also provides the facility to delete, copy and paste criteria and alternatives if necessary.

3.2 Assessment Model Building

Having identified the options and the assessment criteria, and implemented the criterion hierarchy in IDS, in the second phase, we need to address the following three issues and build the assessment model accordingly; (a) how the performance of each option can be measured on each criterion, (b) what weights should be assigned to each criterion so that its relative importance can be represented and (c), what is the preference or risk attitude of policy makers towards each assessment grade or value in the measurement scale of each criterion. Those three elements together with the criterion hierarchy built earlier constitute an assessment model which is used for assessing all the policy options in an IA problem. The three issues are discussed in turn in the following sub-sections.

Assessment Criterion Definition

Issue (i) is concerned with how performances can be measured on each criterion. The simplest cases are when the performance of each option can be measured numerically on a criterion without uncertainty, such as the pound and euro changeover costs if it can be estimated more or less accurately. It is more complicated for other cases. If qualitative judgements are unavoidable, there is an issue of how to reduce subjectivity and increase consistency in the assessment. If the performances are associated with certain random factors, the issue is then how to clarify and represent the uncertainty in the model so that the risks associated with the uncertainty can be revealed and examined.

On qualitative criteria, the performance of each option is commonly assessed by grades. For example, the impact of UK Euro membership in many areas can only be measured qualitatively and a frequently used set of measurement grades are:

- Very negative
- Negative
- Neutral
- Positive
- Very positive

One problem with qualitative grades is that the meaning of a grade may mean different standards for different people. To reduce subjectivity, it is also a common practice to clearly define the standards of all grades.

For a quantitative criterion with probability uncertainty, traditionally the expected or mean value is used to represent the performance of an option on the criterion. This, however, introduces information losses. Ideally the probability distribution of a performance should be preserved and the associated risks be explicitly explored in an assessment process.

The IDS software is designed with a focus on supporting the model building process of IA problems with both qualitative and quantitative criteria

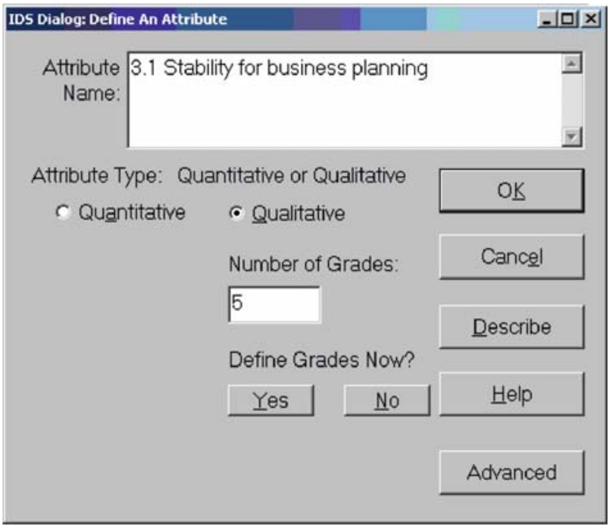

Fig. 2. Define qualitative criteria

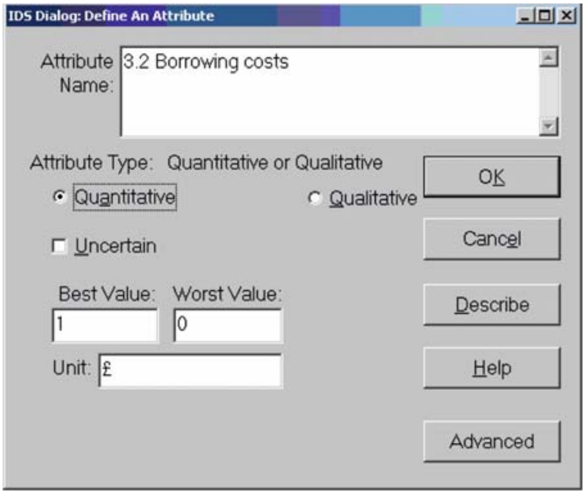

Fig. 3. Define quantitative criteria

under various types of uncertainties. It starts by prompting users to classify a criterion into one of the three logical categories: qualitative, quantitative (without uncertainty) and quantitative with uncertainty (Figs. 2 and 3).

For a qualitative criterion, further interfaces are provided for users to define assessment grades, their corresponding standards and utilities (Figs. 4, 5 and 8). Late on, at the data collection and handling stage, when the performance of an alternative on this very criterion is assessed and rated, the

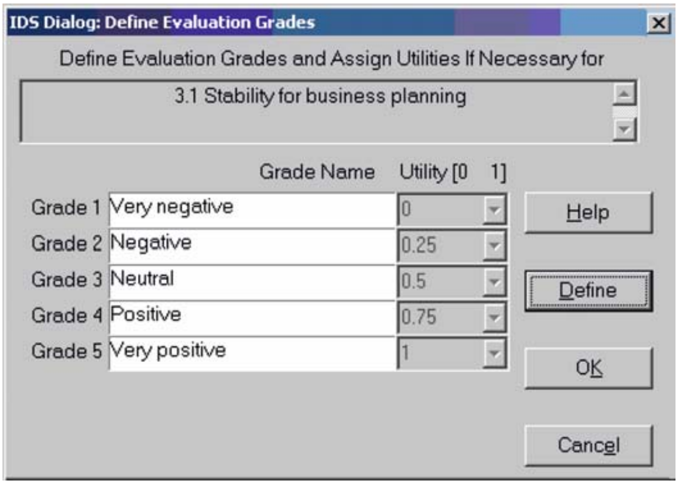

Fig. 4. Define assessment grades

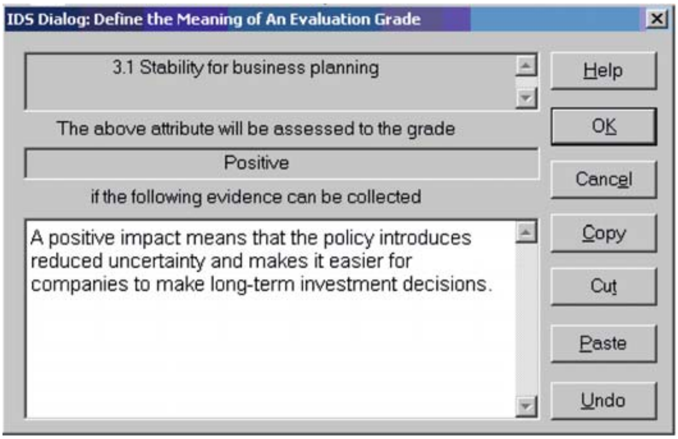

Fig. 5. Define assessment grade standards

grading standards defined here will be conveniently accessible so that users can make a reference to it to ensure the consistency of the assessment.

For a quantitative criterion without uncertainty, IDS prompts users to identify the performance variation range of alternative options on it (Fig. 3), and the preferences of policy makers towards the different performances. If the performance of any alternative on the criterion is anticipated to be a probability distribution, then the "Uncertain" box (Fig. 3) should be checked and later in the data collection stage users will have the flexibility to record the performance of an alternative on the criterion using a distribution.

Relative Importance of Criteria and Weight Elicitation

Issue (ii) is concerned with the role each criterion can play in an assessment or its weight assignment. The assignment process involves significant subjective judgements and need to be supported in order to get a satisfactory set of weights.

In IDS, there are a couple of interfaces dedicated to support criterion weight elicitation. The first one is the visual assignment window (Fig. 6). From this window, a number of methods can be used for eliciting and recording the weights through an interactive graph. One is the direct assignment method [20] and is used when policy makers have more or less decided what weight to give to each criterion. The second one is the SMART (Simple Multi-Attribute Rating Technique) [30] method, which assigns 10 points to the least important criterion and more than 10 to the second least important criterion and so on. The third one is the SWING method [30], which is somehow opposite to SMART. It gives 100 points to the most important criterion and less than 100 points to other criteria. To apply any of the three methods in IDS, users need only to drag and drop each bar in the interactive graph to an appropriate height.

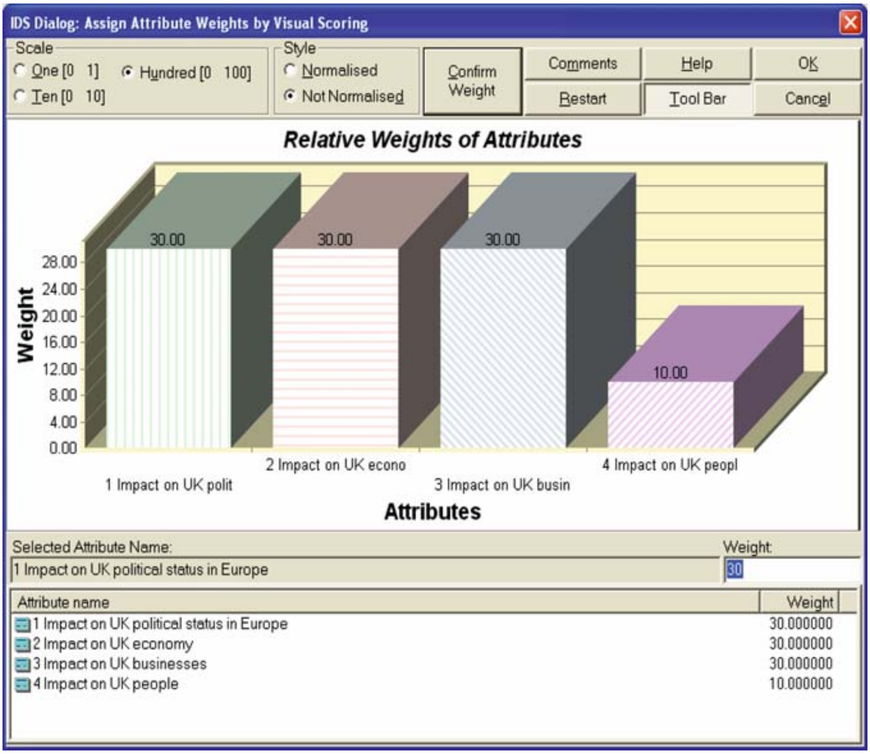

Fig. 6. Weight assignment by interactive graph

The second interface is for supporting weight assignments using paired comparison. It considers only two criteria at a time. This is a frequently used method due to the simplicity of the idea, even though the derived process is quite tedious. From the interface, the comparisons can be carried out between either all possible pair combinations, or one criterion and each of the others ($n-1$ pairs if the number of criteria in consideration is n) [22, 24]. Once the comparisons are finished, the set of weights best fit the comparisons is then calculated and any inconsistency noted by IDS.

When there are multiple stakeholders, and a consensus set of weights can not be achieved, average weights or weight intervals given by members may be used. The intervals of weights can then be used to guide the sensitivity analyses in the next phase for weight fine-tuning (Fig. 7).

Elicitation of Preference of Policy Makers

Issue (iii) is concerned with the preference or risk attitude measurement of policy makers towards the performances of an alternative on each criterion. The measurement is accomplished by using a common scale, normally between 0 and 1 with 0 corresponding to the least and 1 the most preferred levels of a performance respectively. Such a common scale is referred to as utility function in decision theory [14]. For example, the impact of UK's Euro membership on

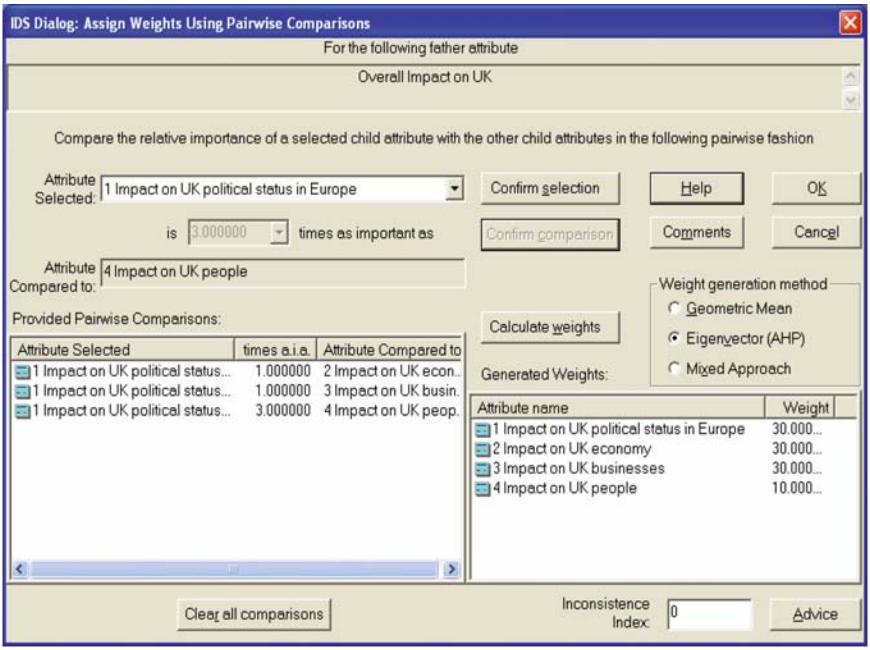

Fig. 7. Weight assignment by pairwise comparison

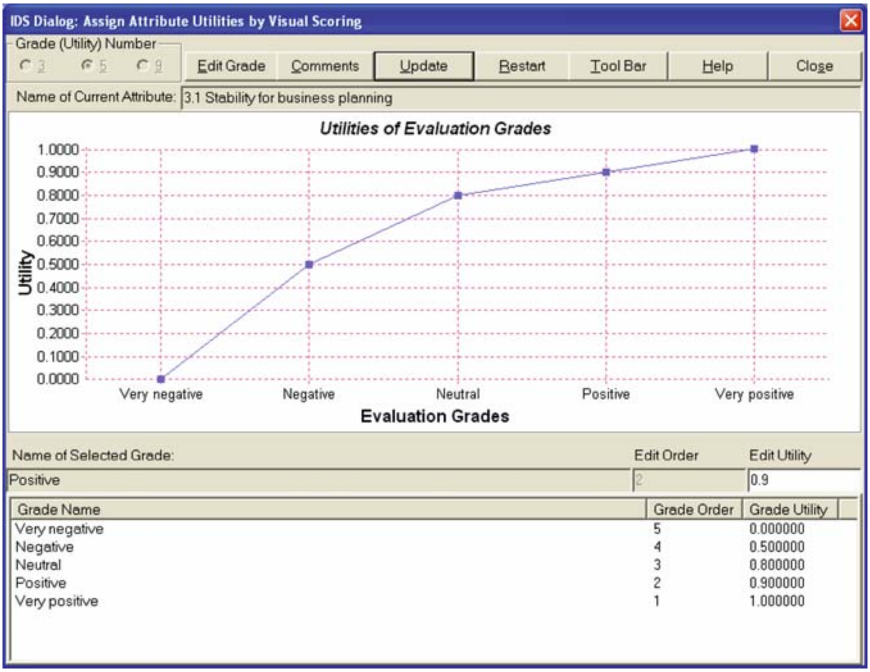

Fig. 8. Interfaces for defining utility functions

Stability for Business Planning is measured by the following 5 grades: Very Negative, Negative, Neutral, Positive and Very Positive. If the policy makers assign utility of 0, 0.5, 0.8, 0.9 and 1 to each of the five grades respectively, the utility function for this criterion may look like the curve shown in Fig. 8. If the policy makers wish to assign different utilities to the grades, it is supported in the IDS by an interactive interface (Fig. 8) where users can drag and drop the points on the curve to a desired position.

As indicated by (11), from utility functions and the performance distributions of alternative options, scores can be calculated and ranking can be generated based on the scores. Therefore one of the purposes of utility functions is to facilitate the comparability of alternatives on each criterion at any level of the hierarchy. Through the use of utility functions, alternative options can be assessed on each criterion using its own most appropriate scale first and then the assessments are transformed to the common scale. IDS has such information transformation procedures [40] built-in to ensure that, although different assessment grades are used, policy makers' preferences are equivalently preserved in the transformation processes and properly presented in the aggregated outcomes.

3.3 Data Collection and Data Handling

Having established and implemented an assessment model using the IDS software, our attention can now turn to assessing individual policies. To assess the impact of a policy in each area (or on each criterion), data need to be collected from different sources, including looking at historical data and seeking expert opinions on the potential costs and benefits, tangible or intangible, of implementing and enforcing a policy. There are inevitably uncertainties in the estimates and judgements. IDS provides a number of interfaces to support data collection and input processes. The aim of the supports are to help improve consistency in judgements, clarify and reduce uncertainties in assessments, and manage the data colleted.

There are three different interfaces for data input in IDS, each for one of the three types of criteria as discussed in Sect. 3.2: quantitative (without uncertainty), quantitative with uncertainty, and qualitative.

Entering assessment data on quantitative type of criteria is straightforward therefore it is not discussed further. If there is uncertainty in quantitative assessments, they can normally be represented as probability distributions. For example, suppose the "Pound-Euro changeover costs" if UK adopts the Euro are estimated to be 3, 3.5 and 4 billion pounds sterling with probability of 30, 50 and 20% respectively. IDS then provides both interfaces to accept the information as it is and an algorithm to properly aggregate the information in the data so that the effects associated with the uncertainties can be revealed in the outcomes.

For qualitative type of criteria, the support to reduce subjectivity in assessments is from two fronts. One is the provision of an evidence mapping interface (Fig. 9). It displays the assessment standards, as defined earlier in the assessment framework (Fig. 5), and the related evidence and judgements, collected and entered by users at the current stage, side by side so that the comparison of a performance against the standards are made easier. In this way the assessments made by different assessors are geared to follow the same standards and improved consistency can be achieved. On the second front, if a performance matches a mixed grade standards, users have the flexibility to assign portions of the performance to a number of grades using belief degrees as discussed in Sect. 2.1 (Fig. 10). In this way, the assessment can be made more objective and accurate.

The supports from IDS also include the structured recording of the assessment knowledge and performance evidences for traceability and future references. From the data and the recorded knowledge, an assessment report for each policy option can be generated automatically. This can further save time, and improve accuracy and efficiency in report preparation.

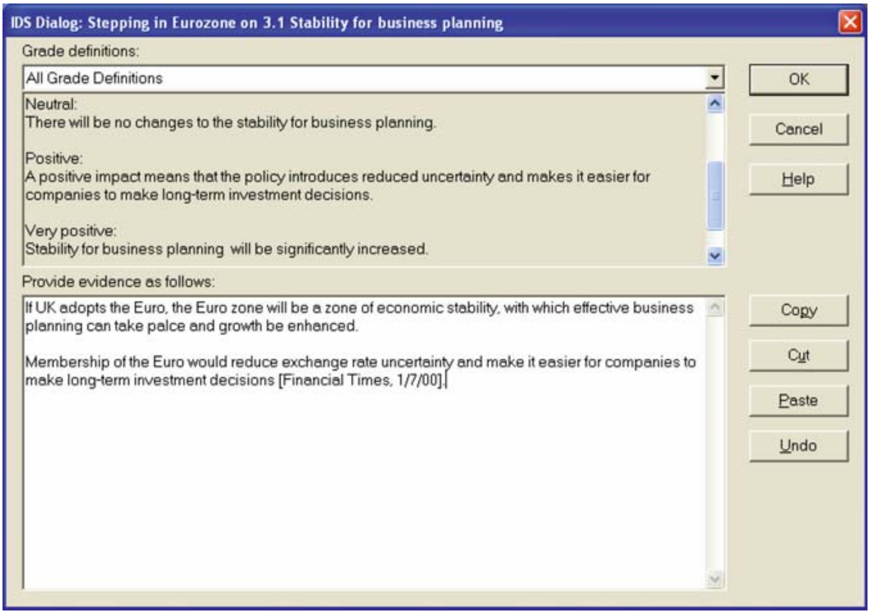

Fig. 9. Making qualitative assessment through evidence mapping

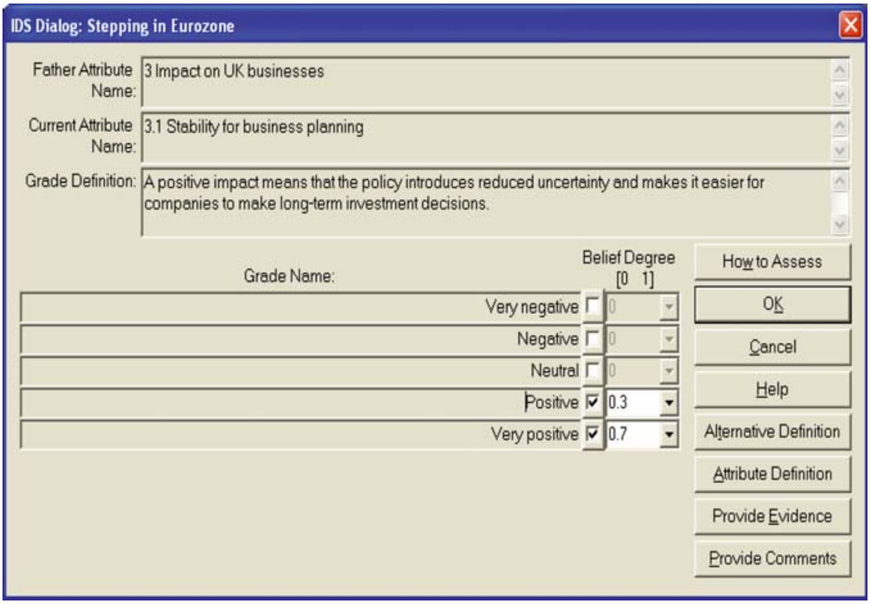

Fig. 10. Making qualitative assessment using belief degrees

3.4 Outcome Interpretation and Sensitivity Analysis

Outcomes and Interpretations

IDS generates different assessment results in both numerical and textual formats. To help the interpretation and the communication of the results, numerical ones are normally presented using graphs, including overall assessment scores of policy options, the potential performance variation ranges when there is missing information in an assessment (Fig. 11), and performance distributions (Fig. 12). Those graphs enable the comparisons among alternative policy options and are available on any selected areas at different levels of the assessment criteria hierarchy.

Ranking is based on overall assessment scores, a weighted sum of utilities of the grades in the aggregated performance distribution of each option, with belief degrees as weights as calculated by (11). The dark grey area in the ranking score graph (Fig. 11) indicates that there is some missing information in the assessment of the option "Not Join" and its performance score can be as high or low as the value marked by the top or bottom of the dark grey area respectively. The height of the dark grey area indicates the combined effect of the missing information. In the example shown in Fig. 11, the effect is small and will not affect the ranking no matter whether the missing information turns out to be in favour of the option or not.

Figure 12 shows the distributed overall performance of the two options regarding the UK Euro membership, based on the information collected in a

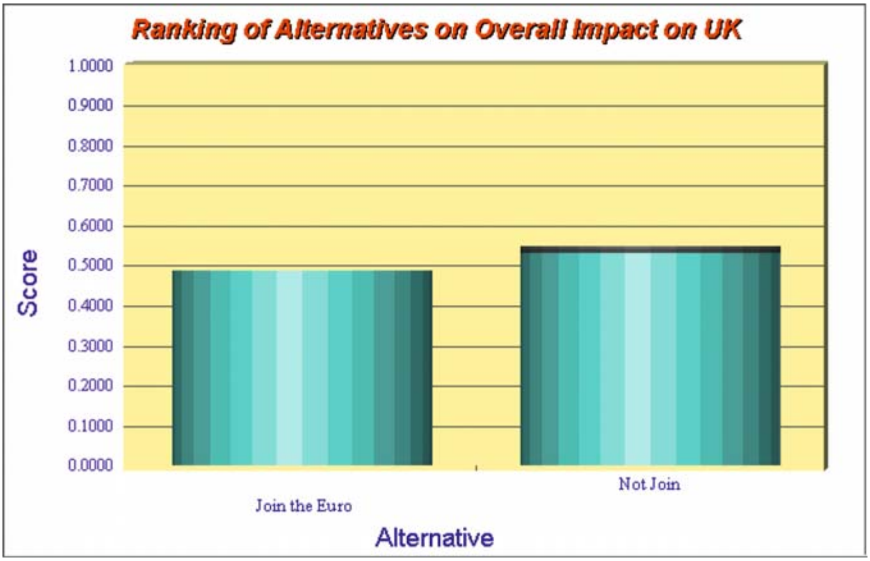

Fig. 11. Ranking of alternatives and performance variations due to uncertainty

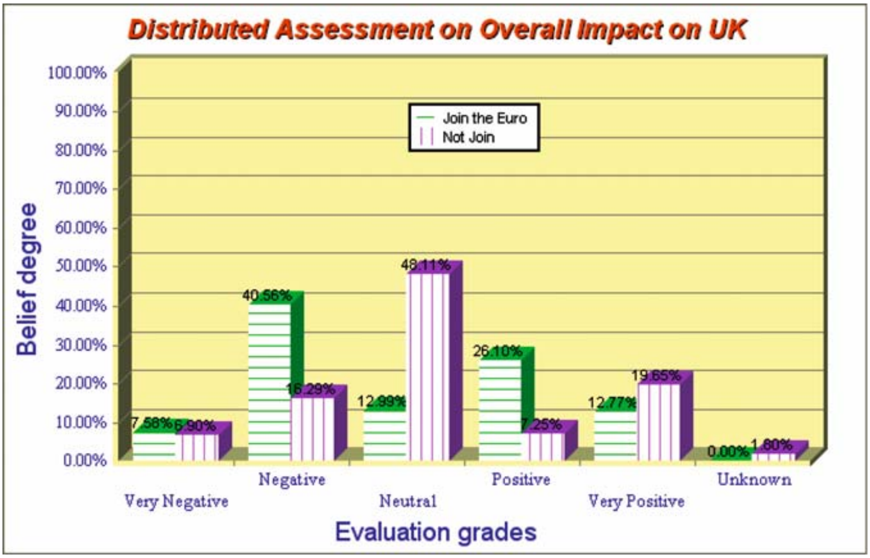

Fig. 12. Performance distributions

study. It reveals the performance composition in different categories (grades), and sheds light on why one option may be better than the other. Note that the portion of missing information is also revealed as a percentage in the Unknown category. The distribution shows that there are both negative and positive impacts if UK joins the Euro, and mostly neutral impacts if not. Such information allows policy makers to make an informed selection. If it is desirable to find the best or worst performing areas for an option, IDS provides a searching function for the purpose so that policy makers knows where exactly the risks are if going for the option.

To improve transparency in policy making processes, those graphical outcomes are available at not only the overall level represented by the top criterion, but also any level in the criteria hierarchy. Performances of all or selected options can also be compare on a selected set of criteria across different levels of the hierarchy.

To save time in assessment report preparation, IDS generates a tailor made assessment report for each option based on the evidence recorded and the assessment model, highlighting key areas to consider for each option. The assessment model, including assessment criteria, assessment grades and grading standards, and assessment results on every attribute, can all be saved in text files. The text files provide a basis with accurate and essential information for generating a detailed report. Together with a range of graphical display of outcomes, the report should help to communicate the assessment outcomes effectively.

Sensitivity Analysis

Sensitivity analysis is regarded as one of the very important step in any decision analysis process. It examines the effects of changes in some of the assumptions and judgements made during assessment processes, including parameters such as attribute weights, shapes of utility curves, and belief degrees assigned to the grades in an assessment. As those judgements and assumptions are somehow subjective in nature and difficult to be precise, sensitivity analysis will help to reveal how robust the outcomes, such as rankings of alternative options, are and therefore help decision makers to understand any risks involved in taking a particular course of action.

There are a range of sensitivity analysis functions supported by IDS which allow most parameters to be changed and the effects displayed. Three typical graphical sensitivity analysis interfaces are briefly described below.

The first type is interactive charts displaying the effects of changes in criterion weights and belief degrees assigned to a performance. For example Fig. 13 is a graph for examining the ranking changes of the 2 policy options in the Euro problem (join or not join the Euro) when the weight for the criterion "Impact on UK Business" changes. The current weight is 30 and the option "Not Join" is ranked higher than "Join". However, the graph shows that the ranking order will change if the weight becomes 40 or larger. If any weight is around a sensitive zone, the graph helps to draw the attention of policy

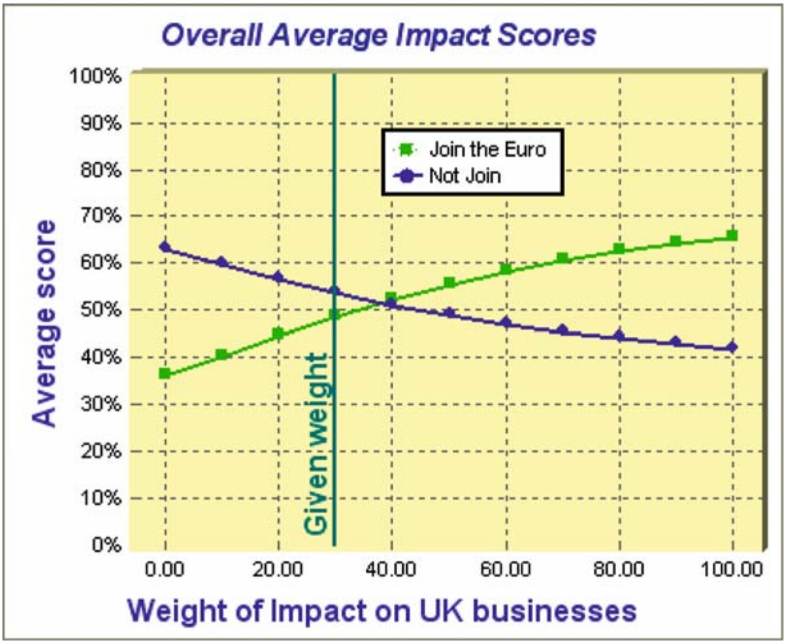

Fig. 13. Performance distributions

makers to the weight which may need to be re-examined and elicited using a number of approaches.

The second type of graphs shows the combined effects of different parameter changes on outcomes. This type of sensitivity analysis is normally referred to as global sensitivity analysis in literature [23, 36]. For example, Fig. 11 is one of such graphs displaying the combined effect of missing information in the assessments of "Not Join" on a number of criteria. Capable of providing global sensitivity information is a unique feature of IDS while most tools allow only one parameter to be changed at a time during sensitivity analyses.

The third type is the so called cost benefit or trade-off analysis graphs. It displays the scores of all alternative options on only two criteria at a time. For example, if the two criteria are "Costs" and "Benefits", as shown in Fig. 14, the two policy options in the Euro problem can then be positioned in the graph according to their performances on the two criteria. This type of graph allows users to exam whether the potential benefits of joining the Euro are worth the costs.

Model Fine Tuning

Impact assessment problems are complicated and it is unlikely to establish satisfactory models for the problems straight away. It is expected that the modelling phases need to be revisited from a number of times to make some adjustments on parameters such as weights after sensitivity analysis. It may also be necessary to check if there are any missing factors that need to be taken into account, or redundant attributes that need to be deleted. At the same time, the policy makers may need to challenge their own intuitions and rethink the problem and their preferences. Therefore the four phases of the

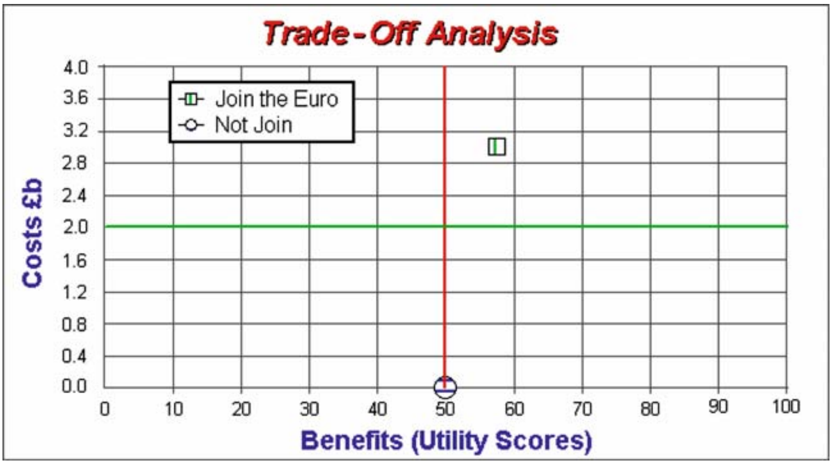

Fig. 14. Cost benefits analysis

MCDA process may need to be repeated until the policy makers are relatively satisfied with the model. The resulted model is termed as requisite, instead of optimal by [18]. This process is incisively summarised by French [9, p. 110].

"The decision makers begin the analysis ill at ease, discomforted by some half-perceived choice before them. As the analysis proceeds, their perceptions, beliefs and preferences evolve, guided by the consistency inherent in the underlying theory. Initially, the models used are very simple. But, gradually as intuitions emerge, the models are refined. A cyclic process is followed in which models are built, the output reflected upon and examined for sensitivity, intuitions emerge leading to revision of the models, and so on. This process is stopped when no further intuitions emerge."

4 Concluding Remarks

Policy making is a complicated process involving dealing with heterogeneous types of data with uncertain and missing information. As such, it needs to be supported with appropriate methodologies and tools. The ER approach and its software implementation, the IDS tool, are purposefully developed for dealing with such complication in IA assessment problems. Through a wide range of applications in supporting many complicated assessment activities, it is demonstrated that IDS is a flexible tool capable of handling data with uncertainties and providing more transparent, informative and reliable outcomes.

The capabilities of the ER approach are achieved through the use of a belief decision matrix to model an assessment problem. The use of belief decision matrix provides the following four advantages.

1. It helps maintain the originality of information in data. Using a conventional decision matrix, the distributed performance information, such as the one shown in (1) has to be approximated by a single value or grade which inevitably introduces information losses or distortion. Therefore the assessment of an option can be more reliably and realistically represented by a belief decision matrix than by a conventional decision matrix.
2. It provides policy makers with flexibility to collect and document assessment information in formats that are appropriate to certain circumstances, such as in single numerical values, probability distribution or subjective judgements with belief degrees. Consequently, it helps strengthen both the confidence and commitment levels of policy stakeholders in their chosen courses of action.
3. It allows all available information embedded in different data formats, including qualitative and incomplete data, to be maximally incorporated in the assessment processes, which again leads to more reliable outcomes.
4. It allows the assessment outcomes to be presented more informatively, which helps the effective communication of the outcomes.

The IDS software is developed to facilitate the application of the ER approach and realise its potential. It provides not only the technical support to apply the ER approach through friendly interfaces, but also cognitive support in the assessment process, and knowledge management, report generation and data presentation facilities. Encouraged and requested by users of IDS, a web based version of the tool has also been developed [33] and the UK Euro membership assessment example is made available online, which is accessible from the web site www.e-ids.co.uk.

The main limitation of the ER approach may be that people who are used to using conventional decision matrices for modelling MCDA problems may find that using belief decision matrices may look complicated, in particular for modelling purely quantitative MCDA problems. With the support of the IDS software and the power of modern computers, the complication associated with data processing is less a concern. To conclude this chapter, it may be noted that modelling an assessment problem using a conventional decision matrix is the same as using a belief decision matrix if all belief degrees in the latter are either 0 or 1.

Acknowledgement

This work forms part of the projects supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under Grant No: EP/F024606/1. The work was also partially supported by the grants from the Natural Science Foundation of China (NSFC) under Grant No.: 70540430127 and Grant No.: 70631003.

References

1. Balestra, G. and Tsoukias A. (1990), Multicriteria analysis represented by artificial intelligence techniques, The journal of the Operational Research Society, Vol. 41, No. 5, 419–430.
2. Belton, V. and Stewart, T.J. (2002), Multiple Criteria Decision Analysis — An Integrated Approach, Kluwer Publishers, London.
3. Beynon, M., Cosker, D. and Marshall, D. (2001), An expert system for multi-criteria decision making using Dempster Shafer theory, Expert Systems with Applications, Vol. 20, 357–367.
4. BRE (Better Regulation Executive) (2006), The Tools to Deliver Better Regulation. Available from http://www.cabinetoffice.gov.uk/regulation/documents/ria/pdf/consultation.pdf, accessed Feb. 2007.
5. Buchanan, B.G. and Shortliffe, E.H. (1984), Rule-Based Expert Systems, Addison-Wesley, Reading, MA.
6. Chen, L.H. (1997), An extended rule-based inference for general decision-making problems, Information Sciences, Vol. 102, 111–131.

7. Chin, K.S., Yang, J.B., Lam, J. and Guo, M. (in press), An evidential reasoning-interval based method for new product design assessment, IEEE transactions on Engineering Management (in press).
8. EFQM (1999), The EFQM Excellence Model, European Foundation for Quality Management, Brussels, ISBN 90-5236-082-0.
9. French, S. (1989), Readings in Decision Analysis, Chapman and Hall, London.
10. George, C. and Kirkpatrick, C. (eds.) (2007), Impact Assessment and Sustainable Development European – Practice and Experience, Edward Elgar Publisher, Northampton, MA.
11. Huynh, V.N., Nakamori, Y., Ho, T.B. and Murai, T. (2006), Multiple-attribute decision making under uncertainty: The Evidential Reasoning approach revisited, IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, Vol. 36, No. 4, 804–822.
12. Hwang, C.L. and Yoon, K. (1981), Multiple Attribute Decision Making Methods and Applications, Springer, Berlin Heidelberg New York.
13. Keeney, R.L. (1992), Value Focused Thinking, Harvard University Press, Cambridge, MA.
14. Keeney, R.L. and Raiffa, H. (1976), Decisions with Multiple Objectives, Cambridge University Press, Cambridge, MA.
15. Liu, J., Yang, J.B., Wang, J. and Sii, S.K. (2002), Review of uncertainty reasoning approaches as guidance for maritime and offshore safety-based assessment, Journal of UK Safety and Reliability Society, Vol. 23, 63–80.
16. Manoliadis, O.G. and Vatalis, K.I. (2003), An Environmental Impact Assessment Decision Analysis System for Irrigation Systems, Selected from papers presented at the 8th International Conference on Environmental Science and Technology, 8–10 September 2003, Lemnos, Greece. Available from http://www.gnest.org/journal/Vol5_No2/09_Manoliadis.pdf (accessed in Feb 2007).
17. OECD (1997), Regulatory impact analysis: Best practices in OECD countries, Paris, OECD.
18. Phillips, L.D. (1984), A Theory o Requisite Models, London School of Economics and Political Science, Decision Analysis Unit.
19. Porter, L. and Tanner, J. (1998), Assessing Business Excellence, Butterworth-Heinmann, London.
20. Pöyhönen, M. and Hämäläinen, R.P. (2001), On the convergence of multiattribute weighting methods, European Journal of Operational Research, Vol. 129, No. 3, 106–122.
21. Roy, B. and Vanderpooten, D. (1997), The European school of MCDA: Emergence, basic features, and current works, European Journal of Operational Research, Vol. 99, No. 1, 26–27.
22. Saaty, T.L. (1988), The Analytic Hierarchy Process, University of Pittsburgh.
23. Saltelli, A., Tarantola, S. and Chan, K. (1999), A role for sensitivity analysis in presenting the results from MCDA studies to decision makers, Journal of Multi-Criteria Decision Analysis, Vol. 8, No. 3, 139–145.
24. Sen, P. and Yang, J.B. (1994), Design decision making based upon multiple attribute evaluation and minimal preference information, Mathematical and Computer Modelling, Vol. 20, No. 3, 107–124.
25. Seppälä, J. (2001), Decision Analysis Frameworks for Life-Cycle Impact Assessment, Journal of Industrial Ecology, Vol. 5, No. 4, 45–68.

26. Shafer, G.A. (1976), Mathematical Theory of Evidence,Princeton University Press, Princeton, NJ.
27. Siow, C.H.R., Yang, J.B. and Dale, B.G. (2001), A new modelling framework for organisational self-assessment: development and application, Quality Management Journal, Vol. 8, 34–47.
28. Sonmez, M., Yang, J.B., Graham, G. and Holt, G.D. (2002), An evidential reasoning based decision making process for pre-qualifying construction contractors, Journal of Decision Systems, Special issue on Decision Making on Urban & Civil Engineering, Vol. 11, 355–381.
29. Stewart, T.J. (1992), A critical survey on the status of multiple criteria decision making theory and practice, OMEGA International Journal of Management Science, Vol. 20, No. 5–6, 569–586.
30. von Winterfeldt, D. and Edwards, W. (1986), Decision Analysis and Behavioural Research, Cambridge University Press, Cambridge, MA.
31. White, C.C. (1990), A survey on the integration of decision analysis and expert systems for decision support, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 20, No. 2, 358–364.
32. Wright, G. and Goodwin, P. (1999), Rethinking value elicitation for personal consequential decisions, Journal of Multi-Criteria Decision Analysis, Vol. 8, 3–10.
33. Xu, D.L. (2005), Web-based assessment via the evidential reasoning approach, Workshop on Human-Computer Interface Issues in e-Democracy, Manchester UK.
34. Xu, D.L. and Yang, J.B. (2002), MCDM analysis using IDS on whether UK should join the Euro, MCDM Winter Conference in Semmering, Austria.
35. Xu, D.L. and Yang, J.B. (2003), Intelligent decision system for self-assessment, Journal of Multi-criteria Decision Analysis Vol. 12, 43–60.
36. Xu, D.L. and Yang, J.B. (2004), Making decisions under uncertainties using the evidential reasoning approach, Proceedings of the 15th Mini-EURO Conference on Managing Uncertainty in Decision Support Models, Portugal.
37. Xu, D.L. and Yang, J.B. (2005), An intelligent decision system based on the evidential reasoning approach and its applications, Journal of Telecommunications and Information Technology, No. 3, 73–80.
38. Yager, R.R. (1987), On the Dempster–Shafer framework and new combination rules, Information Sciences, Vol. 41, No. 2, 93–137.
39. Yager, R.R. (1995), Decision-making under various types of uncertainties, Journal of Intelligent and Fuzzy Systems, Vol. 3, No. 4, 317–323.
40. Yang, J.B. (2001), Rule and utility based evidential reasoning approach for multiple attribute decision analysis under uncertainty, European Journal of Operational Research, Vol. 131, 31–61.
41. Yang, J.B. and Sen, P. (1997), Multiple attribute design evaluation of large engineering products using the evidential reasoning approach, Journal of Engineering Design, Vol. 8, 211–230.
42. Yang, J.B. and Singh, M.G. (1994), An evidential reasoning approach for multiple attribute decision making with uncertainty, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 24, 1–18.
43. Yang, J.B. and Xu, D.L. (2002), On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty, IEEE Transactions on Systems, Man and Cybernetics Part A: Systems and Humans, Vol. 32, 289–304.

44. Yang, J.B. and Xu, D.L. (2002), Nonlinear information aggregation via evidential reasoning in multiple attribute decision analysis under uncertainty, IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, Vol. 32, 376–393.
45. Yen, J. (1989), GERTIS: A Dempster–Shafer approach to diagnosing hierarchical hypotheses, Communications of the ACM, Vol. 32, No. 5, 573–585.
46. Zhang, Z.J., Yang, J.B. and Xu, D.L. (1989), A hierarchical analysis model for multiobjective decision making, In Analysis, Design and Evaluation of Man-Machine System (Selected Papers from the 4th IFAC/IFIP/IFORS/IEA Conference, Xian, P.R. China, September 1989), Pergamon, Oxford, UK, 13–18.
47. Zhao, M.Y., Cheng, C.T., Chau, K.W., and Li G. (2006), Multiple criteria data envelopment analysis for full ranking units associated to environment impact assessment, International Journal of Environment and Pollution, Vol. 28, No. 3/4, 448–464.