JOURNAL OF MULTI-CRITERIA DECISION ANALYSIS
J. Multi-Crit. Decis. Anal. **12**: 43–60 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com) **DOI**: 10.1002/mcda.343

Intelligent Decision System for Self-Assessment

DONG-LING XU* and JIAN-BO YANG†
Manchester School of Management, University of Manchester Institute of Science and Technology, PO Box 88, Manchester M60 1QD, UK

ABSTRACT

Many small and medium enterprises (SMEs) in the UK use the beta (Business Excellence Through Action) approach to the EFQM Excellence Model to conduct business excellence self-assessment, which is in essence a multiple criteria decision analysis (MCDA) problem. This paper introduces a decision support software package called Intelligent Decision System (IDS) to implement the beta approach. It is demonstrated in the paper that the IDS-beta package can provide not only average scores but also the following numerical results and graphical displays on:

- Distributed assessment results to demonstrate the diversity of company performances
- The performance range to cater for incomplete assessment information
- Comparisons
  - between current performances and past performances,
  - among different companies
  - among different action plans.
- Strengths and weaknesses

The IDS-beta package also provides a structured knowledge base to help assessors to make judgements more objectively. The knowledge base contains guidelines provided by the developers of the beta approach, best practices gathered from research on award winning organizations, evidence collected from companies being assessed and comments provided by assessors to record the reasons why a specific criterion is assessed to a certain grade for a company. Four small UK companies, the industry partners of the research project, have carried out the preliminary self-assessment using the package. The results and experience of the application are discussed at the end of the paper. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: decision support system; MCDA; quality award; self-assessment; business excellence; the evidential reasoning approach

1. INTRODUCTION

In increasingly competitive business environments, many organizations adopt the total quality management (TQM) approach to strive for business excellence. To monitor the progress towards business excellence, thousands of organizations across the world use self-assessment on a regular basis (Porter and Tanner, 1998).

There are a few popular business excellence models that provide standard criteria against which an organization can measure its performances. Quality award criteria set in the Japanese Deming Prize, the American Malcolm Baldrige National Quality Award (Porter and Tanner, 1998) and the European Quality Award (EQA) (EFQM, 1999) are among the most popular ones. Many small businesses in the UK use the beta (Business Excellence Through Action) approach (UKEF, 2000) to conduct self-assessment as their first step towards business excellence. Both EQA and the beta approach are based on the European Foundation for Quality Management (EFQM) model (EFQM, 1996, 1999). All of these models have a common feature. That is, they use a number of criteria with a hierarchical structure.

Self-assessment based on the business excellence models is a typical multiple criteria decision analysis (MCDA) problem (Yang *et al.*, 2001;

*Correspondence to: Ling Xu, Manchester School of Management, University of Manchester Institute of Science and Technology, PO Box 88, Manchester M60 1QD, UK. E-mail: ling.xu@umist.ac.uk
†E-mail: jian-bo.yang@umist.ac.uk
Contract/grant sponsor: EPSRC;
contract/grant number: GR/N65615/01;
Contract/grant sponsor: EC;
contract grant number: IPS-2000-00030

Received 30 April 2002
Accepted 23 December 2003

Belton and Stewart, 2002). To assess business excellence using these models, average scores are normally generated as performance indicators and both strengths and weaknesses are identifies to support the assessments. The advantage of such a method is its simplicity and practicality. However an average score does not provide sufficient information on the diversity of business performance or indicate where the business is doing well and where it needs to do better. Therefore, strengths and weaknesses are identified to supplement average scores. However, questions have been raised about the consistency between average scores and strengths and weaknesses (van der Wiele *et al.*, 1995; Siow *et al.*, 2001).

Recently, significant effort has been made by the authors and their colleagues to introduce a new and more complicated MCDA method, the evidential reasoning (ER) approach into self-assessment exercise (Yang, 2001; Yang *et al.*, 2001; Siow *et al.*, 2001). Several projects have been funded by the EPSRC of UK and the EC to conduct research on how to use the ER approach to assist the self-assessment process. A number of papers and research reports have been generated and published as the results of the research activities. The results show that the ER approach can help to reduce subjectivity in the self-assessment process and produce more beneficial information to an organization.

The promising preliminary results have encouraged further research and tests on applying the ER approach to help small businesses to conduct the self-assessment based on the beta approach and a business innovative capability assessment model developed by 12 European organizations (Xu et al., 2003). The beta approach is used by small businesses in the UK, with the help of licensed advisors. The goal of the beta approach is to help businesses to take a measured look at their performances and develop prioritised action plans. This paper will describe how the ER approach and its software implementation 'Intelligent Decision System (IDS)' (Xu and Yang, 2001) can be applied to implement the beta approach and how the implemented package, the IDS-beta package, can assist self-assessment and to achieve the goals of the beta approach.

In the following section, the beta approach will be outlined and the ER approach and the IDS software briefly explained. The process of implementing the beta approach using IDS will be described. The features and advantages of using the ER approach for self-assessment will be discussed using the data provided by the four UK companies which have conducted preliminary self-assessment using the IDS-beta package as part of an on-going research project.

2. THE BETA APPROACH

The beta approach is based on the EFQM model, tailor made for SMEs (under 50 employees) by the UK Excellence Federation (UKEF, 2000). It is used by UK SMEs to conduct self-assessment and aims to give businesses a head start in their path to excellence.

The beta approach uses the same nine major criteria as in the EFQM model. The first five criteria are referred to as enabler criteria and the other four as results criteria. They are

- Leadership,
- Policy and Strategy,
- People (Policy and Management),
- Partnerships and Resources,
- Processes,
- Customer Results,
- People Results,
- Society Results, and
- Key Performance Results.

The path to excellence cannot be achieved in one step but through continuous improvement. The continuous improvement follows the cycle: Plan–Do–Check–Act (PDCA). Based on this recognition, it is suggested in the beta approach that business should check how it is doing at each stage of the PDCA cycle for each of the nine criteria.

There are in total 58 checklist items in the beta approach under the nine top criteria. For each of the *enabler* criteria, the checklist covers every aspect of the PDCA cycle. For each criterion, the beta approach starts with a 'Do' statement to check what you actually *do* in that area. Then it looks under the surface by asking you why you do what you do—is it *planned*? Do you *check* and see whether your plans are working and achieving what you wanted to?

For example, for the leadership criterion, the checklist consists of the following items (UKEF, 2000):

1.1 Do: As leaders we set the direction for our business.

1.5 Plan: We identify our role as leaders, and plan how we'll carry it out.
1.6 Check: We regularly review our leadership to see that it's working effectively.
1.7 Act: As leaders we act to improve our leadership.

To help assessors to understand the checklist, further explanations are given for each of the items. For example, for item 1.1, the following is given:

1.1 Do: As leaders we set the direction for our business.

In other words

(a) We develop the business's fundamental purpose—what we are here for and where we are going.
(b) We develop meaningful values and beliefs for the business.
(c) We personally communicate the key messages: our values; our purpose and direction; our strategy, plans and targets.
(d) We make our priorities clear so that every one knows what matters most.

For each checklist item in the five enabler criteria, the beta approach uses the following six grades to measure how a business is performing:

- Nothing happening
- We're making a start
- We're getting somewhere
- It's working for us
- It's our normal practice
- We are seriously sophisticated.

For each of the four *result* criteria, the checklist covers the 'Check' and 'Act' areas and directs assessors' attention to look for evidence. For example, the following is one of the checklist items in the Key Performance Results criterion:

9.1 Check: We have a range of results to show what we are achieving as a business.

In other words,

(a) We measure our financial results.
(b) We measure our non-financial results.
(c) Our measurements cover all the area that are important to us to check our performance.

Financial, for example	Non-financial, for example
Cash flow	Market share
Revenue	Sales volume
Gross margin	Stock turnover

To match the main focus of the checklist in the result criteria, the beta approach uses the following six grades to assess each checklist item:

- None
- Few
- Some
- Quite a few
- Many
- Most

The complete description of the beta approach is available from the UK Excellence Federation (email: mail@uk-excellence.org).

3. THE EVIDENTIAL REASONING (ER) APPROACH & INTELLIGENT DECISION SYSTEM (IDS)

As its name implies, the Evidential Reasoning approach differs from traditional MCDA methods in that it uses evidence-based reasoning processes to reach a decision.

For more than two decades, many MCDA methods have been developed, such as multiple attribute utility theory (MAUT) and analytical hierarchy process (AHP) (Saaty, 1988) (Jacquet-Lagreze and Siskos, 1982; Belton and Stewart, 2002). Most of these methods are suitable for solving small scale MCDA problems without uncertainty. Unfortunately, real life MCDA problems, such as the business excellence self-assessment, are often of large scale and alternatives are often assessed using a mixture of qualitative and quantitative attributes with certain degrees of uncertainties. As such, there has been a growing need to develop scientific methods and tools for dealing with MCDA problems under uncertainty in a way that is rational, reliable, repeatable, and transparent (Stewart, 1992; Dyer *et al.*, 1992).

As part of the effort to deal with MCDA problems with uncertainties and subjectivity, the evidential reasoning (ER) approach has been developed using the concepts from several disciplines, including decision science (in particular utility theory), artificial intelligence, statistical analysis, fuzzy set theory, and computer technology (Zhang *et al*., 1989; Yang and Singh, 1994; Yang and Sen, 1994; Yang, 2001; Yang and Xu, 2002a, b).

There have been five major milestones in the development of the ER approach. The first one was the introduction of the belief degree concept into the decision matrix (Zhang *et al*., 1989). This brought a more versatile way in which a MCDA problem with uncertainties can be modelled. Traditionally, a MCDA problem is modelled using a so-called decision matrix, with each alternative being measured by a single value on an attribute. In the ER approach, a MCDA problem is described using an extended decision matrix with a belief structure or a belief decision matrix for short, with each alternative being assessed by a vector of paired elements. The paired elements are possible criterion values (grades) and their associated degrees of belief. A belief matrix allows more information to be contained in a model and the decision makers are no longer forced to aggregate decision information into a single value when the original information is a distribution. For example, if there is evidence to suggest that the business performance on one criterion can be assessed to a few grades, an assessor may do so, and associate each grade with a belief degree, as long as the total degrees of belief is 100% or less. This is not accommodated in a traditional decision matrix. Therefore the ER framework not only provides flexibility in describing a MCDA problem, it also prevents any loss of information due to the conversion from a distribution to a single value in the modelling process.

The second milestone was the introduction of the Dempster–Shafer theory (Buchanan and Shortliffe, 1984) into the ER framework so that the distributive information contained in a belief decision matrix could be aggregated to produce rational and consistent assessment results. For years, the authors searched for appropriate mathematical tools to fulfil such a task and the Dempster–Shafer theory was chosen because of its powerful evidence combination rules and its reasonable requirements to apply those rules (Buchanan and Shortliffe, 1984; Lopez de Mantaras, 1990; Yang and Singh, 1994; Yang and Sen, 1994).

The third milestone was the development of the rule and utility based information transformation techniques to transform various sets of evaluation standards to a unified set so that all criteria of both a quantitative and qualitative nature can be assessed in a consistent and compatible manner in the ER framework (Yang, 2001). This to certain extent mirrors the traditional normalization techniques used to handle quantitative criteria with different units in MCDA problems. The key difference is that in the ER framework the new techniques should in a sense preserve the distributive nature of information represented in the belief structure. It has been proved that by using the developed information transformation techniques not only the expected utilities of the original and the transformed assessments are equivalent but the degrees of incompleteness or completeness in the original assessments are also preserved rationally.

The fourth milestone is the enhancement of the approximate reasoning process of the original ER approach and the development of a rigorous reasoning process. Although the Dempster–Shafer theory has been used as the mathematical tool for information aggregation in the ER framework, it would provide irrational synthesis results in its original format when used to deal with conflicting evidence. Significant modifications have been made since the theory was first used in the ER approach to deal with MCDA problems (Yang, 2001; Yang and Xu, 2002a, b). The ER criteria aggregation process is in general a non-linear process and the non-linearity is decided by the weights of criteria and the way each criterion is assessed (Yang and Xu, 2002b), which is different from traditional MCDA methods such as weighting methods. It is proved that the new reasoning process satisfies the following common sense synthesis rules (CSSR) (Yang and Xu, 2002a):

CSSR 1: *If no sub-criterion is assessed to an evaluation grade at all then the upper-level criterion should not be assessed to the same grade either.*

CSSR 2: *If all sub-criterion criteria are precisely assessed to an individual grade, then the upper-level criterion should also be precisely assessed to the same grade.*

CSSR 3: *If all sub-criterion criteria are completely assessed to a subset of grades, then the upper-level criterion should be completely assessed to the same subset as well.*

CSSR 4: *If sub-criterion assessments are incomplete, then an upper-level assessment obtained by aggregating the incomplete* *basic assessments should also be incomplete with the degree of incompleteness properly expressed.*

The final milestone is the implementation of the ER approach by developing a window based software package, IDS (Yang, 2001; Yang and Xu, 2000, 2002b; Xu and Yang, 2001). As mentioned earlier, the ER approach models MCDA problems using belief matrixes, so inevitably the calculations involved in the aggregation processes could be relatively complicated. Without a user-friendly computer interface to facilitate information collection, processing and display, the task could be rather difficult to accomplish by hands, even for a relatively small scale MCDA problem.

Although the ER approach involves relatively complicated calculations, its computational requirements are linearly proportional to the scale of a MCDA problem, namely the number of criteria and the number of alternatives to be assessed in the problem. The package has been used in a variety of applications, such as motorcycle assessment (Yang and Sen, 1994), general cargo ship design (Sen and Yang, 1995), marine system safety analysis and synthesis (Wang *et al.*, 1995, 1996), executive car assessment (Yang and Xu, 1998), project management (Sonmez *et al.*, 2001) and organizational self-assessment (Yang *et al.*, 2001; Siow *et al.*, 2001). The experiences gained from these applications indicate that for MCDA problems with up to a few thousands of criteria and several alternatives, the calculation time is almost unnoticeable. It has also been proved in these applications that the ER approach not only produces consistent and reliable results for problems that can be solved using conventional MCDA methods, but also is capable of dealing with MCDA problems of the following features, which are difficult to handle using conventional methods without making further assumptions:

- Large number (hundreds) of criteria in a hierarchy,
- Large number of alternatives,
- Uncertainties,
- Mixture of deterministic and random information,
- Mixture of quantitative and qualitative information,
- Incomplete or missing information.

In addition to its mathematical functions, IDS is also a knowledge management tool. It records assessment information including evidence and comments in organized structures, and provides systematic help at every stage of the assessment including guidelines for grading criteria.

In the following sections, we will demonstrate how the IDS software can be used to implement the beta approach and help small businesses to conduct self-assessment.

4. SOFTWARE IMPLEMENTATION OF BETA APPROACH USING IDS

The IDS main window is shown in Figure 1, where there are a tree view window for displaying a hierarchy of criteria, a list view window to display the names of the businesses that are assessed, a menu bar where all IDS functions can be found for file management, model building, information input, result analysis and report, and online help, and a tool bar where short cuts for the frequently used IDS functions are provided.

4.1. Build the criteria hierarchy of beta approach using IDS

4.1.1. Define attributes and their assessment grades. As discussed in Section 2, the beta approach uses the following 3 levels of criteria (or attributes):

- One top level attribute: Overall Performance.
- Nine second level attributes: the nine major criteria.
- Third Level attributes: Checklist items. There are in total 58 checklist items distributed under the nine second level attributes.

When IDS is running for the first time, one alternative and one attribute will be displayed in the right and left pane of the main window, respectively. By selecting an attribute in the window with the left mouse and then clicking the blue arrow button, new attributes can be added around the selected attribute. If the blue button is clicked, then the selected attribute can be defined using the interface as shown in Figure 2. The definition of the attribute includes the following information: its name, whether it is qualitative or

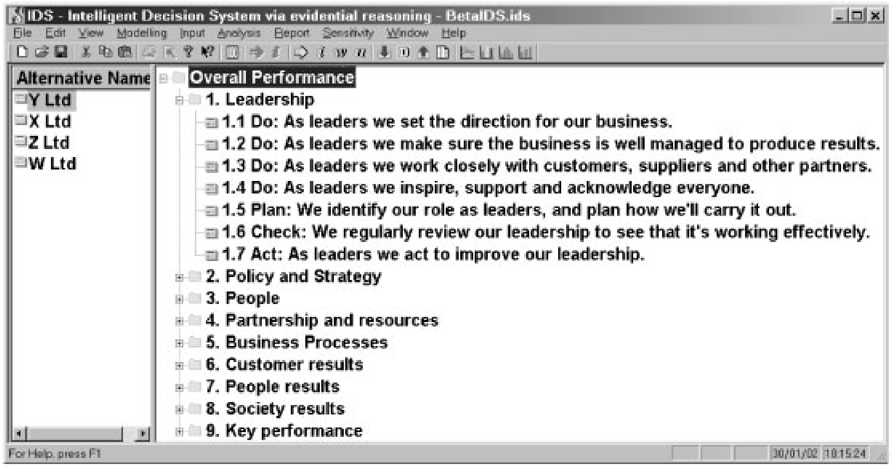

Figure 1. IDS main window and the beta approach criteria hierarchy.

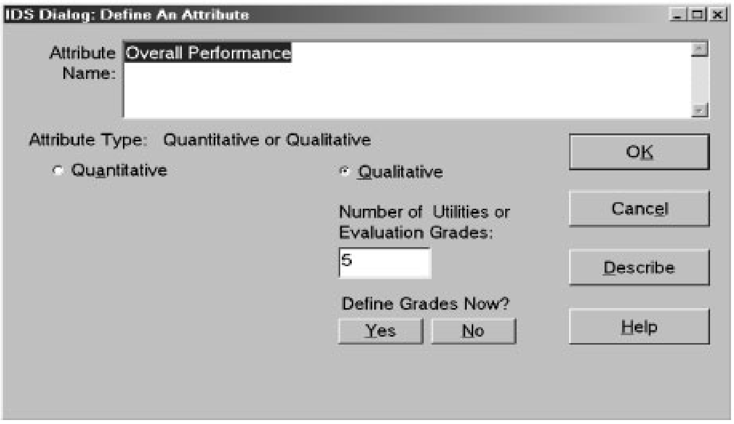

Figure 2. Attribute definition window.

quantitative, if qualitative how many grades will be used to assess it and what are the names, meanings and utilities of the grades (Figure 3), and if quantitative what is the unit for its measurement, whether it is deterministic or probabilistic, and what is the range of its possible values (Figure 4).

Figure 1 shows the built criteria hierarchy of the beta approach using IDS. This hierarchical structure can be changed whenever necessary; existing attributes can be deleted and new attributes added anywhere as required. IDS has no limit on the number of levels of attributes that can be added to the hierarchy. The structure is wholly determined by the requirement of a specific MCDA problem in hand.

Depending on practical appropriateness, different attributes may use different grades. For the Overall Performance criteria, people normally use the following five grades (Dale and Lascelles, 1997): {*Beginner, Committed, Improver, Award Winner, World Class*}, as shown in Figure 3.

As discussed in Section 2, for the five enabler criteria in the second level and their sub-criteria (checklist items) in the third level, the beta approach uses the following set of six grades:

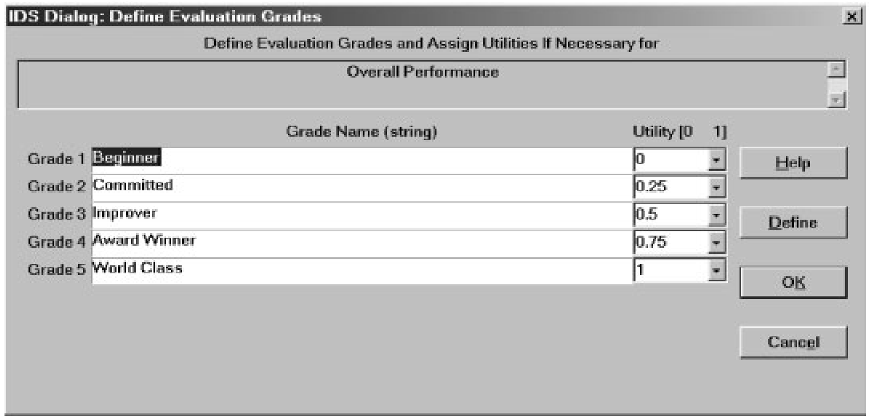

Figure 3. Grade definition window.

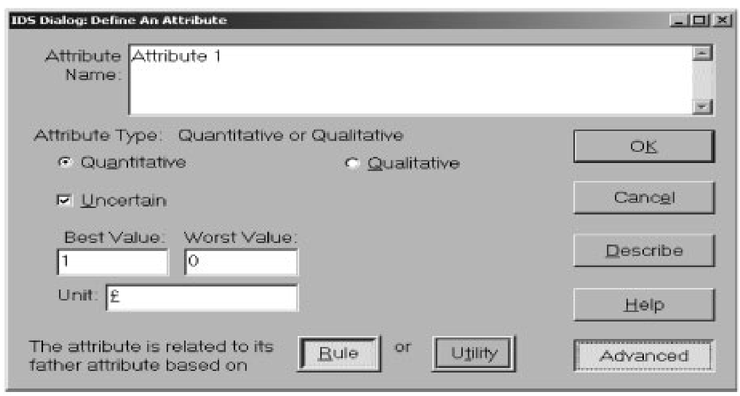

Figure 4. Quantitative definition window.

{*Nothing happening*, *We're making a start*, *We're getting somewhere*, *It's working for us*, *It's our normal practice*, *We are seriously sophisticated*}

For the result criteria and their checklist items the grades are given by

{*None*, *Few*, *Some*, *Quite a few*, *Many*, *Most*}

4.1.2. Convert grades. Note that the three sets of grades are different in numbers and wordings. When IDS detects such differences in grade definitions between two levels of attributes, it will prompt the user to convert the grades of lower level attributes to those of the associated higher-level attribute. The conversion can be conducted using the rule or utility based information transformation techniques (Yang, 2001). If the user does not wish to conduct the conversion himself, the IDS will conduct the conversion based on the assumption that each set of the grades is evenly distributed in the utility space.

As the beta approach is designed for businesses at the early stages of working with the EFQM model, it is specified in the beta approach that the 6th grades for both enabler criteria and result criteria are used as indicators that a business is doing so well that it is beyond the best scope of the beta approach and the business should adopt a more sophisticated approach. In other words, the grades are so designed in the beta approach that the 5th grade is the top grade and the 6th grade will not earn more marks than the 5th one. The conversion between the Overall Performance grades and the grades of its sub-attributes such as Leadership and Customer Satisfaction should be based on utility equivalence. To do so, what is needed is to specify the utilities for each of the grades of the sub-criteria. By default, IDS assumes that the 6 grades of the sub-criteria are evenly distributed in the utility space. However, based on the above understanding of the beta approach, it is assumed that grades 1 to 5 are distributed evenly over the utility space with grade 6 having the same utility as grade 5. Figure 5 shows the IDS interface for directly assigning the grade utilities of the Leadership criteria. Note that the utilities shown in Figure 5 are defined in a consistent way with the scores of each grade used in the EFQM model (Yang *et al.*, 2001). In this case, they are indeed defined to be evenly distributed in the utility space. In general, the utilities of grades should be estimated using appropriate methods such as the probability assignment method (Sonmez, 2002).

Note that the beta approach does not require quantifying the grades. Also note that in IDS it is not required to quantify grades in order to generate the overall distributive performance for a business. The above discussion for quantification is introduced for two purposes. One is to provide a business with a numerical indicator of their business performances. The other is to demonstrate that IDS is capable of providing what a traditional MCDA method can provide.

In IDS, criteria are aggregated using the evidential reasoning approach which combines degrees of belief rather than operating on utilities (or scores). This differentiates the evidential reasoning approach from weighting methods or other traditional MCDA methods (Hwang and Yoon, 1981). The properties and rationality of the ER approach are theoretically analysed and proved in separate papers (Yang, 2001; Yang and Xu, 2002a, b).

4.1.3. Build a knowledgebase. In the interfaces as shown in Figure 2–5, there are buttons such as 'Define' and 'Describe'. These buttons allow decision makers to record text information, such as further details about an attribute if there is any (Figure 2) and guidelines for assigning a specific grade to an attribute (Figure 3). Guidelines could be weighted to reflect their relevant importance in assessment. The information can be retrieved at a click of a button and is helpful when assessors need to clarify the definition of an attribute and the standard for a grade. As the information is built into the model and provided to assessors at the right place at the time when it is needed, it helps to increase the consistency and reduce subjectivity of the scores, even for inexperienced assessors.

For each checklist item in the beta approach, there is further explanation as described in Section 2. For example the checklist item 9.1 'Check: We have a range of results to show what we are

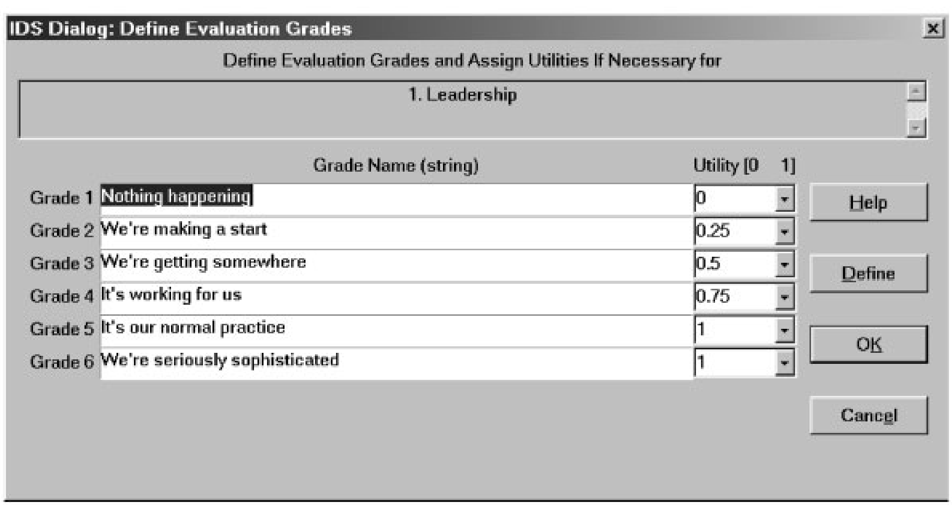

Figure 5. Assign grade utilities.

achieving as a business' has the following explanations:

In other words

(a) We measure our financial results.
(b) We measure our non-financial results.
(c) Our measurements cover all the areas that are important to us to check our performance.

Financial, for example	Non-financial, for example
Cash flow	Market share
Revenue	Sales volume
Gross margin	Stock turnover

The explanations clarify what the item means in detail. They are valuable knowledge and can be recorded by clicking the 'Describe' button in Figure 2. The knowledge recording interface is shown in Figure 6. In the example in Section 4.2, it will be shown how and when the knowledge can be retrieved and used.

4.1.4. Assign attribute weights. It is common that each attribute is of different importance to its upper level attribute. In the beta approach, it is specified that the Customer Results is weighted 2 whilst the Society Results 0.6 out of 10.

IDS provides a few interfaces to allow users to assign weights. The interfaces are activated by first clicking the attribute that a weight needs to be assigned to and then the blue W button. If the weights are known, which is the case for the beta approach, then the visual scoring interface as shown in Figure 7 can be used. The weight can be assigned to each attribute by either dragging the corresponding bar in the graphics pane or by typing in the blank box under the 'Weight' heading on the right.

In case weights are not known and need to be determined, IDS provides a few methods and interfaces to help decision makers to assign weights, including the eigenvector method, the geometric mean method and the direct assignment method. It should be noted that the absolute values of the weights are not important as long as they are relatively correct among attributes which are at the same level and share the same upper-level attribute. Also note that in the beta approach the weights of the criteria have been decided by the EFQM model. So the users merely use IDS to enter the weights rather than eliciting new weights to the criteria.

4.1.5. Comments. After the four steps described above are finished, the beta approach is fully integrated into the IDS software and the IDS-beta package is ready for self-assessment exercises. The above beta approach implementation process is not a task for an assessor. It can be built by beta approach developers or certified providers. Once built, the IDS-beta package can be distributed to assessors in businesses as an expert system to assist the self-assessment exercise. The assessors will not need to repeat the above modelling process, but

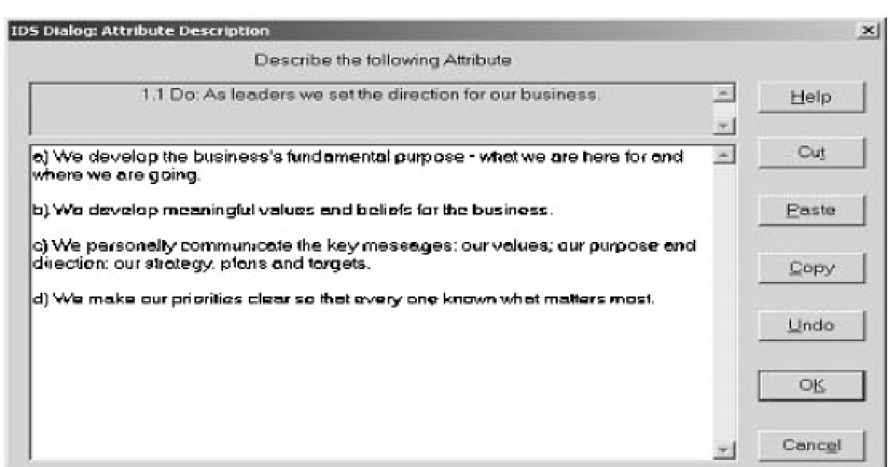

Figure 6. Interface for recording expert knowledge.

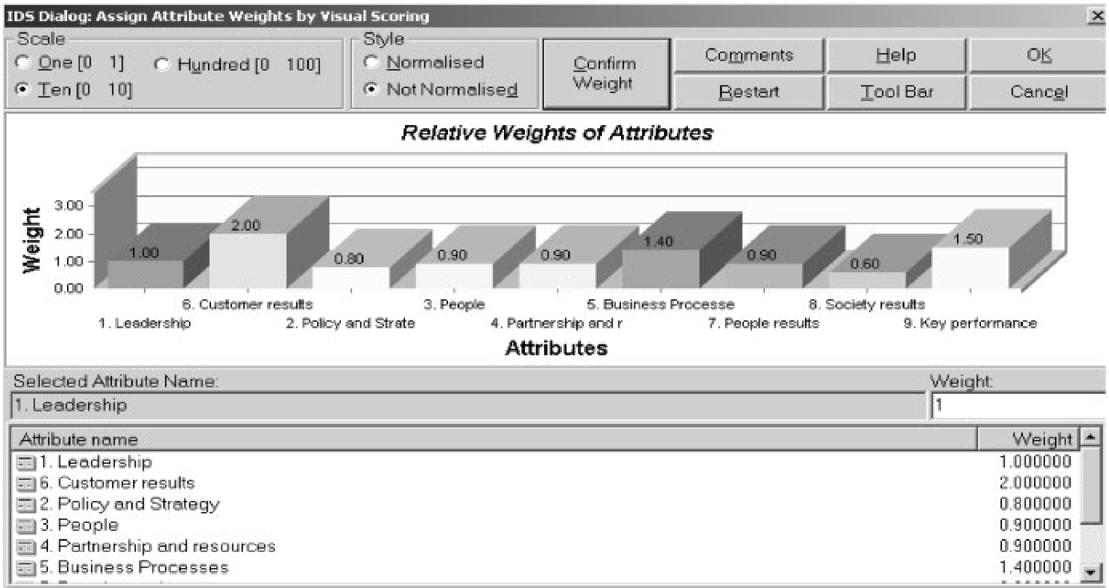

Figure 7. Interface for assigning weights to attributes.

to concentrate on collecting information and making assessments as described in the following sections.

4.2. Collect assessment information using the IDS-beta package

To collect assessment information using the IDS-beta package, the first step is to add the assessed company to the alternative list. This can be accomplished by a click on the Alternative Name pane and then the yellow arrow button  in the main window (Figure 1). By right clicking the newly added alternative object, the name and further description can be edited and recorded. The first company using the IDS-beta package to conduct the self-assessment is Y Ltd (not a real name). Y Ltd is a company located in the North West of the UK and manufactures stainless steel components for pipeline industry. The other three companies are also located in the North West region of the UK. To protect the interests of the companies, we named them as X Ltd, W Ltd and Z Ltd, respectively.

To input assessment information for an alternative (a company), the alternative has to be selected from the list on the left pane of the main window (Figure 1) by clicking it once. Once the alternative is highlighted (blue when the cursor is still in the Alternative pane or grey when outside the pane), all information entered as described below will be associated with the selected alternative.

It should be noted that only the bottom level attributes need to be assessed directly. Hence information input and recording is always related to the bottom level attributes. The assessment results of higher-level attributes are calculated automatically using the ER approach and based on the performances of their sub-attributes. Therefore the assessors need only to concentrate on assessing the bottom-level attributes.

In the beta approach, the bottom-level attributes are the checklist items. By a double click on any of the items listed in the main window (Figure 1), for example item 9.1, the interfaces shown in Figure 8 will be activated.

If only a fast diagnosis assessment or a perception based assessment is required, assessors need only to tick the appropriate grades and select or type the associated degrees of belief in the interface. There is no limit on how many of the defined grades can be assessed to an attribute, as long as the total degree of belief is equal to or less than 1 (100%). In a group decision making situation or when the information results from a survey, the degrees of belief may represent the percentage of people who have assessed the grades to the attribute.

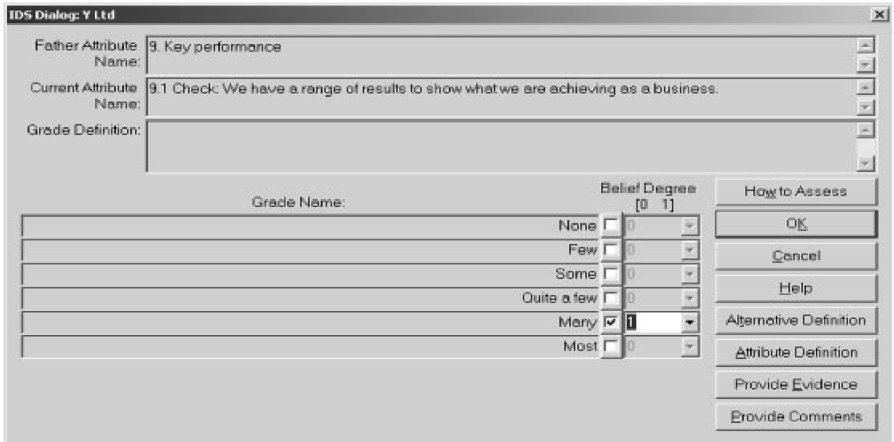

Figure 8. An interface for inputting assessment information.

There are a few buttons on the right-hand side of the assessment window (Figure 8). If a serous assessment is required, evidence should be collected and recorded using the Provide Evidence button. The evidence collected should be compared with the guidelines built in the knowledge base (Section 4.1.3) and comments on any thoughts related to the assessment can also be recorded for future references. By clicking the Provide Comments button, the evidence mapping window as shown in Figure 9 will be displayed. There are three panes in this window. The first one is the Grade Definitions set in the beta approach and used as guidelines for awarding a grade. The second one displays the evidence collected. As the evidence collected and the relevant guidelines are displayed side by side, the grade assessing process is made easier. The third one is directly below the previous two and used for inputting and editing any relevant comments, such as concern and reasons as to why the attribute is assigned the given grades and the associated degrees of belief.

Other supporting functions available from the buttons on the right hand side of the assessment window (Figure 8) include IDS help on How to Assess, IDS content sensitive Help, Alternative Definition and Attribute Definition (Figure 10). The first two help buttons provide access to the built-in IDS help system. The latter two help buttons provide access to the knowledgebase built into the package earlier as described in Section 4.1.3.

Based on the comparisons between evidence and guidelines, grades and the related degrees of belief can then be assessed to the attribute. We call this process an evidence mapping process. This process should be carried out on every bottom attribute for serious assessment. The information can be saved and retrieved from the File menu and modified by double clicking the attribute concerned whenever necessary. Once the process is finished, the assessment results will become available by clicking the button. The button can be pressed at any time or after any modification.

Once the button is clicked, the IDS software starts to syntheses all the information inputted to the bottom level attributes for all alternatives and generate assessment results for all the upper levels attributes. The calculation is carried out on the background and takes only a fraction of a second for the four companies. In next section, the process of using IDS to examine and present the generated results will be demonstrated and the results will be discussed.

4.3. View assessment results

The IDS model can generate different types of assessment results graphically, such as the ranking of the alternatives assessed, the performance trend of an assessed business if two or more rounds of assessment results are available and the diversity of its performance. The appearance of those graphs can be edited and the graphs can be exported to Word documents or PowerPoint files. IDS also has a search function to help identify strengths and weaknesses. The model can also be used to analyse the effect of action plans by simulating different improvement scenarios. The

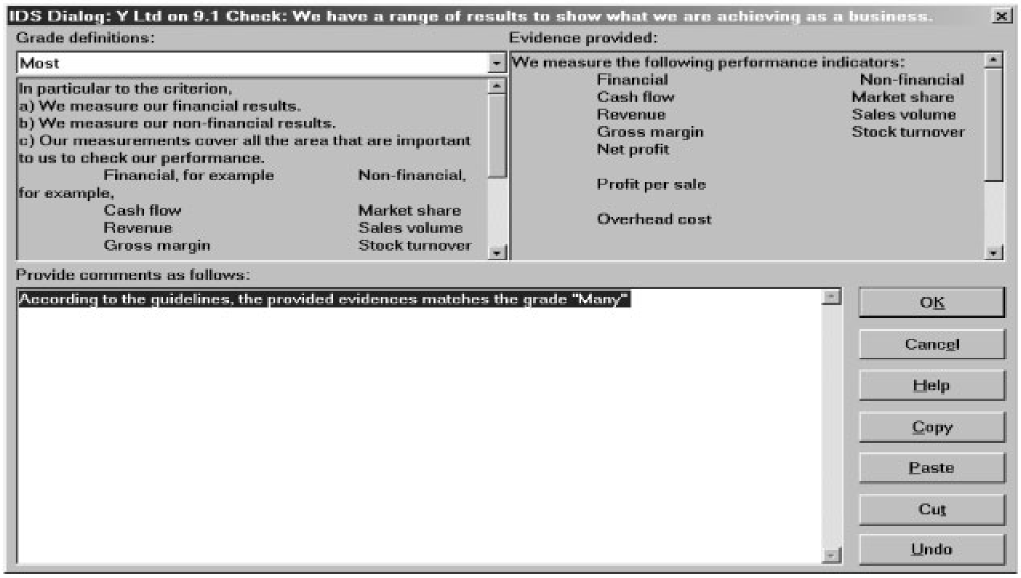

Figure 9. Evidence mapping window.

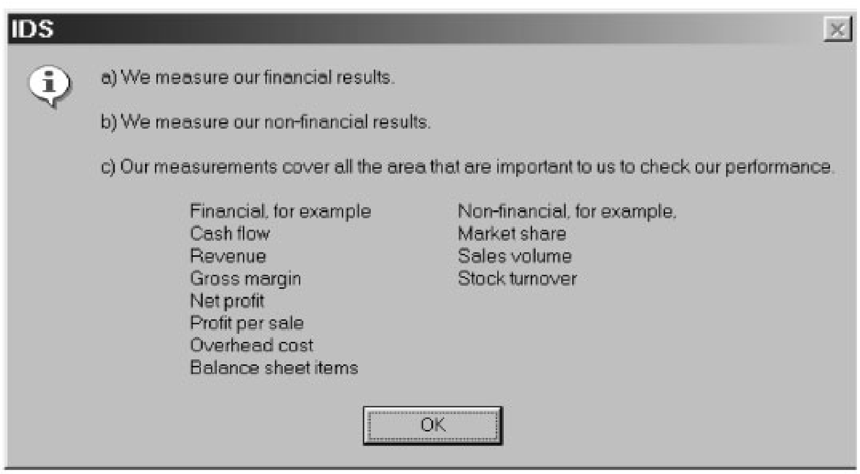

Figure 10. Information on attribute definition.

examples of the results are demonstrated in the following subsections.

4.3.1. Simple ranking. Similar to all other MCDA methods and tools, IDS can produce a simple ranking for all the alternatives assessed. The ranking of the four assessed companies is displayed in Figure 11. It indicates that Y Ltd achieved 75% (award winning position) across all beta criteria.

4.3.2. Performance range. If there is any incomplete assessment where the total degree of belief assigned to the grades for any attribute (Figure 8) is less than 1, then the incompleteness will propagate to the upper level criteria. The ER approach uses the concepts of belief degrees to

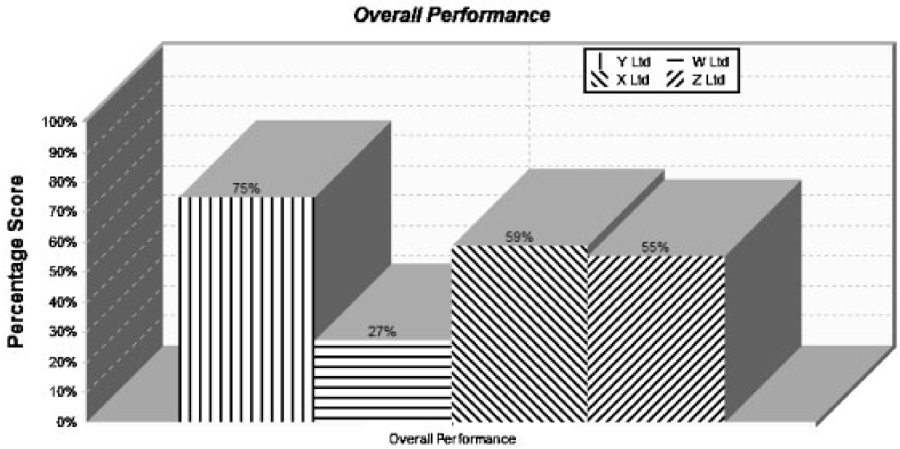

Figure 11. Performance ranking of the assessed companies.

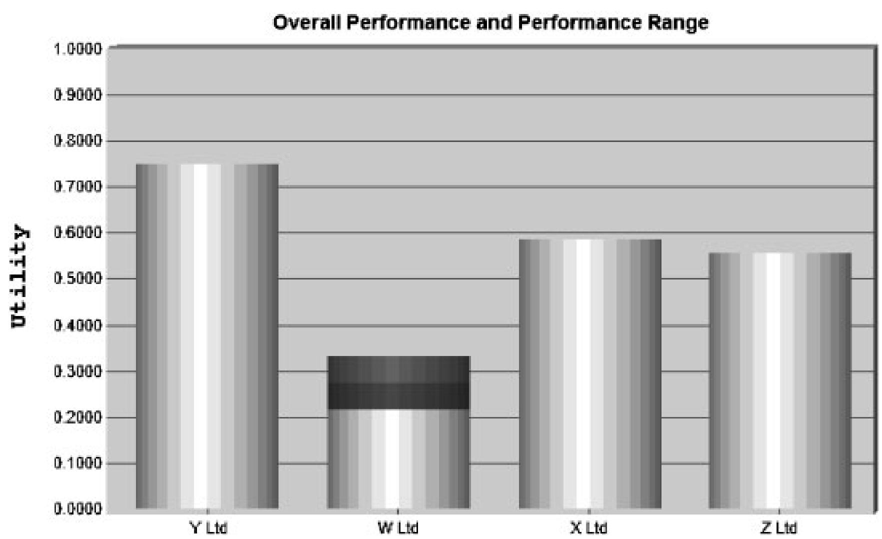

Figure 12. Performance range in case of uncertainty.

handle incomplete information and can provide the lower and upper bounds of the belief degrees. Since the information for assessing W Ltd on some criteria is incomplete, for example, the assessment results are indicated by a range of scores as shown in Figure 12. If the unknown portion of the information is favourable to the company, the company may have the best possible score of about 33%. If the missing information is unfavourable, then the worst possible score will be 22%. The average of the two is 27%, which is used for ranking. For each of the other three companies, the worst possible, average and best possible scores are all the same. This indicates that there is no missing assessment information for the three companies.

4.3.3. Distributed assessment results. The average scores or the range of scores discussed in Sections 4.3.1 and 4.3.2 are convenient for ranking, but they tell nothing about the diversity of the companies' performance. Conventional MCDA methods are unable to provide distributed performance information for an alternative on any upper level criterion.

Different from other MCDA tools, IDS uses the ER approach to aggregate assessment information. The distributed performance information is preserved during the ER aggregation process from bottom level attributes to top-level attribute. Therefore IDS can provide distributed performance for an alternative on any attribute in the hierarchy. Figure 13 is produced by IDS to show the distributed overall performance of Y Ltd. As discussed in Section 4.3.1, Y Ltd scored 75% on average. From Figure 13, it is clear that most of its criteria are assessed to the top 3 grades (30% world Class, 45% Award Winner and 22% Improver) and about 3% to the lower grades. If those low performance areas cause any concerns, IDS will search and display those areas as described below.

4.3.4. Strengths and weaknesses analysis. With the help of the distributed assessment results, strengths and weaknesses can be identified by searching for areas which have heavy weights and where low and high grades are assessed to. IDS provides such a searching function. Figures 14 and 15 display the areas where Y Ltd has achieved 100% and below 50% of the scores, respectively. The areas of strengths and weaknesses are listed as the *X*-axis labels in the two graphs.

4.3.5. Cross comparison between companies. Comparisons among selected alternatives can be carried out on any number of criteria at any level in the hierarchy. The following graph (Figure 16) shows the average scores (in the range of 0 to 1) of the four companies over the 10 criteria at the top two levels.

The comparison can also be focused on the performance diversity on any criterion at any level. The graph in Figure 17 shows that the performance of Z Ltd is mostly in the middle grades, the performance of W Ltd is more at the beginner's level with about 12% unknown information and the performance of Y Ltd is at the Award Winner's level.

4.3.6. Non-graphical report. In addition to examining the assessment results graphically, the information stored in computer such as the knowledge base, assessment information and assessment results can be saved in text files in a structured format. The contents of the files can be used in reports of other formats such as a Word document. Figure 18 shows how the function can be activated from the Report menu of the IDS main window.

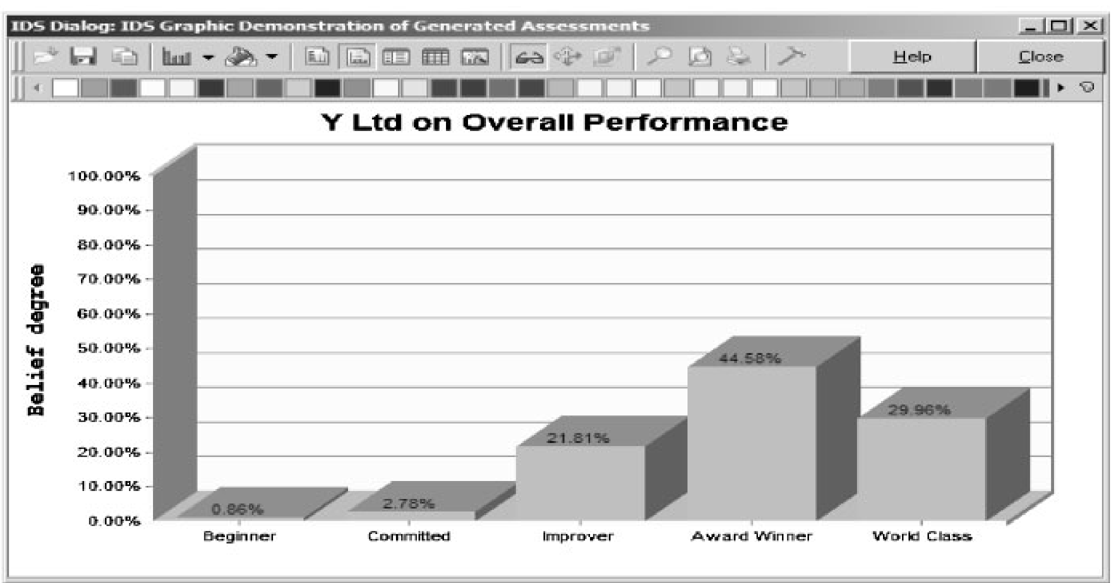

Figure 13. Distributed performance of the Y Ltd.

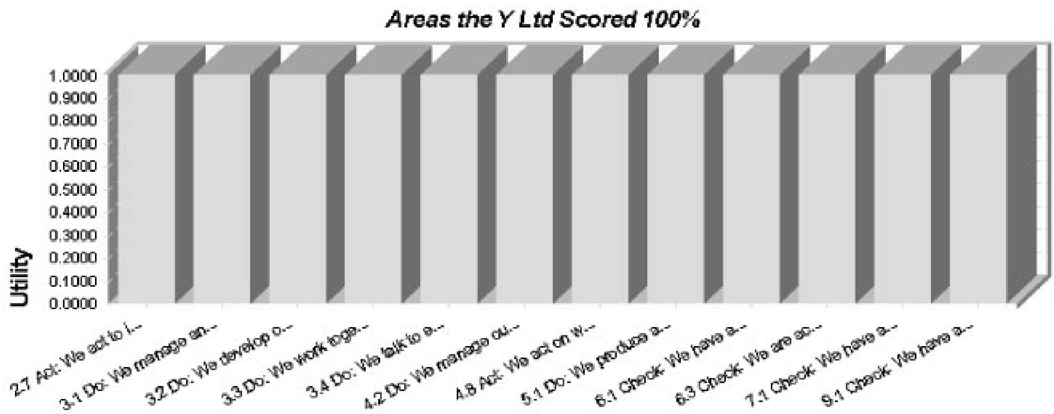

Figure 14. Areas of strengths.

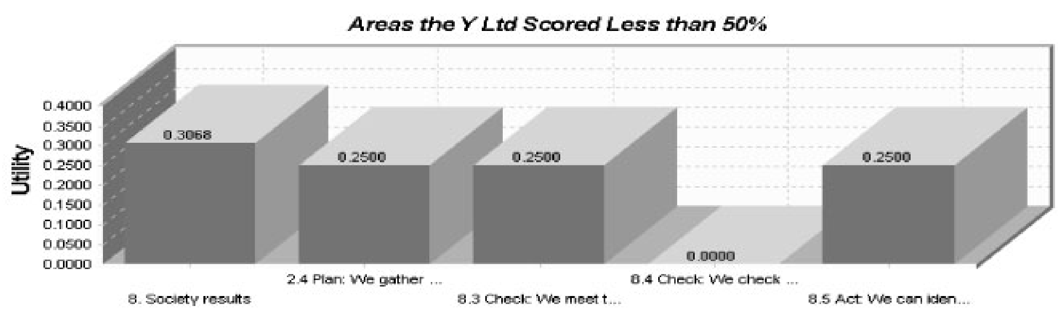

Figure 15. Areas of weaknesses.

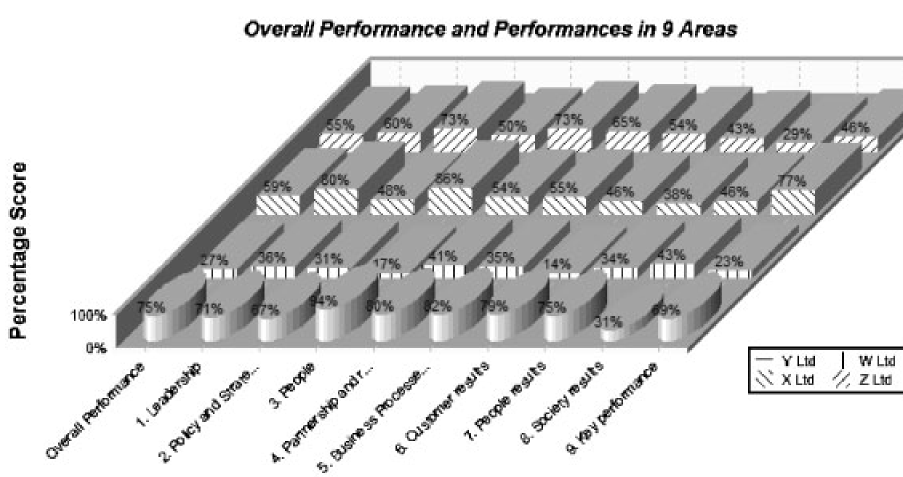

Figure 16. Comparisons of the four companies on the overall performance and the nine major criteria.

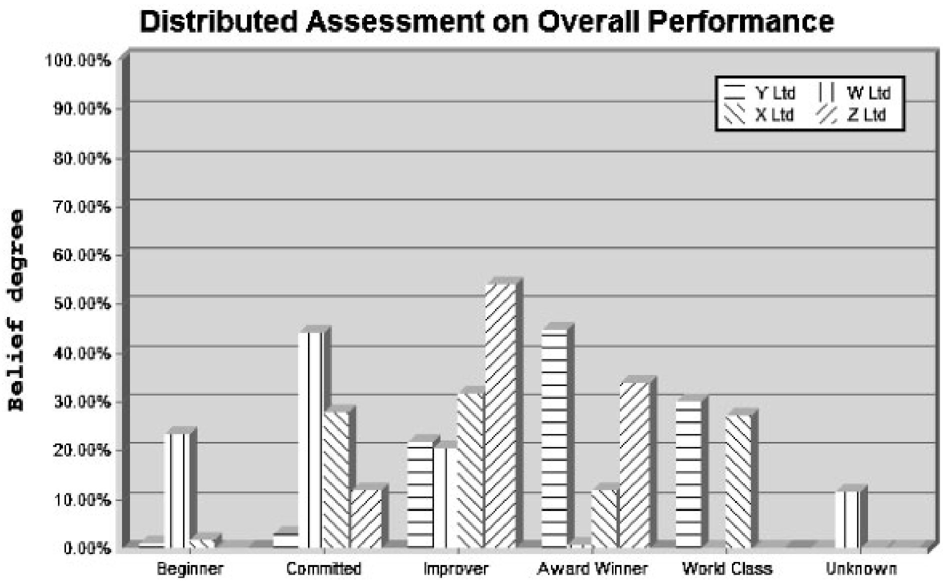

Figure 17. Comparison of the four companies on the overall performance: distributed results.

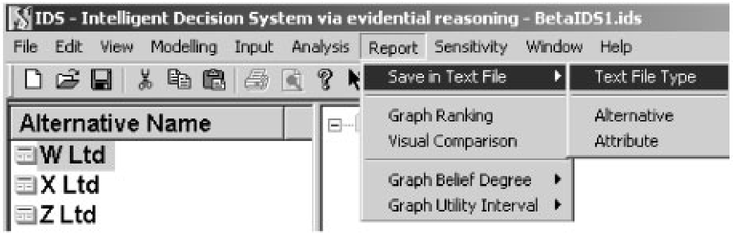

Figure 18. The IDS report menu for saving information in text format.

5. CONCLUDING REMARKS

To ensure continuous improvement, a company needs to assess its current performances against its past performances or industry standards. This requires vigorous self-assessment on a regular basis. With the help of IDS, the self-assessment process can be made easier, more efficient and consistent. To use the IDS-beta package for future self-assessment, the assessment results of a previous round can be copied as the starting point of a new round. IDS provides such ‘copy’ and ‘paste’ functions. Based on previous results, new assessment information can be entered for changed areas. The effects of the change can be examined immediately in comparison with previous results.

The buttons for various help and guidelines are appropriately placed for easy access and can help to make sure that grades are awarded based on consistent standards. The comments recorded in the earlier round can also help to conduct assessment in similar situations. Comments on new situations will help future assessments.

The conformation of the ER approach to the five common sense synthesis rules will always guarantee that the assessment results are rational. Any good or bad grades shown in overall performance or any higher level criteria can always be traced back to areas where a company has done well and where it needs to improve. Results can be examined at any level and on any criterion. Therefore the assessment process is truly transparent.

It is not only easy to compare current performances with past or targeted performances, but also possible to simulate and predict the future performances of any improvement action plans. Unlike AHP (Belton, 1986), assessment results generated by the ER approach will not be affected by newly added alternatives. Therefore the assessment results and process are always repeatable and reusable.

The feedback from the four UK small companies which have tested the package is encouraging. More requests have been received to implement other European countries' TQM models using IDS. There have been increasing interests in other application areas such as supplier assessment and product selections. The ER approach is under continuous development and the IDS package is also in the process of continuous improvement towards excellence.

ACKNOWLEDGEMENTS

The work reported in this paper was carried out purely for research purposes, involving project partners only and with no commercial activity. It was supported by the UK Engineering and Physical Science Research Council (EPSRC) under the grant number GR/N65615/01 and the European Commission (EC) under the grant number IPS-2000-00030. The authors are grateful to the British Quality Foundation for its permission to allow them to use the beta approach in the research and to publish the results. The authors are in debt to the four managing directors of the industrial partners of the projects for their support of the research and participation in the preliminary self-assessments using the IDS-beta package. The authors also thank the colleagues of the Manchester School of Management of UMIST in the UK, who supported the research.

REFERENCES

Buchanan BG, Shortliffe EH. 1984. *Rule-Based Expert Systems*. Addison-Wesley: Reading, MA.

Belton V, Stewart TJ. 2002. *Multiple Criteria Decision Analysis—An Integrated Approach*. Kluwer Academic Publishers: Dordrecht, ISBN 0-7923-7505-X.

Belton V. 1986. Comparison of the analytic hierarchy process and a simple multi-attribute value function. *European Journal of Operational Research* **26**: 7–21.

Dale BG, Lascelles DM. 1997. Total quality management adoption: revisiting the levels. *The TQM Magazine* **7**(6): 418–428.

Dyer JS, Fishburn PC, Steuer RE, Wallenius J, Zionts S. 1992. Multiple criteria decision making, multiattribute utility theory: the next ten years. *Management Science* **38**(5): 645–654.

EFQM. 1999. *The EFQM Excellence Model*. European Foundation for Quality Management: Brussels.

EFQM. 1996. *The European Quality Award Application Brochure*. European Foundation for Quality Management: Brussels.

Hwang H, Yoon K. 1981. *Multiple Attribute Decision Making—Methods and Applications*. Springer: New York.

Jacquet-Lagreze E, Siskos J. 1982. Assessing a set of additive utility functions for multicriteria decision making, the UTA method. *European Journal of Operational Research* **10**: 151–164.

Labein. 2001. *Premie Assessment Guide*. Fundacion Labein, Parque Tecnologico, ED. 101, Zamudio (Bizkaia), Spain.

Lopez de Mantaras R. 1990. *Approximate Reasoning Models*. Ellis Horwood Limited: Chichester, England.

Porter L, Tanner J. 1998. *Assessing Business Excellence*. Butterworth-Heinmann: London.

Saaty TL. 1988. *The Analytic Hierarchy Process*. University of Pittsburgh: Pittsburgh.

Sen P, Yang JB. 1995. An investigation into the influence of preference modelling in ship design with multiple objectives. *Proceedings of the Sixth International Symposium on Practical Design of Ships and Mobile Units*, Seoul, Korea, 17–22 September.

Siow CHR, Yang JB, Dale BG. 2001. A new modelling framework for organisational self-assessment: development and application. *Quality Management Journal* **8**(4): 34–47.

Sonmez M, Yang JB, Holt GD. 2001. Addressing the contractor selection problem using an evidential reasoning approach. *Engineering, Construction and Architectural Management* **8**(3): 198–210.

Sonmez M, Graham G, Yang JB, Holt GD. 2002. Applying the evidential reasoning approach to prequalifying construction contractors. *Journal of Management in Engineering* **18**(3): 111–119.

Sonmez M. 2002. Investigation into utility-based methods and evidential reasoning approach for complex decision analysis. *Ph.D. Thesis*, Manchester School of Management, UMIST.

Stewart TJ. 1992. A critical survey on the status of multiple criteria decision making theory and practice. *OMEGA International Journal of Management Science* **20**(5–6): 569–586.

UKEF. 2000. *Business Excellence Through Action* (*BETA*^{TM}). UK Excellence Foundation.

van der Wiele T, Williams ART, Kolb F, Dale BG. 1995. Assessor training for the European Quality Award. *Quality World Technical Supplement*, March: 12–18.

Wang J, Yang JB, Sen P. 1996. Multi-person and multi-attribute design evaluations using evidential reasoning based on subjective safety and cost analysis. *Reliability Engineering and System Safety* **52**: 113–127.

Wang J, Yang JB, Sen P. 1995. Safety analysis and synthesis using fuzzy sets and evidential reasoning. *Reliability Engineering and System Safety* **47**(2): 103–118.

Xu DL, McCarthy G, Yang JB, Greatbanks R. 2003. Computer supported self-assessment of business innovative capability. *Decision Support Systems*, submitted.

Xu DL, Yang JB. 2001. Introduction to multi-criteria decision making and the evidential reasoning approach. *Working Paper Series No. 0106*, Manchester School of Management, University of Manchester Institute and Technology, ISBN: 1 86115 111 X, May.

Yang JB, Xu DL. 2002a. On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. *IEEE Transactions on Systems, Man and Cybernetics Part A: Systems and Humans* **32**(3): 289–304.

Yang JB, Xu DL. 2002b. 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* **32**(3): 376–393.

Yang JB. 2001. Rule and utility based evidential reasoning approach for multiple attribute decision analysis under uncertainty. *European Journal of Operational Research* **131**(1): 31–61.

Yang JB, Dale BG, Siow CHR. 2001. Self-assessment of excellence: an application of the evidential reasoning approach. *International Journal of Production Research* **39**(16): 3789–3812.

Yang JB, Xu DL. 2000. Intelligent Decision System via Evidential Reasoning. *Version 1.1*. Cheshire, England.

Yang JB, Xu DL. 1998. Knowledge based executive car evaluation using the evidential reasoning approach. In *Advances in Manufacturing Technology*, vol. XII, Baines, Taleb-Bendiab, Zhao (eds). Professional Engineering Publishing: London; 741–749.

Yang JB, Sen P. 1996. Preference modelling by estimating local utility functions for multiobjective optimisation. *European Journal of Operational Research* **95**(1): 115–138.

Yang JB, Sen P. 1994. A general multi-level evaluation process for hybrid MADM with uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics* **24**(10): 1458–1473.

Yang JB, Singh MG. 1994. An evidential reasoning approach for multiple attribute decision making with uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics* **24**(1): 1–18.

Zhang ZJ, Yang JB, Xu DL. 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 (1990)).

