Qualitative Aggregation of Information in Fuzzy Attribute Model

V.I. Levchenko and A.A. Savinov
Institute of Mathematics
Moldavian Academy of Sciences
Academiei 5, 277028 Kishinev
Moldavia (CIS)
Phone: (373+2) 73-81-46, 73-81-30
E-mail: 23LSII@math.moldova.su

April 29, 2007

Abstract

In the paper a qualitative approach to the problem of aggregation of information is discussed. The method is based on the notion of fuzzy attribute model of the aggregation problem in which information is represented in the form of fuzzy constraints on the values of attributes. The aggregate is built by carrying out logical inference on the knowledge base which represents the decision maker preferences in the form of fuzzy logical dependencies between attributes.

Key words: fuzzy attribute model, aggregation by inference, expert estimations, decision making

1 Introduction

The problem of obtaining one aggregate characteristic of the object proceeding from its particular characteristics is highly important in practice of collective expertise, multycriteria decision problems and many other fields.

For example, in any collective estimation of the object by several experts the problem of obtaining some total estimation arises absolutely naturally. This estimation may be interpreted as a collective opinion of the experts group about this object. Analogical situation arises in different problems of decision making in presence of several criteria, when there are object estimations on different (often contradictory) criteria and it is necessary to obtain the only estimation reflecting the properties of the object on the whole number of criteria.

Traditional method for dealing with similar problems as a rule consists in constructing suitable formula or a procedure allowing one to fulfill aggregation of particular estimations. The estimation obtained as a result is declared an aggregate of the particular estimations. How much adequately the aggregate reflects, e.g., the expert group opinion or the object estimation on the whole complex of criteria depends on the degree of suitability of the formula or procedure chosen. Therefore, it is just the substantiation of the selection of the aggregation procedure that often is a weak place in such approach.

The method considered in the paper allows formulating the interdependencies connecting particular estimations with aggregate explicitly in the language of some fuzzy attribute model of the aggregation problem. The aggregate estimation results from the logical inference on the knowledge about dependencies between the aggregate and particular estimations. Such an approach is especially suitable for decision maker allowing him to formulate in rather simple and clear form the rules reflecting his own view of the dependency the aggregate estimation on the source particular estimations.

In particular, using fuzzy attribute model of the aggregation problem allows obtaining a qualitative idea of the aggregate but sufficient for practical decision making in one situation or another.

2 The Aggregation Problem

Let the particular estimation $A_1, A_2, \dots, A_n$ be given concerning the object from some class $\mathcal{C}$. For any element (object) $m \in \mathcal{C}$ the estimation value $A_i(m)$ is thought of as a result of "measurement" of the object m in some scale S_i so that

$$A_i : \mathcal{C} \to S_i \quad (i = 1, 2, \dots, n)$$

It is supposed that there exists some aggregate estimation A which may be obtained proceeding from the particular estimations $A_1, A_2, \ldots, A_n$. It is considered that, in principle, the dependency A on $A_1, A_2, \ldots, A_n$ is known. At the informal level the later is the decision maker view of the connections between A and $A_1, A_2, \ldots, A_n$; the decision concerning the object m depends on the value $A(m)$.

The problem consists in constructing a method by which one is able to infer the aggregate estimation $A(m)$ concerning any object $m \in \mathcal{C}$ provided that the particular estimations $A_i(m)$ $(i = 1, 2, \ldots, n)$ are presented. Obviously this problem formulation requires some explanations and specification. Therefore let us give several concrete aggregation problems in above sense.

2.1 Finding Aggregate Opinion of the Expert Group

Let the group of experts $A = \{A_1, A_2, \ldots, A_n\}$ has the goal to estimate some object m. It is required from the particular estimations $A_i(m)$ $(i = 1, 2, \ldots, n)$ to obtain the common opinion of the whole expert group A concerning the object m, i.e., to build the aggregate of the particular estimations of all experts $A(m)$.

The most simple situation arises if homogeneous in some sense group is considered when we suppose that all experts fulfill the same functions and have equal competence, i.e., their contributions to the aggregate. Then (under particular conditions on the kind of the estimations $A_i(m)$), e.g., the arithmetical mean of all particular estimations can be used as the aggregate:

$$A(m) = \frac{1}{n} \sum_{i=1}^{n} A_i(m)$$

It is also rather simple to take into account different expert "weights" expressing their degree of competence, i.e., different their views into the common opinion. In this case the weighted mean can be used as the aggregate instead of the arithmetical mean.

However there are the expertise techniques where experts fulfill different functions. For example, in method of court it is supposed that there are three groups of experts playing the roles of advocator, procurator and arbitrator. In this case building the aggregate on the particular expert estimations is not so obvious. In practice more complex cases of the decision making are possible, where the aggregate dependency on the particular estimations may be rather complicated.

2.2 Aggregation of Rankings Fulfilled on Different Criteria

The typical representative of such kind problems may serve the problem "what car to choose". Buyer considers as possible decisions, say, five cars: 1, 2, 3, 4, 5. There are particular rankings of these cars indicating preference:

(A_1)	by cost	— 2, 3, 5, 1, 4
(A_2)	by economy	— 1, 5, 4, 2, 3
(A_3)	by design	— 2, 5, 1, 3, 4
(A_4)	by comfort	— 3, 5, 2, 4, 1

In the given example either the aggregate ranking expressing the purchase preference on a number of criteria A_1, A_2, A_3, A_4, or the aggregate in the form of fuzzy estimation of each of the cars 1, 2, 3, 4, 5 on the universe {Buy, No Buy} may be offered to the buyer (decision maker) as the aggregate.

In this situation it is absolutely obvious that the method of aggregation of the particular estimations depends on the types and possibilities of the buyer and therefore it is difficult to propose universal procedure for calculating the aggregate estimation.

Similar difficulties may occur in solving the problem of aggregation of generalized rankings, for example, where each place in rankings is assigned specific semantics.

2.3 Aggregation of the Time Sequence of Estimations

By the time sequence of estimations we mean a number of object estimations obtained in different time moments. In such a case the aggregation must express the object estimation for a particular time interval.

As an example of the time sequence of estimations we can take appraisals of the pupil knowledge on some subject obtained, e.g., for a year. The problem consists in putting down one mark total for this period. If the teacher (decision maker) when finding the mark total besides absolute mark values also considers, e.g., the distribution of positive and negative marks in the beginning, the middle and in the end of the period, the stability of marks, the tendency to improvement or worsening and others, then using as aggregate the arithmetical mean (or similar) estimation will hardly satisfy more adequate aggregate procedure arises.

2.4 Aggregation of Heterogeneous Parameters

Suppose that concerning the fixed group of people appraisals can be given on the attribute A_1 = SEX, A_2 = HEIGHT, A_3 = WEIGHT, A_4 = PHONE NUMBER. It is required to solve the problem which may be written as follows:

$$\begin{array}{l} \text{SEX}(m) = \text{Male} \\ \text{WEIGHT}(m) = 180 \text{ cm} \\ \text{WEIGHT}(m) = 75 \text{ kg} \\ \text{PHONE}(m) = \text{12-34-56} \\ \hline \text{TOTAL}(m) = \text{?} \end{array}$$

Specifying what is meant by TOTAL obviously is possible only if the problem is put into the concrete situation. For example, a criminal is large or middle in height, about 80 kg in weight and lives apparently near the city Z. Then the aggregate estimation has to split up all people from $\mathcal{C}$ by the degree of suspectability.

In this situation building an adequate formula with the help of traditional methods for finding the aggregate seems quite problematic.

The considered examples of aggregation problems show that it is highly desirable to define one general procedure making it possible to describe constructively the aggregation method in different situations. For this reason it is necessary to have means for direct using knowledge about the decision maker aggregate dependency on the particular estimation. This knowledge may include:

- knowledge about techniques and circumstances of obtaining the particular estimations;
- knowledge about interdependencies between the aggregate and particular estimations;
- knowledge about the aggregate destination.

In particular, the latter supposes that the decision maker realizes how much the result (aggregate) must be accurate from the point of view of its destinations. This also may be taken into account when constructing the aggregation method.

3 Forming the Attribute Model of the Aggregation Problem and Aggregation by Logical Inference

General idea of solving the aggregation problem by means of logical inference on some knowledge base consists in the following. The system of attributes $\mathcal{P} = \{P_1, \ldots, P_k\}$ is built which, on the one hand, is connected with the estimations $A_1, A_2, \ldots, A_n$ concerning the objects from class $\mathcal{C}$ and, on the other hand, is arranged so that it is convenient for the decision maker to express his views of the aggregate estimation dependency on the particular estimations $A_1, A_2, \ldots, A_n$.

Suppose that each attribute P_i takes its values on the finite set $\mathcal{V}_i$. The aggregate estimation A also may be considered as an attribute taking its values on the finite set $\mathcal{V}_A$.

The decision maker forms his views of the aggregate in the terms of the attribute system $P_1, \ldots, P_k$ and A in the form of connections between the values of attributes. Thus the subset π is selected in the Cartesian product $\Omega = \mathcal{V}_1 \times \mathcal{V}_2 \times \ldots \times \mathcal{V}_k \times \mathcal{V}_A$ (in the case of fuzzy dependencies π is also fuzzy). The triple

$$\langle \mathcal{P} \cup A, \{\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k, \mathcal{V}_A\}, \pi \rangle$$

is called the attribute model of the aggregation problem.

It is supposed that for any object $m \in \mathcal{C}$ the particular estimations $A_i(m)$ ($i = 1, 2, \ldots, n$) generate some values of attributes $P_1, P_2, \ldots, P_k$. In other words, the object particular estimations generate the description of this object in the attribute system $P_1, P_2, \ldots, P_k$.

Thus calculation of the aggregate $A(m)$ comes to the finding admissible values of attribute A in the model of the problem provided that the values of attributes $P_1, P_2, \ldots, P_k$ are given. This can be fulfilled by means of logical inference described in [3,4].

Unfortunately it seems rather difficult to indicate regular procedure for building the attribute model adequate to the aggregation problem being solved: this stage requires that informal problem analysis be fulfilled and it must be carried out by the cognitolog together with the decision maker.

In order to illustrate this approach let us consider a variant of the attribute model for the aggregation problem described in section 2.

3.1 Attribute Model in the Problem "What Car to Choose"

Suppose that the decision maker (buyer) would like to obtain the aggregate in the form of fuzzy estimation on the universe {Buy, No Buy}. As attributes P_1, P_2, P_3, P_4 we can take, for example

$P_i =$ be in ranking on the criteria A_i $(i = 1, 2, 3, 4)$ with the values on the set {Beginning, Middle, End}.

Attribute A takes values on the set {Buy, No Buy}.

In this attribute model the decision maker describes his system of preferences in the form of rules, examples and so on. For example, for thrifty buyer of modest means it may look like this:

1. IF $P_1 =$ End, THEN $A =$ No Buy
2. IF $P_1 =$ Beginning & $P_2 =$ Beginning, THEN $A =$ Buy
3. IF $P_1 \neq$ End & $P_2 =$ Beginning & $P_3 =$ Beginning, THEN $A =$ Buy

For a buyer for which the car cost is of lower importance these rules may look as follows:

1. IF $P_3 =$ Beginning & $P_4 =$ Beginning, THEN $A =$ Buy
2. IF $P_1 =$ Beginning & $P_2 =$ Beginning & $P_3 =$ Beginning, THEN $A =$ Buy

3. IF P_3 = End, THEN A = No Buy

Of course, in reality the problem attribute model may include considerably more attributes and rules.

Suppose that the rankings of five cars (section 2.2) are directed to the input. In order to determine the values of attributes for any car m, specific procedures are given (possibly by expert way). Suppose that we are going to use fuzzy attribute model. Then this procedure has source rankings to assign the values of attributes P_1, P_2, P_3, P_4. To do it, e.g., fuzzy notions "be in the beginning (middle, end) of ranking" may be used which are represented by fuzzy sets with the help of membership functions (Fig. 1).

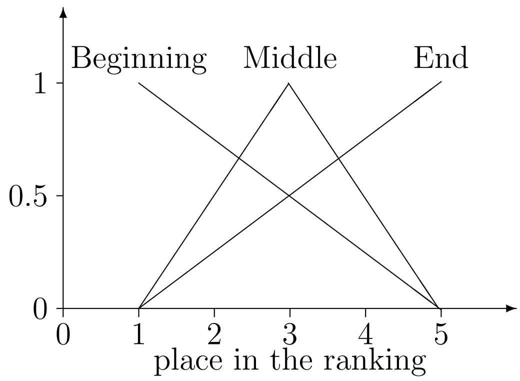

Figure 1: Fuzzy notions for rankings

Then for the car number 1 we have the following values of attributes:

$P_1(1) = \{\text{Beginning} : 0.25, \text{Middle} : 0.5, \text{End} : 0.75\}$
$P_2(1) = \{\text{Beginning} : 1, \text{Middle} : 0, \text{End} : 0\}$
$P_3(1) = \{\text{Beginning} : 0.5, \text{Middle} : 1, \text{End} : 0.5\}$
$P_4(1) = \{\text{Beginning} : 0, \text{Middle} : 0, \text{End} : 1\}$

Fuzzy values of attributes obtained are directed to the input of inference engine (the authors used the system shell EDIP-FUZZY [1,2]) and by means of carrying out logical inference on the knowledge base representing the decision maker preference system the value of aggregate A is obtained in the form of fuzzy set

$$A(1) = \{\text{Buy} : \alpha_1, \text{No Buy} : \alpha_2\},$$

where α_1 and α_2 are some numbers from [0,1].

3.2 Attribute Model of the Aggregation Problem "Search for Criminal"

For the problem "search for criminal" (section 2) the attribute model may look as follows:

P_1 : HEIGHT = {High. Middle, Low}
P_2 : WEIGHT = {Large, Middle, Small}
P_3 : SEX = {Male, Female}
P_4 : PHONE = {Begins with 55 or 56, Others}
A : SUSPECTABILITY = {High, Middle, Low}

To transform the source estimations HEIGHT and WEIGHT the corresponding fuzzy notions "High", "Large" etc. may be used, each of them being characterized by certain membership function (Fig. 2).

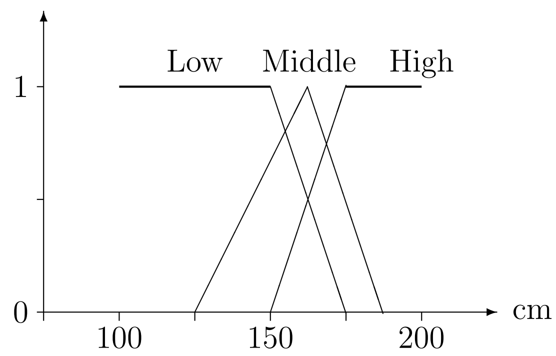

Figure 2: Fuzzy notion HEIGHT

The dependency (in correspondence with the known circumstances) between the aggregate A and attributes P_1, P_2, P_3, P_4 is given by an investigator (decision maker) with the help of the rules, for example

1. IF SEX = Female, THEN A = Low

2. IF SEX = Male & HEIGHT $\neq$ Low & WEIGHT = Large & PHONE = Begins with 55 or 56, THEN A = High

3. IF SEX = Male & HEIGHT $\neq$ Low & WEIGHT = Small, THEN A = Low

Concrete man m with the help of fuzzificators described above obtains fuzzy values of attributes P_1, P_2, P_3, P_4 which are directed to the input of logical inference on the corresponding knowledge base. As a result we obtain the aggregate value in the form

$$A(m) = \{\text{High} : \alpha_1, \text{Middle} : \alpha_2, \text{Low} : \alpha_3\}$$

4 Conclusions

a) As it is evident from the examples, the approach proposed in the paper can be effectively applied if the attribute model of the aggregation problem is sufficiently simple (by a number of attributes, their values and rules). It means that such approach may be applied to the aggregation problems which permit qualitative formulation.

b) The knowledge connecting the aggregate and the model attributes must be sufficiently complete, i.e., it is required that the decision maker have clear problem and its aims formulation and, what is more, the ability to formulate explicitly his own system of preferences.

c) Transition from the particular estimations to the values of model attributes may require the presence of fuzzificators. Their building is the separate problem consisting in construction of membership functions which formalize corresponding fuzzy notions.

References

[1] V.I. Levchenko and A.A. Savinov, The representation of fuzzy knowledge in the diagnostic expert system shell EDIP, 2nd Int. Conf. on Fuzzy Logic and Neural Networks—IIZUKA'92, July 17–22, Iizuka, Japan, (1992).

[2] V. Levchenko and A. Savinov, The matrix representation of fuzzy knowledge and its application to the expert systems design, Computer Sci. J. of Moldova **1**(1), 62–84, (1993).

[3] A.A. Savinov, Fuzzy propositional logic for the knowledge representation, 1st European Congress on Fuzzy and Intelligent Technologies—EUFIT'93, Aachen, Germany, September 7–10, (1993).

[4] A.A. Savinov, Fuzzy propositional logic, Fuzzy Sets and Systems **60**(1), 9–17, (1993).