CBR Supports Decision Analysis with Uncertainty

Ning Xiong and Peter Funk

School of Innovation, Design and Engineering
Mälardalen University
SE-72123 Västerås, Sweden
{Ning.Xiong,Peter.Funk}@mdh.se

Abstract. This paper proposes a novel approach to case-based decision analysis supported by case-based reasoning (CBR). The strength of CBR is utilized for building a situation dependent decision model without complete domain knowledge. This is achieved by deriving states probabilities and general utility estimates from the case library and the subset of cases retrieved in a situation described in query. In particular, the derivation of state probabilities is realized through an information fusion process which comprises evidence (case) combination using the Dempster-Shafer theory and Bayesian probabilistic reasoning. Subsequently decision theory is applied to the decision model learnt from previous cases to identify the most promising, secured, and rational choices. In such a way we take advantage of both the strength of CBR to learn without domain knowledge and the ability of decision theory to analyze under uncertainty. We have also studied the issue of imprecise representations of utility in individual cases and explained how fuzzy decision analysis can be conducted when case specific utilities are assigned with fuzzy data.

Keywords: Case-based decision analysis, case-based reasoning, decision model, similarity, basic probability assignment, information fusion.

1 Introduction

Decision making is prevalent in solving many engineering, health care and management problems. It has also gained increasing importance for intelligent agent systems [1] to interact with the environment autonomously. The main challenges in most practical decision problems are how to cope with uncertain characteristics in the environment and how to make choices in the presence of these uncertain features. Decision theory [2, 3] has offered useful tools in analyzing uncertain situations to identify the "best" course of actions from a reasonable perspective. However, practical applications of decision theory entail formulating a real world problem into a perfect decision model, which may be hard to achieve in many circumstances due to complexity, poor domain knowledge, as well as incomplete information.

A more pragmatic method to make decisions is to visit previous similar situations as reference. It was argued in [4] that decision making under uncertainty is at least partly case-based. With a case-based method we don't require fully understood

L. McGinty and D.C. Wilson (Eds.): ICCBR 2009, LNAI 5650, pp. 358–373, 2009.

domain knowledge for building a precise decision model. The research into this realm is strongly supported by the methodology of case-based reasoning (CBR) [5]. Recently CBR has been widely employed as decision support for explanation [6, 7], label ranking [8], as well as recommendation and advice giving [9-13] in numerous practical applications.

This paper proposes a novel approach to support decision analysis using CBR methods. The power of CBR is utilized for creating a situation dependent decision model from past similar experiences. Further, decision theory is applied to the decision model learnt from past experiences to find out optimal, rational, and low risk solutions. With such integration we create a unified framework in which CBR and decision theory can complement each other. CBR helps decision analysis dealing with complicated problems with poor domain knowledge and incomplete information, while decision theory helps CBR handling uncertain information and features in the problem domain.

The kernel of the proposed work is the case-based learning of a decision model. This is a bit different from the common practice in many CBR systems where finding solutions to the query case appears the main goal of the CBR task. What we seek here is to derive, from previous experiences, a probabilistic characterization of the current situation in terms of likelihoods, risks and probable consequences. We hope this would offer a useful means to tackle the inherent nature of uncertainty in a CBR process, in particular when similar situations don't have similar solutions.

The paper is organized as follows. Section 2 outlines the proposed approach for case-based decision analysis at a general level. We explain derivation of state probabilities for a query situation in section 3, which is followed by estimation of general utilities of actions under states in section 4. Then, in section 5, we discuss decision analysis based on a decision model learnt from cases. Section 6 presents some related work. Finally this paper is concluded in section 7.

2 Case-Based Decision Analysis: The Proposed Approach

This section outlines the proposed approach for case-based decision analysis. We start with basics about the decision tree as a decision model. We shall then present the general idea of creating a decision tree from cases to support decision analysis.

2.1 Decision Model for Decision Analysis

The decision problem for an agent can be abstracted as follows. Given an environment with possible states s_1, s_2, …,s_n, the agent has to make a choice from a set of alternative actions $\{a_1, a_2, \ldots, a_m\}$. The outcome or consequence of an action is dependent on the real state of the environment. A general utility function has been defined for all possible outcomes regarding actions and states. By u_{ij} we denote the general utility of performing action a_i when state s_j is true, i.e., $u_{ij} = U(a_i \mid s_j)$. But the agent has no exact knowledge about the state of the environment, only a probability distribution of the states is available for decision analysis.

This (decision) problem can also be modelled as a decision tree as shown in Fig. 1, where p_i refers to the probability of state s_i ($i=1\ldots n$). The availability of such a model is prerequisite to apply well founded decision analysis methods such as Bayesian decision theory [2] and the principle of general risk constraints [14] for making profitable, secured, and rational choices

However, constructing a perfect decision tree to abstract an underlying situation is not trivial. It requires thorough understanding of the circumstance and detailed domain knowledge for elicitation of all relevant information. In many cases it is hard to define accurate values for probabilities concerning states of the environment and general utilities regarding actions and states in a decision tree. First of all, estimates for probabilities of states are very likely to be subjective or imprecise. It was observed in [15] that most people usually can not distinguish between probabilities roughly ranging from 0.3 to 0.7. Moreover, general utilities regarding actions and states correspond to a sort of generalized information which is hard to explicate without deep domain knowledge. Instead of giving utility in a general sense, users in real life would feel more natural and confident to specify individual utility scores associated with specific cases by evaluation of concrete results therein. Later we will show in the paper how both the state probabilities and the (general) utilities in the decision tree can be estimated from previous cases for a new situation by using a case-based approach.

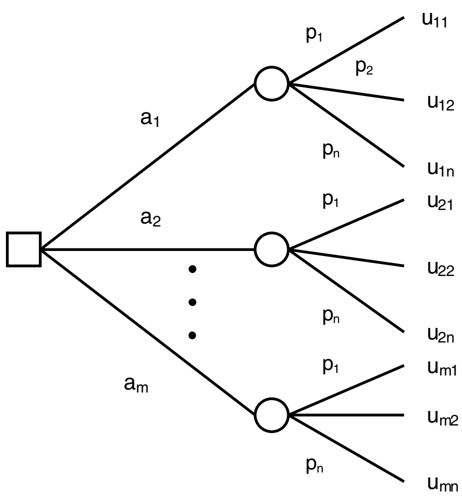

Fig. 1. A decision problem modelled as a decision tree

2.2 Case-Based Learning of Decision Trees

We consider decision trees as vehicles for carrying knowledge and information about candidate actions and their probable consequences. The content of the vehicle is situation dependent. In different situations we may have different alternatives, varying probabilities and different consequences. Here we propose a case-based approach to creating situation dependent decision trees. The basic idea is to derive the right content of the decision model by resorting to previous similar cases with respect to a given new situation. This approach is different from conventional ways CBR works to recommend final solutions based on a subset of retrieved cases. Contrarily, in this paper, we apply CBR in an intermediate stage for creation of a qualified decision model, which can then be utilized by various decision analysis methods to find out rational, justified choices.

A procedure for case-based learning of decision trees is shown in Fig. 2. It starts with similarity matching between a new situation and previous cases in the case library. Every case in the case library receives a similarity score according to a predefined similarity metric. We will not detail the issue of similarity measures due to the scope of this paper, but interested readers can refer to the references [16-19] for recent advancements of similarity modelling in CBR research. After similarity matching, a subset of cases that get the highest similarity scores or pass a specified similarity threshold are selected and retrieved. In the next step, we perform probability and utility derivation based on the subset of retrieved cases and the case library. The purpose is to exploit the information residing in the cases to acquire probabilities of environment states in the current situation as well as (general) utility estimates of alternative actions given different states. Finally, the derived probability and utility values are entered into the decision tree for decision analysis.

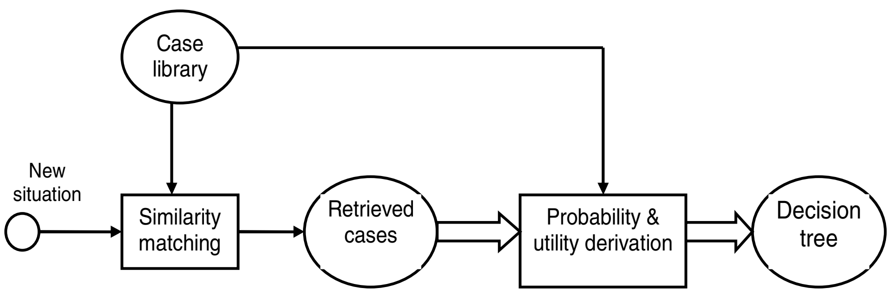

Fig. 2. Case-based learning of decision trees

As basic notation, we assume that a case C_j in the case library is indexed by a 4-tuple $C_j=(B_j, E_j, A_j, U_j)$, where

• B_j is the description of the situation associated with the case. It can, for instance, consist of a set of observed or user-acquired attribute values.
• $E_j=(P_j(s_1), P_j(s_2), \ldots, P_j(s_n))$ represents the known probability distribution for states $s_1, s_2, \ldots, s_n$ in the situation associated with the case. States are usually not observable but reflect internal properties of the environment. Sometimes the probability of a state in a case is also notated as $P_j(s_i) = P(s_i \mid C_j)$.
• A_j denotes an action that was performed in the situation associated with the case.
• U_j is an individual utility score evaluating the outcome of performing action A_j in the situation associated with the case. Hence it is also notated as $U(A_j|C_j)$ later in the paper.

3 Deriving State Probabilities from Previous Cases

The procedure for deriving probabilities of states based upon available cases is depicted in Fig. 3. Given a new target situation *Q*, we look for its similar cases in the case library and a subset of cases is retrieved according to the rule of *KNN* (*k* nearest neighborhoods) or a specified similarity threshold. The retrieved cases are then delivered along with their similarity degrees to the block "information fusion" for assessing the probabilities of states in the new situation *Q*. The information fusion block is further divided into two successive steps, as will be described in subsections 3.1 and 3.2 respectively. The first step concerns evidence combination using the Dempster-Shafer theory (simply D-S theory) [20-21] to yield initial beliefs in states. The D-S theory enables distinguishing different cases in the information fusion process according to their similarity degrees. The second step aims to refine these initial beliefs into final probability evaluations via probabilistic reasoning.

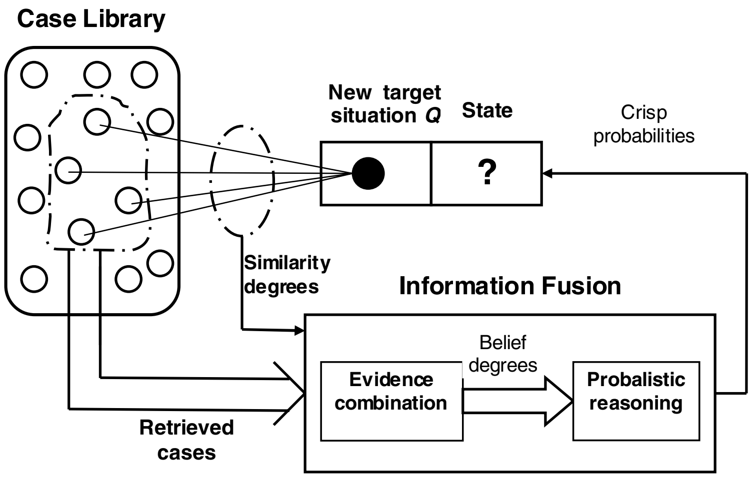

Fig. 3. Derivation of state probabilities based on cases

3.1 Reasoning Degrees of Belief Using the D-S Theory

We consider every retrieved case as a source of information. The evidence combination rule of the Dempster-Shafer theory is employed to aggregate information from relevant cases for assessing the degrees of beliefs in possible states in the query situation.

3.1.1 Evidence Combination Rule of the D-S Theory

The D-S theory is a powerful tool tackling uncertainty. But we do not intend to have an extensive discussion of it in this paper. We shall only introduce some basic concepts of this theory that are relevant for our task of belief aggregation from multiple cases.

In the D-S theory, a sample space of the problem domain is called a "frame of discernment", notated as X. It is assumed that one's total belief due to a piece of evidence can be partitioned into various probability masses, each assigned to a subset of X. These probability masses are specified by basic probability assignment (BPA), which is a function m performing mapping from the power set of X to the interval [0, 1] satisfying:

$$m(\emptyset) = 0 \tag{1}$$

$$\sum_{F \subseteq X} m(F) = 1 \tag{2}$$

In particular the subsets F of X such that $m(F) > 0$ are called the focal elements of the D-S belief structure.

Owing to imprecision of information, we can not figure out exact probability values for arbitrary subsets of X from a BPA function. The following two measures are therefore introduced to impose bounds on the probability of a hypothesis.

Let hypothesis F be a subset of X, the belief of F, denoted $Bel(F)$, is defined as

$$Bel(F) = \sum_{G \subseteq F} m(G) \tag{3}$$

The plausibility of F, denoted $Pl(F)$, is defined as

$$Pl(F) = \sum_{G \cap F \neq \emptyset} m(G) \tag{4}$$

It was shown in [22] that, for any subset F of X, we have the inequality below

$$Bel(F) \leq P(F) \leq Pl(F) \tag{5}$$

This reads that the belief and plausibility measures provide lower and upper bounds on the probability of a hypothesis.

Suppose there are two bodies of evidences over the same frame of discernment, but induced from independent information sources. The BPA functions associated with the two bodies of evidences are m_1 and m_2 respectively. The task now is to combine the evidence related functions m_1 and m_2 into an aggregated (basic) probability assignment function $m_{12} = m_1 \oplus m_2$. According to the evidence combination rule of the D-S theory, the basic probability mass for a hypothesis F ($F \subseteq X$), incorporating both pieces of evidences, is calculated as follows:

$$m_{12}(\emptyset) = 0, \qquad m_{12}(F) = K\left(\sum_{F_1 \cap F_2 = F} m_1(F_1) m_2(F_2)\right) \tag{6a}$$

$$K = \left(1 - \sum_{F_1 \cap F_2 = \emptyset} m_1(F_1) m_2(F_2)\right)^{-1} \tag{6b}$$

The above combination rule reads that $m_{12}(F)$ is calculated from the summation of the products $m_1(F_1)m_2(F_2)$ where the intersection of F_1 and F_2 equals F. The quantity K plays the role of normalization such that the sum of basic probability numbers for all subsets of X equals one. K is computed on all pairs of F_1 and F_2 that have no intersections with each other. Next we shall use this combination rule to estimate the degrees of belief in various states based on the cases retrieved from the case library.

3.1.2 Combining Retrieved Cases as Evidences

Suppose Nr cases are retrieved from the case library after similarity matching. Without loss of generality, we denote the set of retrieved cases by

$$E = \{C_1, C_2, \cdots, C_{Nr}\} \tag{7}$$

The similarity degrees of these retrieved cases against the query situation are given by $Sim = \{\alpha_1, \alpha_2, \cdots, \alpha_{Nr}\}$ where α_j represents the degree of similarity of case C_j. Our task here is to aggregate the information of the cases in E to acquire combined degrees of belief in various states in the current situation.

Obviously the frame of discernment, X, in our problem domain is the set of states in the environment. In order to apply the evidence combination rule stated above, we first have to interpret the probability distributions in individual cases into a form complying with the D-S belief structure. This can be easily done by restricting the focal elements of the belief structure to individual states as singleton subsets of X. Hence the probability distribution in a case C_j can be interpreted as a basic probability assignment function written as

$$BP(C_j) = \{(s_i, P_j(s_i)), i = 1 \cdots n\} \tag{8}$$

where s_i denotes a state in the environment and $P_j(s_i)$ is the probability that state s_i is true in the situation described by case C_j.

Now consider the basic probability assignment function that is induced by the evidence of a retrieved case C_j. Let $m(i, j)$ be the basic probability value to which the hypothesis that state s_i is true is supported by case C_j as evidence. This probability mass should be reduced from function (8) with similarity degree α_j as discounting factor. Hence we have

$$m(i, j) = \alpha_j \cdot P_j(s_i) \quad i = 1, \cdots, n; \quad j = 1, \cdots, Nr \tag{9}$$

As the sum of basic probabilities of states is now smaller than one according to (9), we introduce an extra subset S containing all possible states. The subset S receives the remaining probability mass unassigned to any individual state. Thus we can write

$$m(S, j) = 1 - \sum_{i=1}^{n} m(i, j) = 1 - \sum_{i=1}^{n} \alpha_j P_j(s_i) = 1 - \alpha_j, \quad j = 1, \cdots, Nr \tag{10}$$

Having established basic probability assignments induced by retrieved cases, we now attempt to aggregate these assignment functions into an overall assessment using the evidence combination rule. Denote E_t as the set of the first t retrieved cases as follows:

$$E_t = \{C_1, C_2, \cdots, C_t\} \tag{11}$$

Let $m(i, E_t)$ be the basic probability mass to which the hypothesis that state s_i is true is supported by all evidences (retrieved cases) in E_t. By $m(S, E_t)$ we denote the remaining probability mass unassigned to individual states after all evidences in E_t

have been combined. The algorithm to fuse case information according to the evidence combination rule can be formulated in a recursive form as follows:

$$m(i,E_{t+1}) = K_{t+1}\big(m(i,E_t)m(i,t+1) + m(i,E_t)m(S,t+1) + m(S,E_t)m(i,t+1)\big) \quad (12a)$$
$$i = 1 \cdots n$$

$$m(S,E_{t+1}) = K_{t+1}m(S,E_t)m(S,t+1) \quad (12b)$$

$$K_{t+1} = \left(1 - \sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n} m(i,E_t)m(j,t+1)\right)^{-1} \quad t = 1 \cdots Nr-1 \quad (12c)$$

where K_{t+1} is a normalizing factor to make the sum of the basic probability values induced by the evidences in E_{t+1} equal one. It bears noting that, to start with the above recursive form, we have $m(i,E_1) = m(i,1)$ and $m(S,E_1) = m(S,1)$. The final outcomes of this combination procedure are $m(i, E_{Nr})$ and $m(S, E_{Nr})$, which correspond to the basic probability values after incorporating all retrieved cases as evidences.

In terms of the belief function defined in (3), the probability mass $m(i, E_{Nr})$ also represents the degree of belief in state s_i after considering all retrieved cases. Hence the combined degrees of belief are directly given by

$$\beta_i = m(i,E_{Nr}) \qquad i = 1 \cdots n \quad (13a)$$

$$\beta_S = m(S,E_{Nr}) = 1 - \sum_{i=1}^{n} \beta_i \quad (13b)$$

where β_S refers to the degree of belief unassigned to any individual state after all retrieved cases have been incorporated. It indicates a degree of ignorance or incompleteness of information in the generated assessment.

Further, from the plausibility definition in (4), the value of plausibility for the hypothesis that state s_i is true is equal to $\beta_i+\beta_S$. It is the upper bound of the likelihood for the truth of state s_i. The lower bound for the likelihood of state s_i is reflected by the belief degree β_i. In other words, we obtain the interval $[\beta_i, \beta_i+\beta_S]$ as estimate of the probability for state s_i (i=1…n) by using the D-S combination rule. In the next subsection we shall discuss how to refine these initial estimates to obtain crisp probability values of states by doing probabilistic reasoning.

3.2 Reaching Final Probabilities via Probabilistic Reasoning

The probability intervals derived from the D-S rule can be refined via probabilistic reasoning. Without any prior knowledge, the initial probability $P_0(s_i)$ for a state is defined by equally distributing the unassigned probability β_S among all states. Thus we have

$$P_0(s_i) = \beta_i + \frac{\beta_S}{n} \qquad i = 1 \cdots n \quad (14)$$

Then we perform probability updating based on the Bayes theorem.

As cases in the case base were collected independently of each other, we utilized the similar relation with every retrieved case, $sr(C_j)$, as an independent observation to update the prior probabilities according the Bayes theorem. Define H_j as the set of the first j observations (similar relations) as follows:

$$H_j = \{sr(C_1), sr(C_2), \cdots, sr(C_j)\} \tag{15}$$

The Bayesian reasoning for probabilities of states can be summarized in a recursive form by

$$\begin{aligned} P(s_i \mid H_{j+1}) &= \frac{P(s_i \mid H_j) \cdot P(sr(C_{j+1}) \mid H_j, s_i)}{\sum_{k=1}^{n} P(sr(C_{j+1}) \mid H_j, s_k) \cdot P(s_k \mid H_j)} \\ &= \frac{P(s_i \mid H_j) \cdot P(sr(C_{j+1}) \mid s_i)}{\sum_{k=1}^{n} P(sr(C_{j+1}) \mid s_k) \cdot P(s_k \mid H_j)} \quad j = 0, \cdots, Nr-1 \end{aligned} \tag{16}$$

Note that we have $P(s_i \mid H_0) = P_0(s_i)$ to start with this recursive form.

It can be seen from Eq. (16) that, to update probabilities of states, we need the conditional probability $P(sr(C_j)|s_i)$ for all the retrieved cases C_j (j=1…Nr). Such probability can be regarded as the likelihood of a randomly picked case from the case base being similar to C_j provided that the state in this case is known as s_i. Hence we have

$$P(sr(C_j) \mid s_i) = \sum_{\substack{C \in Case\ \ Base \\ C\ \ similar\ \ to\ \ C_j}} P(C \mid s_i) \quad i = 1 \cdots n, \quad j = 1 \cdots Nr \tag{17}$$

Further, we apply the Bayes theorem and transform the probability $P(C|s_i)$ to the following form:

$$P(C \mid s_i) = \frac{P(C) \cdot P(s_i \mid C)}{\sum_{C_k \in Case\ \ base} P(s_i \mid C_k) \cdot P(C_k)} \quad i = 1 \cdots n \tag{18}$$

Since we assume all cases in the case library are equally probable to be selected, Eq. (18) is simplified to

$$P(C \mid s_i) = \frac{P(s_i \mid C)}{\sum_{C_k \in Case\ \ base} P(s_i \mid C_k)} = \frac{P(s_i \mid C)}{\sum_{\forall k} P_k(s_i)} \quad i = 1 \cdots n \tag{19}$$

At this point, it has been obvious that we can calculate the probability $P(C|s_i)$ by using probabilistic information stored in individual cases in the case library, which further enables updating state probabilities in terms of Eqs. (16) and (17).

However, one disadvantage of the above calculation with Bayes theorem is that different similarities (thereby importances) of the cases are not taken into account. For more accurate results, we are not directly adopting such assessment as final probability values. Instead we utilize the probability values yielded from Bayesian reasoning as factors to divide the unassigned probability β_S across various states. This means that every state s_i receives an additional probability mass from β_S in proportion to $P(s_i|H_{Nr})$. This additional mass is then added to the lower bound of probability, β_i, to settle the final probability assessment. In other words, after information fusion in two steps, the probability for state s_i is finalized as

$$P(s_i) = \beta_i + P(s_i \mid H_{Nr}) \cdot \beta_S \qquad i = 1 \cdots n \tag{20}$$

4 Derivation of General Utilities of Actions Given States

The basic idea is to derive the general utility of performing one action under a given state by using information from the case library. However, owing to the fact that no exact information is known about states in cases, case specific utilities recorded can not provide direct answers to our inquiries. As an alternative, we here attempt to estimate this utility with an expected value by considering all those cases in which the underlying action was performed. By *Sub*(*a*) we denote the subset of cases in the case library in which the action *a* was performed. Then the expected value of the general utility of performing action *a* given state s_i can be given by:

$$U(a \mid s_i) = \sum_{C_t \in Sub(a)} U(a \mid C_t) \cdot P_a(C_t \mid s_i) \tag{21}$$

As $U(a|C_t)$ represents the known utility recorded in case C_t, what remains to resolve is the probability $P_a(C_t|s_i)$. By employing the Bayes theorem, this probability is reformulated as

$$P_a(C_t \mid s_i) = \frac{P_a(C_t) \cdot P(s_i \mid C_t)}{\sum_{C_k \in Sub(a)} P_a(C_k) \cdot P(s_i \mid C_k)} \tag{22}$$

Considering that cases in the subset *Sub*(*a*) are equally probable to be picked up, Eq. (22) is reduced to

$$P_a(C_t \mid s_i) = \frac{P(s_i \mid C_t)}{\sum_{C_k \in Sub(a)} P(s_i \mid C_k)} \tag{23}$$

Since $P(s_i|C_k)$ is available as the probability of state s_i in case C_k, we easily resolve Eq. (23), leading to computation of the expected value of the general utility according to Eq. (21). This expected value then enters the decision tree as estimation of the (general) utility of action *a* given state s_i.

5 Decision Analysis Using Case-Based Decision Model

Once a decision model is constructed from cases, it can be applied to analyse and evaluate alternative actions in the current situation, taking into account both likelihoods and probable consequences. We will first introduce a well established principle for doing such analysis of decisions, followed by discussions of how this basic principle can be applied in circumstances when utility values specified in individual cases are fuzzy or imprecise.

5.1 Principle of Maximizing Expected Utility

With complete information in the decision tree derived, we can now compute the expected utility of the various alternative actions. The expected utility of action a_j is defined as

$$EU(a_j) = P(s_1)\cdot U(a_j \mid s_1) + P(s_2)\cdot U(a_j \mid s_2) + \cdots + P(s_n)\cdot U(a_j \mid s_n) \qquad (24)$$

where $P(s_i)$ and $U(a_j|s_i)$ represent the probability and (general) utility values derived from the retrieved cases and the case library respectively. Then a choice should be made among the alternatives according to the principle of maximizing the expected utility [2], which is formulated as follows:

The principle of maximizing expected utility (MEU): In a given decision situation the deciding agent should prefer the alternative with maximal expected utility. That means that alternative a_1 is preferred to a_2 if and only if $EU(a_1) > EU(a_2)$.

The expected utility of an action approximates the mean utility score that will be obtained if an agent or decision maker meets the situation many times and chooses and conducts the same action constantly. In view of this, the significance of the MEU principle is to optimize the long term performance of decision making under uncertainty.

The merit of doing decision analysis after CBR can be illustrated with the following example. Assume that, given a target situation, two cases C_1 and C_2 are retrieved from the case base and they have actions a_1 and a_2 respectively. Both cases are assigned with good utility values as evaluations of their outcomes, but case C_1 is more similar to the target situation. Then, according to CBR alone, action a_1 associated with case C_1 will be judged more suitable as solution to the new situation. Nevertheless, if we further consider more information in the decision tree, we might change our preference after decision analysis.

For instance, suppose that the state probabilities and utilities of actions under possible states (s_1 and s_2) are derived from previous cases as follows:

$$P(s_1 \mid sr(C_1), sr(C_2)) = 0.6 \qquad U(a_1 | s_1) = 70 \qquad U(a_2 | s_1) = 40$$
$$P(s_2 \mid sr(C_1), sr(C_2)) = 0.4 \qquad U(a_1 | s_2) = -90 \qquad U(a_2 | s_2) = 60$$

The expected utilities of a_1 and a_2 are calculated as $EU(a_1)$=6 and $EU(a_2)$=48 respectively in the current situation. Hence we will prefer action a_2 according to the MEU principle. We believe that a_2 is a more rational choice considering the high risk of action a_1 under state s_2. This rational choice is achieved by taking advantage of the case-based decision tree which accommodates more information than case similarity alone.

5.2 When Utility Values from Cases Are Fuzzy

Until now we have assumed that a crisp utility value is assigned in every case in the case library. However, in numerous practical applications, it is frequently difficult for human users and even domain experts to assign an exact utility as their evaluation of the consequences. They would be more likely to say that the outcome in a case should receive a utility score of, say, around 60. Here "*around 60*" is an imprecise value and, in terms of fuzzy set theory [23], it is considered as a fuzzy number. Our intention is to extend the representation to allow for users to specify vague, fuzzy data as evaluations of utility in specific cases. But, by doing this, we will not exclude the possibility that users assign crisp utility values if they prefer. Considering that crisp numbers are special singleton fuzzy numbers, this extension would bring a useful generalization of the theory and methods making our framework applicable to more general types of data and information.

In fuzzy set theory, a fuzzy number is a fuzzy subset of R that is convex and normal. So, in principle, users can define any convex and normal fuzzy subset of R as fuzzy utility in a specific case. But, for reducing computational complexity, we would prefer to recommend triangular fuzzy numbers which are intuitive, simple, and easy to manipulate. A triangular fuzzy number F can be depicted by a 3-tuple: $F = (f_1, f_2, f_3)$, with its membership function being illustrated in Fig. 4.

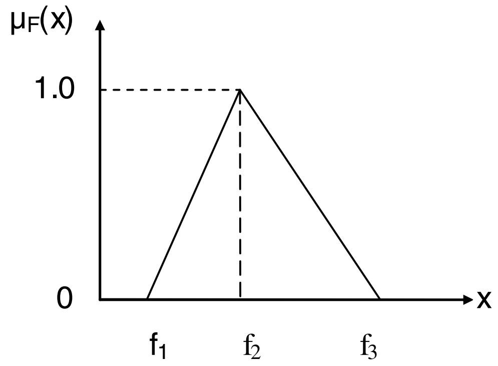

Fig. 4. A triangular fuzzy number

Two nice properties of triangular fuzzy numbers are that the addition of two triangular fuzzy numbers is still a triangular fuzzy number and that the multiplication of a constant with a triangular fuzzy number is still a triangular fuzzy number [24]. That is to say that, given two triangular fuzzy numbers $F = (f_1, f_2, f_3)$, and $G = (g_1, g_2, g_3)$, and a constant γ, we have

$$F + G = (f_1 + g_1, f_2 + g_2, f_3 + g_3) \tag{25}$$

$$\gamma \cdot F = (\gamma \cdot f_1, \gamma \cdot f_2, \gamma \cdot f_3) \tag{26}$$

Owing to the properties depicted in Eqs. (25) and (26), we clearly see that the general utility estimation in (21) and the expected utility of actions in (24) are also triangular fuzzy numbers as long as utilities in individual cases are specified as triangular fuzzy numbers. Consequently, evaluating alternatives according to the EMU principle turns to studying the fuzzy dominance relations between the fuzzy expected utilities represented as fuzzy numbers. This task is not trivial in the sense that a natural order does not exist with the quantities as fuzzy numbers.

Let F_i and F_j be two fuzzy numbers corresponding to the fuzzy expected utilities for alternative actions a_i and a_j respectively. The fuzzy relation that a_i is dominated by a_j (or F_i is dominated by F_j) is defined by the degree of possibility and the degree of necessity of the event $F_i < F_j$, which are given by

$$\Pi(F_i < F_j) = \sup_{x<y} \min\left(\mu_{F_i}(x), \mu_{F_j}(y)\right) \tag{27}$$

$$N(F_i < F_j) = 1 - \sup_{x \geq y} \min\left(\mu_{F_i}(x), \mu_{F_j}(y)\right) \tag{28}$$

Further, we investigate to what extent an action is a dominated one. Since the statement that a_i is dominated becomes true if a_i is dominated by at least one of the other alternatives, we apply an *s*-norm as logical disjunction to connect the dominance relations between a_i and the others. If the maximum operator is adopted as the means for *s*-norm, the degrees of possibility and necessity of alternative a_i being dominated are respectively defined as

$$Poss(a_i) = \max_{i \neq j}\left(\Pi(F_i < F_j)\right) \tag{29}$$

$$Nec(a_i) = \max_{i \neq j}\left(N(F_i < F_j)\right) \tag{30}$$

By means of the possibility and necessity values given in (29) and (30), we actually have defined a fuzzy subset of dominated alternative actions. Finally, we define an α-β-cut of this fuzzy subset to reach a crisp subset *DOM*. The membership function of the crisp subset *DOM* is given by

$$\mu_{DOM}(a_i) = \begin{cases} 1 & Poss(a_i) \geq \alpha \quad and \quad Nec(a_i) \geq \beta \\ 0 & otherwise \end{cases} \tag{31}$$

where α, β ($0 < \beta < \alpha < 1$) are parameters controlling the number of dominated actions. The remaining alternatives are subsequently recommended to the decision maker as non-dominated solutions.

6 Related Work

A probabilistic model for CBR was first proposed in [25]. The basic idea presented there is to consider the CBR principle that similar problems have similar solutions as a “rule of thumb” rather than a universally valid rule. According to this probabilistic model, the conventional CBR principle can be reformulated into a heuristic rule stating that similar problems are at most likely to have similar solutions. The merit of this formulation is that it allows for exceptions to the CBR rule.

Later, a similarity-based inference scheme [26] was developed from the CBR probabilistic model [25] by the same author. The method is to represent information from relevant cases into belief functions for characterizing confidence of alternative solutions for a new problem at hand. Then the belief functions from individual cases are combined in the framework of information fusion. This method can be useful in the overall problem solving process by measuring different confidence levels of different candidate solutions.

We proposed a framework for case-based decision analysis in [27], in which the probabilistic information from individual cases was integrated solely with the Bayes theorem. The weakness is that it can not take into account the different degrees of similarity of retrieved cases in probabilistic calculation. This problem is overcome with the work presented here by using the D-S theory for evidential combination with respect to cases. Our work differs from [25] and [26] in that it does not directly evaluate solutions and their confidences. Instead it aims to produce an intermediate decision model by fusing information from cases to highlight all possibilities and consequences. Decision analysis can then be conducted on the derived decision model to identify the most promising solution in view of expected outcomes.

7 Conclusion

This paper presents a new framework for case-based decision analysis supported by CBR. We claim that CBR and decision theory can complement each other in a coherent, hybrid system. CBR has the strength of creating a situation dependent decision model without domain knowledge. This is achieved by deriving states probabilities and general utility estimates from previous cases through an information fusion process comprising evidence combination and probabilistic reasoning. It follows that more accurate and objective data will be available in the decision model, promoting more reliable results of decision analysis. On the other hand, decision theory helps CBR better tackling the uncertainty issue by considering all probable consequences, risks, and likelihoods rather than similarity of cases alone. This would endow the agent or decision maker with more complete awareness of the situation and environment for making predictive, secured and rational choices. Besides, we have shown that fuzzy numbers can be used to represent case specific utility values for decision analysis and that fuzzy and probabilistic information are well utilized together in our case-based framework.

In future we will apply our approach to support decision making in strategic maintenance scenarios in industry. Therein a machine or production line under investigation can be considered as the environment, and evaluation grades or faults of the machine refer to the internal states of the environment. We plan to not only assess probable grades for machines but also carry out decision analysis based on previous cases to find out effective, rational, low risk counter-measures (maintenance plans, repair alternatives, etc.) as decision support.

References

1. Weiss, G.: Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. MIT Press, Cambridge (1999)
2. Gärdenfors, P., Sahlin, N.E.: Introduction: Bayesian Decision Theory – Foundations and Problems. In: Decision, Probability, and Utility, pp. 1–15. Cambridge University Press, Cambridge (1997)

3. Raiffa, H.: Decision Analysis: Introductory Readings on Choices under Uncertainty. McGraw Hill, New York (1997)
4. Gilboa, I., Schmeidler, D.: Case-Based Decision Theory. The Quarterly Journal of Economics 110, 605–639 (1995)
5. Aamodt, A., Plaza, E.: Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. Artificial Intelligence Com. 7, 39–59 (1994)
6. Doyle, D., Cunningham, P., Bridge, D., Rahman, Y.: Explanation Oriented Retrieval. In: Funk, P., González Calero, P.A. (eds.) ECCBR 2004. LNCS, vol. 3155, pp. 157–168. Springer, Heidelberg (2004)
7. Roth-Berghofer, T.R.: Explanations and Case-Based Reasoning: Foundational Issues. In: Funk, P., González Calero, P.A. (eds.) ECCBR 2004. LNCS, vol. 3155, pp. 389–403. Springer, Heidelberg (2004)
8. Brinker, K., Hullermeier, E.: Label Ranking in Case-Based Reasoning. In: Weber, R.O., Richter, M.M. (eds.) ICCBR 2007. LNCS, vol. 4626, pp. 77–91. Springer, Heidelberg (2007)
9. Coyle, L., Cunningham, P.: Improving Recommendation Ranking by Learning Personal Feature Weights. In: Funk, P., González Calero, P.A. (eds.) ECCBR 2004. LNCS, vol. 3155, pp. 560–572. Springer, Heidelberg (2004)
10. McSherry, D.: Completeness Criteria for Retrieval in Recommender Systems. In: Roth-Berghofer, T.R., Göker, M.H., Güvenir, H.A. (eds.) ECCBR 2006. LNCS, vol. 4106, pp. 9–29. Springer, Heidelberg (2006)
11. Nicholson, R., Bridge, D., Wilson, N.: Decision Diagrams: Fast and Flexible Support for Case Retrieval and Recommendation. In: Roth-Berghofer, T.R., Göker, M.H., Güvenir, H.A. (eds.) ECCBR 2006. LNCS, vol. 4106, pp. 136–150. Springer, Heidelberg (2006)
12. Stahl, A.: Combining Case-Based and Similarity-Based Product Recommendation. In: Roth-Berghofer, T.R., Göker, M.H., Güvenir, H.A. (eds.) ECCBR 2006. LNCS, vol. 4106, pp. 355–369. Springer, Heidelberg (2006)
13. Aamodt, A.: CBR for Advice Giving in a Data-Intensive Environment. In: Proceedings of 10th Scandinavian Conference on Artificial Intelligence, Stockholm, pp. 201–205 (2008)
14. Ekenberg, L., Boman, M., Linneroth-Bayer, J.: General Risk Constraints. Journal of Risk Research 4, 31–47 (2001)
15. Shapira, Z.: Risk Taking: A Managerial Perspective. Russel Sage Foundation, Thousand Oaks (1995)
16. Stahl, A., Gabel, T.: Using Evolution Programs to Learn Local Similarity Measures. In: Ashley, K.D., Bridge, D.G. (eds.) ICCBR 2003. LNCS, vol. 2689, pp. 537–551. Springer, Heidelberg (2003)
17. Cheng, W., Hullermeier, E.: Learning Similarity Functions from Qualitative Feedback. In: Althoff, K.-D., Bergmann, R., Minor, M., Hanft, A. (eds.) ECCBR 2008. LNCS, vol. 5239, pp. 120–134. Springer, Heidelberg (2008)
18. Gabel, T., Riedmiller, M.: Increasing Precision of Credible Case-Based Inference. In: Althoff, K.-D., Bergmann, R., Minor, M., Hanft, A. (eds.) ECCBR 2008. LNCS, vol. 5239, pp. 225–239. Springer, Heidelberg (2008)
19. Xiong, N., Funk, P.: Learning Similarity Metric Reflecting Utility in Case-Based Reasoning. Journal of Intelligent and Fuzzy Systems 17, 407–416 (2006)
20. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
21. Smets, P.: Belief Functions. In: Smets, P., Mamdani, E.H., Dubois, D., Prade, H. (eds.) Non-Standard Logics for Automated Reasoning, pp. 253–277. Academic Press, London (1988)

22. Dempster, A.P.: Upper and Lower Probabilities Induced by a Multi-Valued Mapping. Ann. Math. Stat. 38, 325–339 (1967)
23. Zadeh, L.A.: The Concept of a Linguistic Variable and Its Applications to Approximate Reasoning. Information Sciences 8, 199–249 (1975)
24. Kaufmann, A., Gupta, M.: Introduction to Fuzzy Arithmetic. Theory and Applications. Van Nostrand Reinhold Company Inc. (1985)
25. Hullermeier, E.: Toward a Probabilistic Formalization of Case-Based Inference. In: Proceedings of International Joint Conference on Artificial Intelligence, pp. 248–253 (1999)
26. Hullermeier, E.: Similarity-Based Inferences as Evidential Reasoning. In: Proceedings of 14th European Conference on Artificial Intelligence, pp. 55–59 (2000)
27. Xiong, N., Funk, P.: A novel framework for case-based decision analysis. In: Proceedings of 10th Scandinavian Conference on Artificial Intelligence, Stockholm, Sweden, pp. 141–148 (2008)