Asian Journal of Information Technology 7 (6): 254-259, 2008
ISSN: 1682-3915

An Algorithm of Intelligent Search Dealing with Uncertainty

[1]Ashit Kumar Dutta, [1]Sahar Idwan and [2]Ranjit Biswas
[1]Faculty of Information Technology, Prince Al-Hussein Bin Abdullah-II for IT, Hashemite University, P.O. Box 330127, Zarqa, Jordan
[2]Department of Computer Science, Institute of Technology and Management, Sector-23 A, Gurgaon, Haryana, India

Abstract: In this study, the authors propose an algorithm of intelligent search using vague theory of Gau and Buehrer. The objective of such research is to deal with the imprecise data involved in different kinds of existing searching techniques in a more efficient ways and thus to suggest a new improved version of searching technique under uncertainty which will be helpful in many real life problems of computer science, specially in AI, in Data Mining, in fuzzy DBMS, etc. to list a few only.

Key words: Vague set, fuzzy set, vague number, Most Expected Object (MEO), nlt (x)

INTRODUCTION

The study of search problem is very common in computer science. AI is one of the many fields in which search is an important topic. We search a file, we search a directory, we search for a character in a file, we search an item in a link list etc., are few of infinite number of searching done in computer. There are various methods of searching like Depth-first search (DFS), Breadth-first search (BFS), Hill Climbing, Beam search, Best first search existing in the literature. Each method has got independent merit of its own to address different types of searching in different situations. All existing searching techniques are not always based on precise data. Consequently, to deal with uncertainty in searching, we feel that non-classical logic that fuzzy logic and/or vague fuzzy logic will be the appropriate tool.

In this study, we propose an algorithm of intelligent search using vague theory of Gau and Buehrer (1993). The objective of such research is to deal with the imprecise data involved in different kind of existing searching techniques in a more efficient ways and thus to suggest few improved version of searching techniques under uncertainty which will be helpful in many real life problems of computer science specially in AI.

Preliminaries of vague set theory of Gau and Buehrer (1993): With different aims and objectives, different authors from time to time have made a number of generalizations (Dubois and Prade, 1990; Kaufmann, 1975; Zimmermann, 1991) of Zadeh's (1965) fuzzy set theory. Of these, the notion of Vague theory introduced by Gau and Buehrer (1993) is of interest to us.

In most cases of judgments, evaluation is done by human beings (or by an intelligent agent) where there certainly is a limitation of knowledge or intellectual functionaries. Naturally, every decision-maker hesitates more or less, on every evaluation activity. To judge whether a patient has cancer or not, a doctor (the decision-maker) will hesitate because of the fact that a fraction of evaluation he thinks in favour of truthness, another fraction in favour of falseness and rest part remains undecided to him. This is the breaking philosophy in the notion of vague set theory introduced by Gau and Buehrer (1993). In this study, we introduce a notion of vague algebra by defining vague groups of a group, vague normal groups and study some properties of them.

We present now some preliminaries on the theory of vague sets (VS). In his pioneer research, Zadeh (1965) proposed the theory of fuzzy sets. Since then it has been applied in wide varities of fields like Computer Science, Management Science, Medical Sciences, Engineering problems etc. to list a few only.

Let $U = \{u_1, u_2.........., u_n\}$ be the universe of discourse. The membership function for fuzzy sets can take any value from the closed interval (0, 1). Fuzzy set A is defined as the set of ordered pairs $A = \{ (u, \mu_A (u)): u \in U \}$, where $\mu_A (u)$ is the grade of membership of element u in set A. The greater $\mu_A (u)$, the greater is the truth of the statement that 'the element u belongs to the set A'.

Corresponding Author: Ashit Kumar Dutta, Faculty of Information Technology, Prince Al-Hussein Bin Abdullah-II for IT, Hashemite University, P.O. Box 330127, Zarqa, Jordan

But Gau and Buehrer (1993) pointed out that this single value combines the 'evidence for u' and the 'evidence against u'. It does not indicate the 'evidence for u' and the 'evidence against u' and it does not also indicate how much there is of each. Consequently, there is a genuine necessity of a different kind of fuzzy sets which could be treated as a generalization of Zadeh's fuzzy sets (1965).

Definition 1: A vague set (or in short VS) A in the universe of discourse U is characterized by 2 membership functions given by:

- A truth membership function

$$t_A: U \rightarrow (0, 1) \text{ and}$$

- A false membership function

$$f_A: U \rightarrow (0, 1)$$

Where,

t_A (u) : A lower bound of the grade of membership of u derived from the 'evidence for u'.

f_A (u) : A lower bound on the negation of u derived from the 'evidence against u' and $t_A (u) + f_A (u) = 1$.

Thus the grade of membership of u in the vague set A is bounded by a subinterval (t_A (u), 1-f_A (u)) of (0, 1). This indicates that if the actual grade of membership is μ (u), then $t_A (u) \leq \mu (u) \leq 1\text{-}f_A (u)$. The vague set A is written as $A = \{ < u, (t_A (u), f_A (u)) >: u \in U \}$, where the interval ($t_A$ (u), 1-f_A (u)) is called the vague value of u in A and is denoted by V_A (u).

For example, consider an universe U = {DOG, CAT, RAT}. A vague set A of U could be A = { <DOG, (.7,.2)>, <CAT, (.3,.5)., <RAT, (.4,.6)>}.

Definition 2: A vague set A of a set U with $t_A (u) = 0$ and $f_A (u) = 1 \forall u \in U$ is called the zero vague set of U.

A vague set A of a set U with $t_A (u) = 1$ and $f_A (u) = 0 \forall u \in U$ is called the unit vague set of U.

A vague set A of a set U with $t_A (u) = \alpha$ and $f_A (u) = 1\text{-}\alpha \ \forall u \in U$ is called the α-vague set of U, where $\alpha \in (0, 1)$.

Vague sets have an extra dimension over fuzzy sets: There are a number of generalizations of fuzzy sets of Zadeh done by different authors. For each generalization, one (or more) extra dimension is annexed with a more specialized type of aim and objective. Thus, a number of higher order fuzzy sets are now in literatures and are being applied into the corresponding more specialized application domains. While fuzzy sets are applicable to each of such application domains, higher order fuzzy sets can not, because of its specialization character by birth. Application of higher order fuzzy sets makes the solution-procedure more complex, but if the complexity on computation-time, computation-volume or memory-space are not the matter of concern then a better results could be achieved. Vague sets defined recently by Gau and Buehrer (1993) have also an extra edge over fuzzy sets. Let U be a universe, the set of all students of Calcutta School. Let A be a vague set of all good-in-maths students of the universe U and B be a fuzzy set of all good-in-maths students of U. Suppose that an intellectual Manager M_1 proposes the membership value μ_B (x) for the element x in the fuzzy set B by his best intellectual capability. On the contrary, another intellectual Manager M_2 proposes independently two membership values t_A (x) and f_A (x) for the same element in the vague set A by his best intellectual capability. The amount t_A (x) is the true-membership value of x and f_A (x) is the false-membership value of x in the vague set A. Both M_1 and M_2 being human agents have their limitation of perception, judgment, processing-capability with real life complex situations. In the case of fuzzy set B, the manager M_1 proposes the membership value μ_B (x) and proceed to his next computation. There is no higher order check for this membership value in general. In the later case, the manager M_2 proposes independently the membership values t_A (x) and f_A (x) and makes a check at this base-point itself by exploiting the constraint $t_A (x) + f_A (x) = 1$. If it is not honored, the manager has a scope to rethink, to reshuffle his judgment procedure either on 'evidence against' or on 'evidence for' or on both. The two membership values are proposed independently, but they are mathematically not independent. This is the breaking philosophy of Gau and Buehrer's (1993) vague sets.

SOME POPULAR EXISTING METHODS OF INTELLIGENT SEARCH

Some of the most interesting problems in AI have the frustrating property that there is no good way to solve them. It often happens that, although the solution can be generated a piece, a partial solution falls apart and some or all of the research to get it must be redone.

A search problem is characterized by an initial state and goal-state description. The guesses are called operators: a single operator transforms a state into another state, which we hope is closer to a goal state (a state satisfying the goal-state description). The objective may be to find a goal state, or to find a sequence of operators to a goal state. In addition the problem may require finding just any solution, or an optimal solution: the best solution, measured some way. Sometimes we are not sure whether there is any solution; the object is then to search until a solution is found or we are satisfied that no solution exists.

A commonly cited class of search problems is puzzles. For example, in the missionaries and cannibals puzzle, the problem is this:

Three missionaries and three cannibals are trying to cross a river. As their only means of navigation, they have a small boat, which can hold one or two people. If the cannibals outnumber the missionaries on either side of the river, the missionaries will be eaten; this is to be avoided. Find a way to get them all across.

The initial state has all the travelers on one side of the river and the goal state description has them all on the other. In this case, the goal state is described by being specified completely and the objective is to find the sequence of moves that gets to the end. By contrast, in the scheduling problem the objective is a good schedule and the sequence of moves that gets to it is of little interest.

This distinction, between searching for a state and searching for a path, is important to remember in writing programs, but is of little theoretical significance, so we usually neglect it. This is because the path can always be included in the state. For example, in the missionaries and cannibals problem, a program which searched for the goal state could be made to remember the path by changing the definition of the state to include not only who was where, but also the steps taken so far. The latter would be modified each time a step was taken. The distinction between programs that find any solution and those that find an optimal solution is more interesting.

There are a numbers of searching methods in computer science, decision science and many other areas. Let us consider a simple example: Suppose that Mr.X wants to find some path from one city to another city using a highway map. The starting point in the city, may be called ‘start node’ and the ending point in the city may be called ‘goal node’. To find the appropriate path through the highway map, Mr.X may consider two types of costs:

- Travel cost (in terms of time).
- Travel cost (in terms of distance).

If Mr.X needs to go to the goal city often, he wants to take care of the above two costs. But the data in connection to these costs is not always available and if available, it is not always precise. Therefore, the optimal searching in terms of above costs is not straightforward. The existing method of searching like DFS, BFS do not take care of costs (Fig. 1).

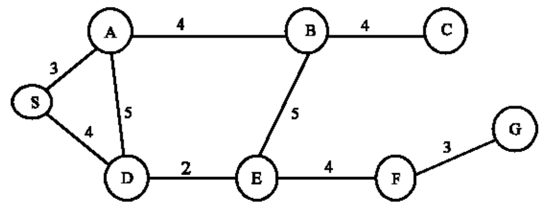

Fig. 1: Stuck in a loop-such as S-A-D-S-A-D-S-A-D

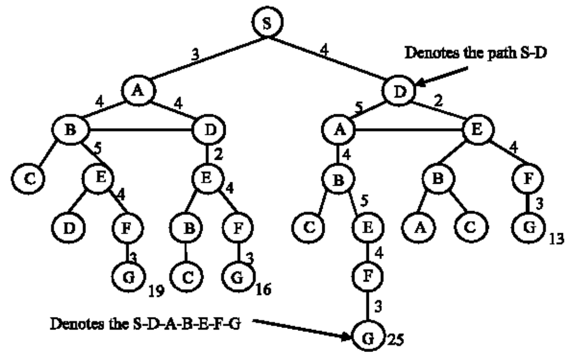

Fig. 2: The start node in a search tree

The most obvious way to find a solution is to look at all possible paths. Of course, we should discard paths that revisit any particular city so that we cannot get stuck in a loop-such as S-A-D-S-A-D-S-A-D....

With looping paths eliminated, we can arrange all possible paths from the start node in a search tree, a special kind of semantic tree in which each node denotes a path.

The Fig. 2 shows a search that consists of nodes denoting the possible paths that lead outward from the start node of the net.

Note that although in a tree denotes a path, there is no room in the diagram to write out each path at each node. Accordingly, each node is labeled with only the terminal node of the path it denotes. Each child denotes a path that is a one-city extension of the path denoted by its parent.

The most important searching methods are DFS, BFS, Hill Climbing, Beam search etc. Depth-first search is a good idea when we are confident that partial paths either reach dead ends or become complete paths after a reasonable number of steps. In contrast, depth-first search is a bad idea if there are long paths, even infinitely long paths, that neither reach dead ends nor become complete paths. In those situation, we need alternative search methods.

Breadth-first search works even in trees that are infinitely deep or effectively infinitely deep. On the other hand, breadth-first search is wasteful when all paths lead to the goal node at more or less the same depth.

Note that breadth-first search is a bad idea if the branching factor is large or infinite, because of exponential explosion. Breadth-first search is a good idea when we are confident that the branching factor is small. We may also choose breadth-first search, instead of depth-first search, if we are worried that there may be long paths, even infinitely long paths, that neither reach dead ends nor become complete paths.

We may be so uninformed about the search problem that we cannot rule out either a large branching factor or long useless paths. In such situations, we may want to seek a middle ground between depth-first search and breadth-first search. One way to seek such a middle ground is to choose nondeterministic search. When doing nondeterministic search, we expand an open node that is chosen at random. That way, we ensure that we cannot get stuck chasing either too many branches or too many levels.

Search efficiency may improve spectacularly if there is a way to order the choices so that the most promising are explored earliest. In many situations, we can make measurements to determine a reasonable ordering. In the rest of this study, we learn about such methods that take advantage of such measurements; they are called heuristically informed methods.

VAGUE SEARCH: A METHOD OF INTELLIGENT SEARCH

In our research we shall deal with such type of search problems in which distance of traversing a path is taken into utmost consideration. By the shortest path we mean the path of which total distance is minimum. An obvious procedure is that one can find out all possible paths (if possible) and then select the best one from them. This procedure is known as British museum procedure. For this one can find out all possible paths by DFS or BFS and the search is to continue till all paths are found.

But in the maximum case the site of the search tree is too large and consequently finding all possible paths is extremely rigorous. Our research here will be useful to solve the search is too large. In many cases, we can improve the existing searching technique by using guesses about distances remaining, as well as facts about distances already accumulated. After all, if guess about distance remaining is good, then that guessed distance added to the definitely known distance already traversed should be a good estimate of the total path length, u (total path length):

u (total path length) = d (already traveled)
+ e (distance remaining),

where, d (already traveled) is the known distance already traveled and e (distance remaining) is an estimate of the distance remaining.

Surely, it makes sense to work hardest on developing the path with the shortest estimated path length until the estimate is revised upward enough to make some other path by the one with the shortest estimated path length. After all, if the guesses were perfect, this approach would keep us on the optimal path at all times.

In general, however, guesses are not perfect and a bad overestimate somewhere along the true optimal path may cause us to wander away from that optimal path permanently.

Note however, that underestimates cannot cause the right path to be overlooked. An underestimate of the distance remaining yields and underestimate of total path length, u (total path length):

u (total path length) = d (already traveled)
+ u (distance remaining),

where, d (already traveled) is the known distance already traveled and where u (distance remaining) is an underestimate of the distance remaining.

Now, if we find a total path by extending the path with the smallest underestimate repeatedly, we need to do no further work once all partial path distance estimates are longer than the best complete far encountered. We can stop because the real distance along a complete path cannot less then an underestimate of that distance. If all estimates of remaining distance can be guaranteed to be underestimates, we cannot blunder.

Of course, the closer an underestimate is to the true distance, the more efficiently we search, because, if there is no difference at all, there is no chance of developing any false movement. At the other extreme, an underestimate may be so poor as to be hardly better than a guess of zero, which certainly must always be the ultimate underestimate of remaining distance. In fact, ignoring estimates of remaining distance altogether can be viewed, as the special case in which the underestimate used is uniformly zero.

Vague numbers are used by Chen (2003) to analyze vague system reliability. We now define two new objects nlt (x) and MEO which will be useful in our presentation next.

Vague number nlt (x): Let x ° R, the set of real numbers. A vague number not less than x, called in short by nlt (x), is such that the true membership value $t_{nlt\,(x)}$ (x) may or may not be equal to unity in nlt (x).

We next present the definition of the Most Expected Object (MEO) of a vague number nlt (x).

Most Expected Object (MEO): Suppose that A is a vague set of a set X with true membership function t_A and false-membership function f_A. The term 'Most Expected Object (MEO)' of A is defined by the following crisp subset of X:

$$MEO(A)=\{x_p\}, \text{if } t_A(x_p)=1$$
$$=\left\{\begin{matrix} x_q: \text{where, } x_q \in X \text{ and} \\ \dfrac{t_A(x_q)}{f_A(x_q)} = \max\left\{\dfrac{t_A(x_i)}{f_A(x_i)} : x_i \in X\right\}\end{matrix}\right\},$$
otherwise

Proposed method: We start by an example. Consider a basic search problem shown in Fig. 1.

Suppose that the number in the Fig. 1 denotes the travel cost (between the corresponding pair of nodes). We want to find the optional path from the starting node S (called the root node) to the destination node G (called the goal node). The term optimal means with minimum travel cost.

In our method we start from S and keep track of all partial paths contending for further consideration. The shortest one is extended one level creating as many new partial paths as there are branches, but we make every time a guess about the distance remaining in terms of a vague number like nlt (x).

Suppose that a guess about the distance remaining is done which is nlt (x). The guess is done by the path searcher. Then a good underestimate of the total path length is given by e = d+x', where d = distance already traveled and x' = least member of MEO (x) not less than x.

Here, x' acts as an underestimate of the remaining distance. Clearly, the underestimates at every step cannot cause the right path to be overlooked. Now, if we find a total path by extending the path with the smallest underestimate repeatedly, we need to do further research once all partial path distance estimates are longer than the best complete path distance so far encountered. The algorithm is given in the next study.

Vague search algorithm:

- Until the first path in the queue terminates at the goal node or the queue is empty.
- Remove the first path from the queue; create new paths by extending the first path to all the neighbors of the terminal node.
- From a one-element queue consisting of a zero-length path that contains only the root node.
- Reject all new paths with loops.
- Add the remaining new paths, if any, to the queue.
- Sort the entire queue by e (where e = d+x'), with least-cost paths in front. {here guess is nlt (x) and x' = least member of MEO (x) not less than x }.
- If the goal node is found, then success; otherwise failure.

An example: Consider the search problem in Fig. 3, which we now solve using the above algorithm. Starting from root node S, there are initially 2 partial paths, which are S-A and S-D. At the node A, d = 13 and suppose that guess made by the decision-maker about the remaining distance is the vague number nlt (90).

Also suppose that MEO (90) = 90' = 93. Therefore, at A, e = d+x' = 106. At the node D, d = 24 and suppose that guess about the remaining distance is the vague number nlt (40).

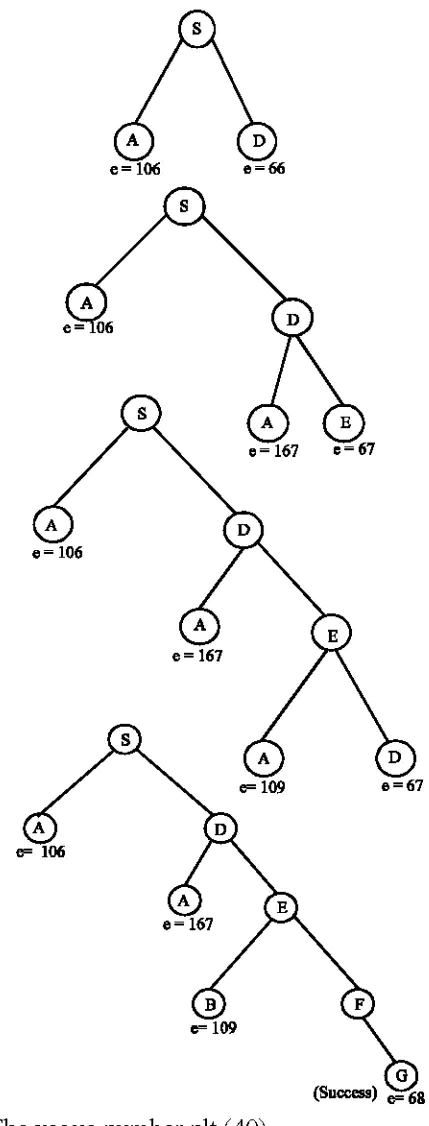

Fig. 3: The vague number nlt (40)

Suppose that MEO (40) i.e. 40' = 42. Therefore, e = d+ x' = 24 +42 = 46.

Clearly, this time D is the node from which we have to search, because D's underestimated path length is 66 which is shorter than that of A.

Expanding D leads to the partial paths S-D-A and S-D-E. The expansion of partial path at every step will look like below (hypothetical). Thus, we find that the intelligent vague optimal path is S-D-E-F-G, which is of length 68.

CONCLUSION

In this study, we have presented an intelligent search technique using vague set theory of Gau and Buehrer. The proposed method is a kind of vague branch and bound search. We have explained the method by an example with hypothetical data.

REFERENCES

Chen, Shyi-Ming, 2003. Analyzing fuzzy system reliability using vague set theory. Int. J. Applied Sci. Eng., 1: 82-88.

Dubois and Prade, 1990. Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York.

Gau, W.L. and D.J. Buehrer, 1993. Vague sets. IEEE. Trans. Syst. Man and Cybernetics, 23: 610-614.

Kaufmann, A., 1975. Introduction to the theory of fuzzy subsets. Academic Press, New York.

Zadeh, L.A., 1965. Fuzzy sets, Inform. Cont., 8: 338-353.

Zimmermann, H.J., 1991. Fuzzy Set Theory. Its Applications, Kluwer Academic Publishers, Boston, USA.