World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:2, No:9, 2008

Pruning Method of Belief Decision Trees

Salsabil Trabelsi
Salsabilt@voila.fr

Zied Elouedi
zied.elouedi@gmx.fr

Khaled Mellouli
khaled.mellouli@ihec.rnu.tn

LARODEC, Institut Supérieur de Gestion de Tunis, 41 Avenue de la Liberté,
2000 Le Bardo, Tunisie

Open Science Index, Computer and Information Engineering Vol:2, No:9, 2008 publications.waset.org/3379/pdf

***Abstract*—** The belief decision tree (BDT) approach is a decision tree in an uncertain environment where the uncertainty is represented through the Transferable Belief Model (TBM), one interpretation of the belief function theory. The uncertainty can appear either in the actual class of training objects or attribute values of objects to classify. In this paper, we develop a post-pruning method of belief decision trees in order to reduce size and improve classification accuracy on unseen cases. The pruning of decision tree has a considerable intention in the areas of machine learning.

Keywords: machine learning, uncertainty, belief function theory, belief decision tree, pruning.

I. Introduction

Decision trees are a simple yet successful technique for supervised classification learning. The visual presentation makes the decision tree model very easy to understand. It has also good classification accuracy compared to other classification techniques. However, the standard decision trees do not well perform their classification task in an environment characterized by uncertainty in data. In order to overcome this limitation, many researches have been done to adapt standard decision tree to this kind of environment. The idea was to introduce theories that could represent uncertainty. Several kinds of decision trees were developed: probabilistic decision trees [11], fuzzy decision trees [17], belief decision trees [2],[3] and possibilistic decision trees [7],[6]. In our work, we will focus on belief decision trees.

The belief decision tree approach is a decision tree technique adapted in order to handle uncertainty about the actual class of the objects in the training set and also to classify objects characterized by uncertain attributes. The uncertainty is represented by the Transferable Belief Model (TBM), one interpretation of the belief function theory.

The theory of belief functions is considered as a useful theory for representing and managing uncertain knowledge. It allows to express partial beliefs in a flexible way. Besides, it permits to handle partial or total ignorance concerning classification parametres.

When a belief decision tree is built from real world databases, many of branches will reflect noise in the training data due to uncertainty. The results are many of undesirable nodes and difficulty to interpret the tree. Our aim is to overcome this problem of overfitting in belief decision tree. In order to reduce the size of the tree and improve classification accuracy. Pruning is a way to cope with this problem. So, our objective in this work is to prune belief decision tree. "How does tree pruning work?" there are two common approaches to tree pruning. Methods that can control the growth of a decision tree during its development are called *pre-pruning* methods, the others are called post-pruning methods. In post-pruning approach, grow the full tree, allow it overfit the data and then post-prune it. It requires more computation than pre-pruning, yet generally leads to a more reliable tree.

In this work, we focused on post-pruning approach. Pre-pruning in belief decision tree has developed in [5] by improving the stopping criteria concerning the value of the selection measure. So, we suggest to develop post-pruning method to simplify the belief decision trees in order to reduce the size and the complexity. This paper is organized as follows: Section 2 provides a brief description of basics of belief function theory. In section 3, we describe the BDT approach. Then, in Section 4, we present the description of our pruning belief decision tree method. Finally in Section 5, we carry simulations to compare BDT without pruning, after pre-pruning and after our post-pruning method.

II. Belief function theory

In this section, we briefly review the main concepts underlying the belief function theory. This theory is appropriate to handle uncertainty in classification problems especially within the decision tree technique. In belief decision trees the uncertainty is represented through the Transferable Belief Model (TBM), one interpretation of the belief function theory.

A. Definitions

The TBM is a model to represent quantified belief functions [15]. Let Θ be a finite set of elementary events to a given problem, called the frame of discernment [14]. All the subsets of Θ belong to the power set of Θ, denoted by 2^{Θ}.

The impact of a piece of evidence on the different subsets of the frame of discernment Θ is represented by a basic belief assignment (bba).

Open Science Index, Computer and Information Engineering Vol:2, No:9, 2008 publications.waset.org/3379/pdf

The bba is a function $m : 2^{\Theta} \rightarrow [0, 1]$ such that:

$$\sum_{A \subseteq \Theta} m(A) = 1 \quad (1)$$

The value *m(A)*, named a basic belief mass (bbm), represents the portion of belief committed exactly to the event A.

Associated with m is the belief function, denoted *bel*, corresponding to a specific bba *m*, assigns to every subset A of Θ the sum of masses of belief committed to every subset of A by m [13].

The belief function *bel* is defined for $A \subseteq \Theta$, $A \neq \emptyset$ as:

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad (2)$$

The plausibility function *pl* quantifies the maximum amount of belief that could be given to a subset A of the frame of discernment. It is equal to the sum of the bbm's relative to subsets B compatible with A.

The plausibility function *pl* is defined as follows:

$$pl(A) = \sum_{A \cap B \neq \emptyset} m(B), \forall A \subseteq \Theta \quad (3)$$

B. Combination

In the transferable belief model, the basic belief assignments induced from distinct pieces of evidence are combined by either the conjunctive rule of combination or the disjunctive rule.

1) The conjunctive rule : When we know that both sources of information are fully reliable then the bba representing the combined evidence satisfies [16]:

$$(m_1 \textcircled{\cap} m_2)(A) = \sum_{B,C \subseteq \Theta : B \cap C = A} m_1(B) m_2(C) \quad (4)$$

2) The disjunctive rule : When we only know that at least one of sources of information is reliable but we do not know which is reliable, then the bba representing the combined evidence satisfies [16]:

$$(m_1 \textcircled{\cup} m_2)(A) = \sum_{B,C \subseteq \Theta : B \cup C = A} m_1(B) m_2(C) \quad (5)$$

C. Decision making

In the transferable belief model, holding beliefs and making decisions are distinct processes. Hence, it proposes two level models:

- *The credal level* where beliefs are represented by belief functions.
- *The pignistic level* where beliefs are used to make decisions and represented by probability functions called the pigninstic probabilities and is defined as:

$$BetP(A) = \sum_{B \subseteq \Theta} \frac{\mid A \cap B \mid}{\mid B \mid} \frac{m(B)}{(1 - m(\emptyset))}, \quad \text{for all} \quad A \in \Theta \quad (6)$$

III. Belief decision trees

A belief decision tree is a decision tree in an uncertain environment where the uncertainty is represented by the TBM. There are two methods to build the tree averaging and conjunctive approaches [3].

- The averaging approach is an extension of the classical approach developed by Quinlan and based on the gain ratio criterion [12].
- The conjunctive approach represented ideas behind the TBM itself and based on a distance criterion.

In our work, we will focus only on BDT in averaging approach and we will use the following notations:

- T: a given training set composed by p objects I_j j=1,...p,
- S: a set of objects belonging to the training set T,
- A: an attribute,
- Θ = {C_1,C_2,.....,C_n}: the frame of discernment made of the n possible classes related to the classification problem.
- $m^{\Theta}\{I_j\}(C)$: the bbm given to the hypothesis that the actual class of object I_j belongs to C $\subseteq \Theta$.

A. The attribute selection measure in averaging approach

The major parameter ensuring the building of a decision tree is the attribute selection measure allowing to determine the attribute to assign to a node of the induced BDT at each step. Under this approach, the attribute selection measure is based on the entropy computed from the average pignistic probabilities computed from the pignistic probabilities of each instance in the node. The following steps are proposed to choose the appropriate attribute:

1) Compute the pignistic probability of each object I_j by applying the pignistic transformation to $m^{\Theta}\{I_j\}$.
2) Compute the average pignistic probability function $BetP^{\Theta}\{S\}$ taken over the set of objects S. For each $C_i \in \Theta$,

$$BetP^{\Theta}\{S\}(C_i) = \frac{1}{|S|} \sum_{I_j \in S} BetP^{\Theta}\{I_j\}(C_i) \quad (7)$$

3) Compute the entropy Info(S) of the average pignistic probabilities in the set S. This Info(S) value is equal to:

$$Info(S) = -\sum_{i=1}^{n} BetP^{\Theta}\{S\}(C_i) \log_2 BetP^{\Theta}\{S\}(C_i) \quad (8)$$

4) Select an attribute A. Collect the subset S_v^A made with cases of S having v as a value for the attribute A. Then, compute the average pignistic probability for objects in subset S_v^A. Let the result be denoted $BetP^{\Theta}\{S_v^A\}$.
5) Compute $Info_A(S)$, as Quinlan:

$$Info_A(S) = \sum_{v \in D(A)} \frac{|S_v^A|}{|S|} Info(S_v^A) \quad (9)$$

Open Science Index, Computer and Information Engineering Vol:2, No:9, 2008 publications.waset.org/3379/pdf

where $D(A)$ is the domain of the possible values of the attribute A and $Info(S_v^A)$ is computed using $BetP^{\Theta}\{S_v^A\}$.

6) Compute the information gain provided by the attribute A in the set of objects S such that:

$$Gain(S,\ A) = Info(S) - Info_A(S) \qquad (10)$$

7) Using the split Info, compute the gain ratio relative to attribute A:

$$Gain\ Ratio(S,\ A) = \frac{Gain(S, A)}{Split\ Info(S, A)} \qquad (11)$$

Where

$$Split\ Info(S, A) = -\sum_{v \in D(A)} \frac{|S_v^A|}{|S|} \log_2 \frac{|S_v^A|}{|S|} \qquad (12)$$

8) Repeat the same process for every attribute A belonging to the set of attributes that can be selected. Next, choose the one that maximizes the gain ratio.

B. Belief decision trees procedures

Like the standard decision tree, the belief decision tree is composed of two principal procedures: the building or the construction of the tree from uncertain data and the classification of new instances that maybe characterized by uncertain or even missing attribute values.

1) Building procedure: Building a decision tree in this context of uncertainty will follow the same steps presented in C4.5 algorithm based the attribute selection measure represented in section 3.1.

2) Classification procedure: Once the belief decision tree is constructed, it is able to classify an object described by an exact value for each one of its attributes [3], we have to start from the root of the belief decision tree, and repeat to test the attribute at each node by taking into account the attribut value until reaching a leaf. As a leaf is characterized by a bba on classes, the pignistic transformation is applied to get the pignistic probability on the classes of the object to classify in order to decide its class. For instance, one can choose the class having the highest pignistic probability.

Belief decision trees also deal with the classification of new instances characterized by uncertainty in the values of their attributes. The idea to classify such objects is to look for the leaves that the given instance may belong to by tracing out possible paths induced by the different attribute values of the object to classify. The new instance may belong to many leaves where each one is characterized by a basic belief assignment. These bba's are combined using disjunctive rule of combination in order to get beliefs on the instance's classes.

IV. Pruning belief decision trees method

A belief decision tree is a classification technique based on decision trees within the framework of belief function theory. Inducing a belief decision tree may lead in most cases to very large trees with bad classification accuracy and difficult comprehension. Several pruning methods have been developed to cope with this problem including minimal cost-complexity pruning [1], reduced error pruning [10], critical value pruning [8], pessimistic error pruning [10], minimum error pruning [9] and error based pruning [12].

All these methods deal with only standard decision trees and not with BDT. So, our objective is to adapt one of these post-pruning methods in order to simplify the belief decision tree and improve its classification accuracy. In our work, we will choose minimal cost-complexity pruning (MCCP) to adapt for pruning belief decision trees. This pruning method is appealing because it performs well in terms of size pruned tree and accuracy. It also produces a selection of trees for the expert to study. Where it is helpful if several trees, pruned to different degree are available.

This section is dedicated to the presentation of our pruning belief decision tree method based on MCCP. We start by explaining how this method works in a certain case, then we present our pruning method in an uncertain case.

A. Minimal cost-complexity pruning in certain case

The MCCP, was developed by Breiman et al [1]. This method is also known as the CART pruning algorithm. It consists of two steps :

1) Generating a series of increasingly pruned trees $\{T_0, T_1, T_2, \ldots, T_n\}$.
2) Selecting the best tree with the lowest error rate on separate test set.

As regarding the first step, T_{i+1} is obtained from T_i by pruning all the nodes having the lowest increase in error rate per pruned leaf denoted α.

$$\alpha = \frac{R(t) - R(Tt)}{NT - 1} \qquad (13)$$

- R(t) is the error rate if the node is pruned, which becomes a leaf belonging to only one class. It is the proportion of training examples which do not belong to this class.
- R(Tt) is the error rate if the node is not pruned. It represents the average of the error rates at the leaves weighted by the number of examples at each leaf.

The method works as follows:

1) Compute α for each (non-terminal) node (except the root) in T_i.
2) Prune all the nodes with the smallest value of α, so obtaining the tree T_{i+1}.
3) Repeat this process until only root is left yields a series of pruned tree.
4) The next step is to select one of these as the final tree. The criterion for selection of the final tree is the lowest mis-classification rate on independent data set. This selection is based only on testing set accuracy.

Open Science Index, Computer and Information Engineering Vol:2, No:9, 2008 publications.waset.org/3379/pdf

B. BDT pruning method based on MCCP

Our objective is to develop a pruning method based on standard minimal cost-complexity pruning to prune belief decision trees in averaging approach. To prune a node in MCCP, we compute the error rate if the node is pruned or not. To do this, we should know at each node or leaf, the number of objects belonging to each class. However, in belief decision tree, the class of the objects are represented by a bba and not by a certain class.

The idea is to use the pignistic transformation. It is a function which can transform the belief function to probability function in order to make decisions from beliefs. This function is used to build the belief decision tree. In this section, we propose the following steps to prune the belief decision tree by adapting MCCP.

1) For each node in the belief decision tree, compute the pignistic probability of each object I_j by applying the pignistic transformation to $m^{\Theta}\{I_j\}$.

2) Compute the sum pignistic probability function $BetP^{\Theta}\{\mathrm{S}\}$ taken over the set of objects S belonging to a node . For each $C_i \in \Theta$,

$$BetP^{\Theta}\{S\}(C_i) = \sum_{I_j \in S} BetP^{\Theta}\{I_j\}(C_i) \qquad (14)$$

$BetP^{\Theta}\{\mathrm{S}\}$ represents the number of objects that belonging to each class $C_i \in \Theta$ for a node.
In this way, we can compute the number of errors of each node. It is the sum of objects not allocated to the class which occurs most frequently.

3) Compute the error rate if the node is pruned, which become a leaf.

R(t)=

$$\frac{\sum_{C_i \in \Theta} (BetP^{\Theta}\{S\}(C_i)) - Max\,(BetP^{\Theta}\{S\}(C_i))}{|T|} \qquad (15)$$

$Max\,(BetP^{\Theta}\{S\}(C_i))$ represents the number of training objects belong to the class which occurs most frequently.

4) Compute the error rate if the node is not pruned.

$$R(T_t) = \sum R(i) \text{ for i = sub-tree leaves} \qquad (16)$$

5) Compute the increase in error per pruned leaf, denoted α.

$$\alpha = \frac{R(t) - R(Tt)}{NT - 1} \qquad (17)$$

Where NT is the number of leaves in the node and $NT - 1$ is the number of pruned leaves

6) Repeat the same process for every node in the belief decision tree only the root.

If a node has the lowest α starting pruned it and obtaining the first pruned tree. The node become a leaf represented by the average bba relative to the objects belong to it. Continue this process until the root is left yields a series of pruned tree.

Example: To explain how to compute the error rate of a node in an uncertain context, we take a node N containing three objects and induced from training set of 10 instances. The class of each object is represented by a bba $m^{\Theta}\{I_j\}$ are defined as follows:

$m^{\Theta}\{I_1\}(C_1)$=0.7; $m^{\Theta}\{I_1\}(\Theta)$=0.3;
$m^{\Theta}\{I_2\}(C_1)$=0.6; $m^{\Theta}\{I_2\}(\Theta)$=0.4;
$m^{\Theta}\{I_3\}(C_1)$=0.95; $m^{\Theta}\{I_3\}(\Theta)$=0.05;

Compute the pignistic probability of each object I_j (see Table 1)

TABLE I
COMPUTATION OF $BetP^{\Theta}\{S\}$

	C_1	C_2	C_3
$BetP^{\Theta}\{I_1\}$	0.8	0.1	0.1
$BetP^{\Theta}\{I_2\}$	0.74	0.13	0.13
$BetP^{\Theta}\{I_3\}$	0.96	0.02	0.02
Sum = $BetP^{\Theta}\{S\}$	2.5	0.25	0.25

The node N has 2.5 objects of class C_1, and 0.25 of class C_2 and 0.25 of C_3.

Error rate if the node is pruned:(see equation 15)

$$R(t) = \frac{3 - 2.5}{10} = 0.05;$$

V. EXPERIMENTATION AND SIMULATION

In our experiments, We have performed several tests and simulations on real databases obtained from the U.C.I. repository: Wisconsin breast cancer database, Balance Scale weight and Congressional voting databases available in [1]. These datasets are modified in order to include uncertainty in classes.

Different results carried out from these simulations will be presented and analyzed in order to evaluate our proposed pruning method for certain and uncertain cases. The size characterized by the number of nodes and leaves in the belief decision tree and the PCC representing the percent of correct classification of the objects belonging to testing set are relevant criteria to judge the performance of a pruning method.

Let us remind that our objective is to reduce the size and improve the classification accuracy of belief decision tree by pruning it. So, we compare the size and PCC resulting from the application of the averaging approach without pruning (S.bef.Prun, PCC.bef.Prun),with pre-pruning [5] (S.aft.Pre.Prun, PCC.aft.Pre.Prun) and with applying our pruning method (S.aft.Prun, PCC.aft.Prun).

[1]http://www.ics.uci.edu/ mlearn/MLRepository.html

Open Science Index, Computer and Information Engineering Vol:2, No:9, 2008 publications.waset.org/3379/pdf

A. Results of the certain case

Tables II and III summarize different results relative to W. Breast Cancer, B. Scale weight and C. Voting databases for the certain case.

TABLE II
EXPERIMENTAL MEASURES (CERTAIN CASE (SIZE))

Degree of uncertainty	Size bef.Prun	Size aft.Pr.Prun	Size aft.Pt.Prun
W. Breast Cancer	274	151	123
B. Scale weight	326	265	109
C. Voting	57	34	29

TABLE III
EXPERIMENTAL MEASURES (CERTAIN CASE (PCC))

Degree of uncertainty	PCC bef.Prun	PCC aft.Pr.Prun	PCC aft.Pt.Prun
W. Breast Cancer	76.6%	77.2%	82.53%
B. Scale weight	60.7%	62.3%	70.9%
C. Voting	95.21%	95.74%	96.53%

From these tables, we can conclude that our pruning method in certain case has good results. There are an improvement of the size in all databases. For W. Breast Cancer, the mean size goes from 274 items to 123 items. For B. Scale weight database, the size of the induced tree goes from 326 items to 109 items. Finally, for C. voting, the size is reduced from 57 to 29 items. There also an increase of the PCC for all databases. For W. Breast Cancer, the PCC goes from 76.6% to 82.53%. For B. Scale weight database, the mean PCC is increased from 60.7% to 70.9%. Finally, for C. Voting, the PCC goes from 95.21% to 96.53%. We can also conclude that pre-pruning reduces size and increases PCC, but not better than post-pruning.

B. Results of the uncertain case

This section presents different results carried out from testing the pruning belief decision tree method in averaging approach on uncertain case. The class of the instances in our databases is certain, so we will modify them and create artificially the uncertainty in class of each object. We take different degrees of uncertainty (low, middle and high) based on increasingly values of probabbility P used to transform the actual class C of each object to a bba: m(C)= 1-P and m(Θ)= P.

TABLE IV
EXPERIMENTAL MEASURES (W. BREAST CANCER, UNCERTAIN CASE)

Degree of.uncertainty	Size bef.Prun	Size aft.Pre.Prun	Size aft.Prun
low degree	399	302	82
middle degree	401	305	87
high degree	444	321	101
Mean	414	309	90

TABLE V
EXPERIMENTAL MEASURES (W. BREAST CANCER, UNCERTAIN CASE)

Degree of.uncertainty	PCC bef.Prun	PCC aft.Pre.Prun	PCC aft.Prun
low degree	67.97%	69.13%	82.38%
middle degree	67.83%	68.38%	82.54%
high degree	66.09%	67.18%	81.29%
Mean	67.29 %	68.23%	82.07%

From Table IV and V, we can conclude that pruning belief decision tree method work well in all degrees of uncertainty for W. Breast Cancer database. The mean size goes from 414 items to 90 items and the mean PCC is improved from 67.29% to 82.07%.

TABLE VI
EXPERIMENTAL MEASURES (BALANCE SCALE WEIGHT, UNCERTAIN CASE)

Degree of.uncertainty	Size bef.Prun	Size aft.Pre.Prun	Size aft.Prun
low degree	508	338	139
middle degree	498	331	147
high degree	522	373	97
Mean	509	347	127

TABLE VII
EXPERIMENTAL MEASURES (BALANCE SCALE WEIGHT, UNCERTAIN CASE)

Degree of.uncertainty	PCC bef.Prun	PCC aft.Pre.Prun	PCC aft.Prun
low degree	58.68%	66.84%	77.81%
middle degree	58.35%	66.58%	77.78%
high degree	63.1%	66.26%	81.7%
Mean	60 %	66.56%	79%

Tables VI and Table VII show that our post pruning method has a good impact on belief decision tree for B. Scale weight database, the size is reduced from 509 items to 127 items and the PCC goes from 60% to 79%.

TABLE VIII
EXPERIMENTAL MEASURES (CONGRESSIONAL VOTING, UNCERTAIN CASE)

Degree of.uncertainty	Size bef.Prun	Size aft.Pre.Prun	Size aft.Prun
low degree	171	133	75
middle degree	158	124	56
high degree	157	111	58
Mean	162	122	63

TABLE IX
EXPERIMENTAL MEASURES (CONGRESSIONAL VOTING, UNCERTAIN CASE)

Degree of.uncertainty	PCC bef.Prun	PCC aft.Pre.Prun	PCC aft.Prun
low degree	94.29%	94.88%	96.78%
middle degree	94.08%	94.29%	96%
high degree	92.27%	92.65%	95.33%
Mean	93.67%	93.94%	96%

For C. Voting database, the Tables VIII and IX show that our pruning belief decision tree method has good results on BDT. The mean size is improved from 162 items to 63 items. The mean PCC is increased from 93.67% to 96%. Like in certain case, our post-pruning method has good results on uncertain case more than pre-pruning.

VI. CONCLUSION

In this paper, we have presented our pruning belief decision tree method with the objective to reduce the size of the induced tree and to improve the classification accuracy in an uncertain context. Pruning is a way to cope with the problem of overfitting. Then, we have shown the different results obtained from simulations and that have been performed on real databases. These experiments have shown interesting results for the performance of our post-pruning method comparing with BDT without pruning and with pre-pruning.

As a future work, we propose the pruning of the belief decision tree induced from the conjunctive approach [4] and comparaison between the both pruning methods in averaging and conjunctive approaches.

REFERENCES

[1] Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J.: Classification and Regression Trees. Wadsworth, Belmont, CA(1984).

[2] Denoeux, T., Skarstien-Bajanger, M.: Induction of decision trees for partially classified data. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. Nashiville, USA (2000) 2923-2928.

[3] Elouedi, Z., Mellouli, K., Smets, P.: Decision trees using the belief function theory. In Proceedings of the international conference on Information Processing and Management of Uncertainty IPMU'2000, Vol.1, Madrid, Spain(2000)141-148

[4] Elouedi, Z., Mellouli, K., Smets, P.: Belief decision trees: Theoretical foundations. In International Journal of Approximate Journal IJAR, Vol.28(2-3), (2001) 91-124.

[5] Elouedi, Z., Mellouli, K., Smets, P.: A Pre-Pruning Method in Belief Decision Trees. The Ninth International conference on Information Processing and Management of uncertainty in knowledge-Based Systems IPMU 2002, Vol.1, Annecy, France(2002) 579-586.

[6] Hüllermeier, E. : Possibilistic Induction in decision tree learning. In Proceedings of the 13 Th European Conference on Machine Learning. ECML2002. Helsinki (2002), Finland (2002) 173-184.

[7] Jenhani, I. Elouedi, Z. Ben Amor, N. & Mellouli, K.:Qualitative inference in possibilistic option decision trees. In Proceedings of the Eight European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty.(2005)

[8] Mingers, J.: Expert systems - rule induction with statistical data. Journal of the operational research society. Vol 38.(1987) 39-47.

[9] Niblett, T. Bratko, I.: Learning decision rules in noisy domains. Proc. Expert Systems 86, Cambridge: Cambridge University Press.(1986)

[10] Quinlan, J.R.: Simplifying decision trees. International Journal of Man-Machine studies, (1987) 221-234.

[11] Quinlan, J.R.: Probabilistic decision tree. In R.S. Michalski, J.G. Carbonell, & T.M Michell (Eds.), Machine Learning, Vol.3, Chap. 5, Morgan Kaufmann(1990) 267-301).

[12] Quinlan, J.R.: C4.5: Programs for Machine learning. Morgan Kaufman, San Mateo, California (1993).

[13] Shafer, G.: A mathematical theory of evidence. Princeton University Press. Princeton, NJ(1976).

[14] Smets, P.: Belief functions. In P. Smets, E. H. Mamdani, D. Dubois, H. Prade (Eds.): Non standard logics for automated reasoning. Academic Press, London (1988) 253-286.

[15] Smets, P., & Kennes, R.: The transferable belief model. Artificial Intelligence,Vol 66,(1994) 191-236.

[16] Smets, P.: The application of belief transferable belief model to diagnostic problems. International Journal of Intelligent Systems, Vol.13, (1998) 127-157.

[17] Zeidler, J.,& Schlosser, M.: Continuous valued attributes in fuzzy decision trees. In Procedings of the sixth International Conference on Information Processing and Management of Uncertainty , Vol.1,. Granada, Spain(1996) 395-400.

Open Science Index, Computer and Information Engineering Vol:2, No:9, 2008 publications.waset.org/3379/pdf