On the Invalidity of Fuzzifying Numerical Judgments in the Analytic Hierarchy Process

Thomas L. Saaty[a*], Liem T. Tran[b]

[a]*University of Pittsburgh, Pittsburgh, PA 15260*
[b]*University of Tennessee, Knoxville, TN 37996*

Keywords: Analytic Hierarchy Process, Perturbation, fuzzy

Summary : Fuzzy set theory has serious difficulties in producing valid answers in decision-making by fuzzifying judgments. No theorems are available about its workability when it is applied indiscriminately as a number crunching approach to numerical measurements that represent judgments. When judgments are allowed to vary in choice over the values of a fundamental scale, as in the Analytic Hierarchy Process, these judgments are themselves already fuzzy. To make them fuzzier can make the validity of the outcome, when the actual outcome is known, worse, as shown by several examples in this paper. Also, improving the consistency of a judgment matrix does not necessarily improve the validity of the outcome. Validity is the goal in decision-making, not consistency, which can be successively improved by manipulating the judgments as the answer gets farther and farther from reality. An example of this is included.

1. Introduction

One reason why we use quantitative methods and algorithms to model the world is to help us understand it better and more accurately in order to control and change it to our liking. When our effort to do that without inquiring as to why, the modeling effort can become a misguided intellectual exercise to publish without concern for the validity of what we are doing. One way to see the fallacy of playing with numbers is to illustrate with examples that have known answers. Fudging the numbers with fuzziness not only increases the complexity of manipulations but also robs the original numbers of their elegance and simplicity to represent judgments and often leads to less desirable, instead of more desirable outcomes. Some authors have done it because, in their words, "it is popular".

There has been some hype in the literature about "improving" some mathematics and in particular numbers, through fuzzification. Fuzzy logic is seen as another means for dealing with uncertainty, along with the traditional probability theory and statistics. Fuzzy logic is derived from fuzzy set theory which deals with approximate rather than precise reasoning as in classical predicate logic. Different from classical set theory, fuzzy set theory permits the gradual assessment of the membership of elements in relation to a set, described with a membership function $\mu \rightarrow [0, 1]$. Fuzzy sets and fuzzy logic have found their theoretical origins and applications in electrical engineering. A list of applications given on the internet includes automobile subsystems, such as ABS and cruise control, air conditioners, the massive engine used in the *Lord of the Rings* films, which helped show huge scale armies create random, yet orderly movements, cameras, digital image processing, such as edge detection, rice cookers, dishwashers, elevators, washing machines and other home appliances, video game artificial intelligence, language filters on message boards and chat rooms for filtering out offensive text and others.

There have been attempts to apply fuzzy concepts to the social sciences, politics and to decision making [see reading list 11-39 in the references at the end of the paper, and there are many more references not cited

---

*Corresponding author
E-mail Address: saaty@katz.pitt.edu (T.L. Saaty), ltran@utk.edu (L.T. Tran)

here]. These by their very nature include judgments under uncertainty that are already fuzzy and may not benefit from further fuzzification to improve their depth for better understanding and control. Yet we know of blind efforts to carry fuzzy techniques wherever there are numbers without questioning the validity of the practice. All numbers are seen to be amenable to fuzzy logic according to certain procedures. In particular, it has been applied in the field of operations research which unlike mathematics, does not insist on proof when a familiar technique is used. Referees of many articles appear to let them through for publication without being critical about the validity of the outcome.

In a widely circulated and highly controversial paper, Charles Elkan in 1994 [7] commented that "*...there are few, if any, published reports of expert systems in real-world use that reason about uncertainty using fuzzy logic. It appears that the limitations of fuzzy logic have not been detrimental in control applications because current fuzzy controllers are far simpler than other knowledge-based systems. In future, the technical limitations of fuzzy logic can be expected to become important in practice, and work on fuzzy controllers will also encounter several problems of scale already known for other knowledge-based systems*". Reactions to Elkan's paper are many and varied, from claims that he is simply mistaken, to others who accept that he has identified important limitations of fuzzy logic that need to be addressed by system designers. In fact, fuzzy logic was not largely used at that time, and today it is used to solve very complex problems in the AI area. Probably the scalability and complexity of the *fuzzy* system will depend more on its implementation than on the theory of fuzzy logic."

Our concern here is with the validity of applying fuzzy thinking to decision making. The first author is particularly concerned about fuzzifying judgments in the Analytic Hierarchy Process because he is its originator. The second author is very knowledgeable about and has done his work in the field of fuzzy sets. Our collaboration arose out of our mutual interest in discovering the true value of using fuzziness in making decisions.

2. Validity, Truth, Falsifiability and Fuzziness

It is known in science as emphasized in the works of the philosopher Karl Popper in the 1930's that a distinction is made between simple existential statements such as: this is a crisp number and therefore it can be transformed into a fuzzy number; and universal statements such as: all numbers before they are subjected to fuzzy thinking are crisp numbers and therefore all numbers can be transformed to fuzzy numbers. To avoid pitfalls in using universal statements, it is proposed that a counter-example be given to deny the truth of such a universal statement. Our purpose is to show that whatever the claim of making numbers fuzzy may be, fuzzification does not necessarily improve the numerical value(s) of a solution in those situations when the true value is already known by other means and is being estimated by a numerical process that represents judgments of involved participants, whether well or poorly informed. Using fuzzy numbers in decision making is inadvisable until precise conditions are given for when the process works well and when it does not. Unquestioned use of fuzzy numbers is unjustified in practice. That is all we want to show, particularly in relation to numbers that are used to represent pairwise comparison judgments subject to uncertainty whose numerical representation is already "fuzzy" and not because of uncertainty because even after fuzzification the numbers used are still uncertain. There is adequate elaborate mathematical theory to justify the stability of the outcome obtained from the numerical judgments without the need to make them fuzzy first and make the theory of their stability even more complex. In addition, examples show that the outcome of fuzzification can be far from an actual value that is already known for comparisons purposes. In this regard experts in multi attribute value theory who made comparisons of outcomes of decision experiments write [3]: *"These experiments demonstrated that the MAVT and AHP techniques, when provided with the same decision outcome data, very often identify the same alternatives as 'best'. The other techniques are noticeably less consistent with MAVT, the Fuzzy algorithm being the least consistent.*"

We should note that fuzziness may work better in metric spaces than in order spaces. The topology of order used in decision making differs considerably from the idea of closeness in metric topology used in engineering and physics. The alternatives of a decision may be forced to be close by a metric approach that ends up losing the right order among them. This may be one cause of the failure of applying fuzzy concepts wholesale to decision and social science problems.

Thus an application may be valid in the sense that it follows the rules prescribed to generate certain kinds of numbers, but the outcome may not be true in the sense of its applicability to the real world. To show that the outcome is not true one resorts to examples.

In the Analytic Hierarchy Process, the inconsistency of judgments is measured by an index based on the principal eigenvalue of the positive reciprocal matrix of judgments. Procedures are known for improving the consistency by identifying the most inconsistent judgment and indicating to what value it should be changed to obtain the maximum improvement in inconsistency due to that judgment. But the judge must consider if it is possible to change that particular judgment by very much according to her/his understanding. If the change made is inadequate to reduce inconsistency, the second most inconsistent judgment is examined carefully and so on. It may be that none of the judgments can be changed enough to reduce the inconsistence to a tolerable level in order to make a decision. In that case more knowledge and understanding are needed and the decision must be postponed, particularly if the criterion with respect to which the comparisons are made is very important. But one could mechanically and automatically change the judgments without consulting the judge, or in a paper up for publication one can do anything one wants to make the inconsistency appear as good as one wishes even bringing it down to a value of zero. Students often try to do that without paying attention to whether there is any meaning in what they are doing by changing their judgments. In fuzzy set practice all the judgments are changed without a systematic way of checking to find out if the change in judgment is acceptable. Incidentally, one does not need fuzzy sets to do that, there are gradient algorithms that accomplish the same purpose.

With regard to making changes in judgments, Professor Patrick Harker, now dean of the Wharton School has written about fuzziness to the first author: "Beyond the mathematical issues there is a fundamental question of human judgment. No one could possibly think of how to change all of the parameters simultaneously; this is simply a mathematical convenience that does not relate at all to human cognition. While one might argue that people could think of changing more than one judgment at a time, changing all $n^2/2$ seems unreasonable. I really believe that the mathematics should point out inconsistencies and guide people, but that people must ultimately make the final call on whether or not the judgments make sense."

What is useful to show now is that improving consistency in the AHP even without applying fuzzy concepts does not imply improving the accuracy of the outcome. In fact the example below shows that with better consistency the outcome is arithmetically worse. We then show that with fuzzy, the same thing happens. So it appears that fuzzifying judgments by improving AHP consistency is the wrong thing to do unless a theory is carefully developed to state the conditions under which it can be done and when it should not be done. We very much doubt that such conditions are easy to state in the area of decision making.

3. Paired Comparisons and the Analytic Hierarchy Process

The AHP uses pairwise comparisons of a knowledgeable person to determine the importance of criteria in a decision. Because most criteria are intangible, it is also important to compare the alternatives with respect to each such criterion. Even when a criterion is tangible and the alternatives have measurements with respect to them, the significance of their values must often (but not necessarily always) be compared using pairwise comparison judgments. Judgments in the AHP are entered in a square matrix of the elements on the left side of that matrix are compared with the same elements listed in the same order above the matrix.

We begin by formulating the condition for a solution in the ideal case where one has measurements $w_1,...,w_n$ for the n criteri or stimuli $A_1, A_2, ... A_n$ in order to become familiar with the structure of the problem of pairwise comaprisons. :

$$Aw = \begin{array}{c} A_1 \\ \vdots \\ A_n \end{array} \begin{bmatrix} w_1/w_1 & \cdots & w_1/w_n \\ \vdots & \cdots & \vdots \\ w_n/w_1 & \cdots & w_n/w_n \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = n \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = nw$$

where A has been multiplied on the right by the column form of the vector of weights $w = (w_1, ..., w_n)$. The result of this multiplication is nw. Thus, to recover the scale from the matrix of ratios, one must solve the problem $Aw = nw$ or $(A - nI) w = 0$ where I is the identity matrix. This is a system of homogeneous linear equations. It has a nonzero solution if and only if the determinant of A-nI, a polynomial of degree n in n (it has a highest degree term of the form n^n and thus by the fundamental theorem of algebra has n roots or eigenvalues), is equal to zero, yielding an nth degree equation known as the characteristic equation of A. This equation has a solution if n is one of its roots (eigenvalues of A). But A has a very simple structure because every row is a constant multiple of the first row (or any other row). Thus all n eigenvalues of A, except one, are equal to zero. The sum of the eigenvalues of a matrix is equal to the sum of its diagonal elements (its trace). In this case the diagonal elements are each equal to one, and thus their sum is equal to n, from which it follows that n must be an eigenvalue of A and it is the largest or *principal* eigenvalue, and we have a nonzero solution. The solution is known to consist of positive entries and is unique to within a multiplicative (positive) constant and thus belongs to a ratio scale. We note that in this case our matrix A is consistent: given any row or any minimal spanning set of entries that interconnects all the elements, the rest of the matrix can be constructed from this set by forming the appropriate ratios. More simply, the matrix satisfies the relationship $a_{ij}a_{jk} = a_{ik}$ for all $i, j. k$. Otherwise the matrix is said to be inconsistent.

When the measurements are not available and judgments are used the matrix takes the positive reciprocal form:

$$A = \begin{bmatrix} 1 & a_{12} & ... & a_{1n} \\ 1/a_{12} & 1 & ... & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 1/a_{1n} & 1/a_{2n} & ... & 1 \end{bmatrix}$$

Although reciprocal with $a_{ji} = 1/a_{ij}$, this matrix need not be consistent.

In general we assume that expert judgments are made to estimate the ratios of the entries in the vector w. Fundamentaly, the second matrix is assumed to be a small perturbation of the first matrix. There are several ways to justify the argument that in order to obtain the vector of priorities from the second matrix we must again solve the eigenvalue problem $Aw = \lambda_{max} w$. One of the simplest is theorem which says that a small perturbation of a matrix yields a matrix whose principal eigenvalue is close to the principal eigenvalue of the unperturbed matrix.

The numerical judgments use the fundamental scale of absolute numbers (invariant under the identity transformation). From logarithmic stimulus-response theory that we do not go into here, we learn that a stimulus compared with itself is always assigned the value 1 so the main diagonal entries of the pairwise comparison matrix are all 1. We also learn that we must use integer values for the comparisons. The numbers 3, 5, 7, and 9 correspond to the verbal judgments "moderately more dominant", "strongly more dominant", "very strongly more dominant", and "extremely more dominant" (with 2, 4, 6, and 8 for compromise between the previous values). Reciprocal values are automatically entered in the transpose position. Consistency of the set of judgments is measured by the *consistency ratio* (C.R.), which we explain now.

The computation

$$n\lambda_{max} = \sum_{i=1}^{n}(\sum_{j=1}^{n}\varepsilon_{ij}) = \sum_{i=1}^{n}\varepsilon_{ii} + \sum_{\substack{i,j=1 \\ i \neq j}}^{n}(\varepsilon_{ij} + \varepsilon_{ji}) = n + \sum_{\substack{i,j=1 \\ i \neq j}}^{n}(\varepsilon_{ij} + \varepsilon_{ij}^{-1}) \geq n + (n^2 - n)/2 = n^2 \quad (1)$$

reveals that $\lambda_{max} \geq n$. Moreover, since $x + 1/x \geq 2$ for all x > 0, with equality if and only

if $x = 1$, we see that $\lambda_{max} = n$ if and only if all $\gamma_{ij} = 1$, which is equivalent to having all $a_{ij} = w_i / w_j$.

The foregoing arguments show that a positive reciprocal matrix A has $\lambda_{max} \geq n$, with equality if and only if A is consistent. As our measure of deviation of A from consistency, we choose the *consistency index*

$$\mu \equiv \frac{\lambda_{max} - n}{n - 1}.$$

We have seen that $\mu \geq 0$ and $\mu = 0$ if and only if A is consistent. These two desirable properties explain the term "n" in the numerator of μ; what about the term "n-1" in the denominator? Since trace $A = n$ is the sum of all the eigenvalues of A, if we denote the eigenvalues of A that are different from λ_{max} by $\lambda_2,...,\lambda_{n-1}$, we see

that $n = \lambda_{max} + \sum_{i=2}^{n} \lambda_i$, so $n - \lambda_{max} = \sum_{i=2}^{n} \lambda_i$ and $\mu \equiv \frac{\lambda_{max} - n}{n - 1} = -\frac{1}{n-1} \sum_{i=2}^{n} \lambda_i$ is the negative average of the non-principal eigenvalues of A.

In order to get some feel for what the consistency index might be telling us about a positive *n*-by-*n* reciprocal matrix A, consider the following simulation: choose the entries of A above the main diagonal at random from the 17 values {1/9, 1/8,…,1, 2,…,8, 9}. Then fill in the entries of A below the diagonal by taking reciprocals. Put ones down the main diagonal and compute the consistency index. Do this 50,000 times and take the average, which we call the *random index*. Table 1 shows the values obtained from one set of such simulations and also their first order differences, for matrices of size 1, 2,…,15.

Table 1
Random index

Order	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
R.I.	0.00	0.00	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49	1.52	1.54	1.56	1.58	1.59
First Order Differences		0.00	0.52	0.37	0.22	0.14	0.10	0.05	0.05	0.04	0.03	0.02	0.02	0.02	0.01

Since it would be pointless to try to discern any priority ranking from a set of random comparison judgments, we should probably be uncomfortable about proceeding unless the consistency index of a pairwise comparison matrix is very much smaller than the corresponding random index value in Table 1. The *consistency ratio* (C.R.) of a pairwise comparison matrix is the ratio of its consistency index : to the corresponding random index value in Table 1.

Fig. 1 below is a plot of the first two rows of Table 1. It shows the asymptotic nature of random inconsistency.
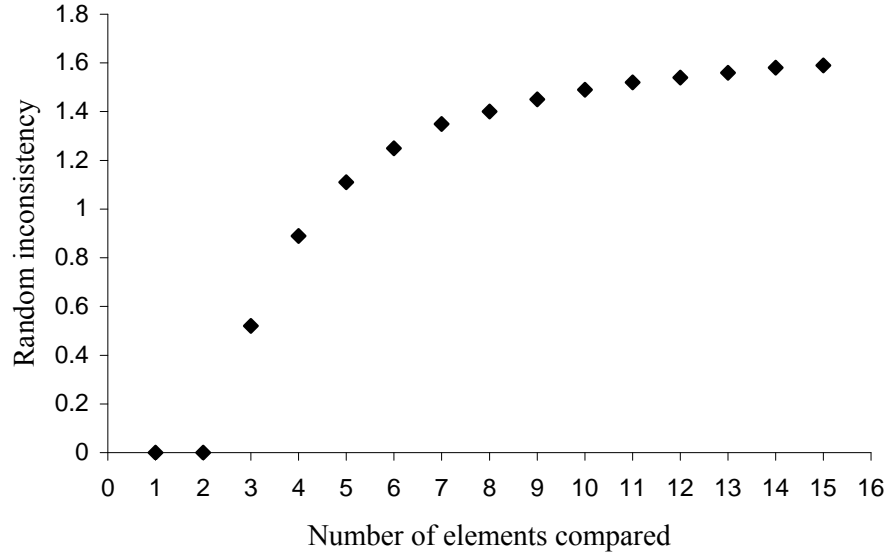
Fig. 1. Plot of Random Inconsistency

Since it would be pointless to try to discern any priority ranking from a set of random comparison judgments, we should probably be uncomfortable about proceeding unless the consistency index of a pairwise comparison matrix is very much smaller than the corresponding random index value in Table 1. The *consistency ratio* (C.R.) of a pairwise comparison matrix is the ratio of its consistency index : to the corresponding random index value in Table 1. The notion of order of magnitude is essential in any mathematical consideration of changes in measurement. When one has a numerical value say between 1 and 10 for some measurement and one wishes to determine whether change in this value is significant or not, one reasons as follows: A change of a whole integer value is critical because it changes the magnitude and identity of the original number significantly. If the change or perturbation in value is of the order of a percent or less, it would be so small (by two orders of magnitude) and would be considered negligible. However if this perturbation is a decimal (one order of magnitude smaller) we are likely to pay attention to modify the original value by this decimal without losing the significance and identity of the original number as we first understood it to be. Thus in synthesizing near consistent judgment values, changes that are too large can cause dramatic change in our understanding, and values that are too small cause no change in our understanding. We are left with only values of one order of magnitude smaller that we can deal with incrementally to change our understanding. It follows that our allowable consistency ratio should be not more than about .10. The requirement of 10% cannot be made smaller such as 1% or .1% without trivializing the impact of inconsistency. But inconsistency itself is important because without it, new knowledge that changes preference cannot be admitted. Assuming that all knowledge should be consistent contradicts experience that requires continued revision of understanding.

If the $C.R.$ is larger than desired, we do three things: 1) Find the most inconsistent judgment in the matrix (for example, that judgment for which $\varepsilon_{ij} = a_{ij}\, w_j \big/ w_i$ is largest), 2) Determine the range of values to which that judgment can be changed corresponding to which the inconsistency would be improved, 3) Ask the judge to consider, if he can, change his judgment to a plausible value in that range. If he is unwilling, we try with the second most inconsistent judgment and so on. If no judgment is changed the decision is postponed until better understanding of the stimuli is obtained. Judges who understand the theory are always willing to revise their judgments often not the full value but partially and then examine the second most inconsistent judgment and so on. It can happen that a judges knowledge does not permit one to improve his or her consistency and more information is required to improve the consistency of judgments.

6

When we speak of perturbation, we have in mind numerical change from consistent ratios obtained from priorities. The larger the inconsistency and hence also the larger the perturbations in priorities, the greater is our sensitivity to make changes in the numerical values assigned. Conversely, the smaller the inconsistency, the more difficult it is for us to know where the best changes should be made to produce not only better consistency but also better validity of the outcome. Once near consistency is attained, it becomes uncertain which coefficients should be perturbed by small amounts to transform a near consistent matrix to a consistent one. If such perturbations were forced, they could be arbitrary and thus distort the validity of the derived priority vector in representing the underlying decision.

The third row of Table 1 gives the differences between successive numbers in the second row. Figure 2 is a plot of these differences and shows the importance of the number seven as a cutoff point beyond which the differences are less than 0.10 where we are not sufficiently sensitive to make accurate changes in judgment on several elements simultaneously.
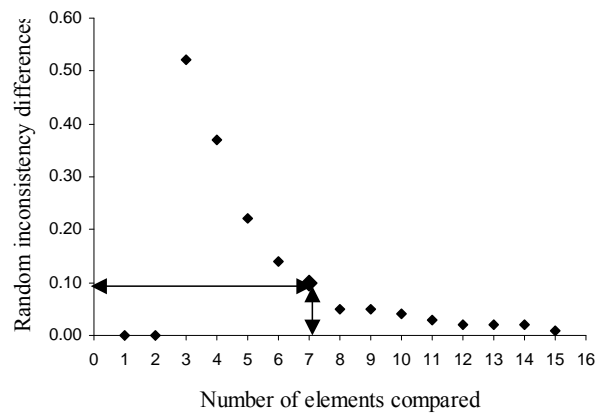

Fig. 2. Plot of First Differences in Random Inconsistency

4. Improving Inconsistency need not improve the Validity of the Outcome

Here is an example where being inconsistent is more accurate because one is within the range of the scale. Assume that the actual values are 12, 6, 1.33, 1, or in normalized form approximately: .590, .295, .066, .049. The judgment matrix of pairwise comparisons of the actual values in the order they are given is presented in Table 2. A consistent matrix ($\lambda_{max} = n, or\ a_{ij}a_{jk} = a_{ik}$) of judgments within the 1-9 Scale is shown in Table 3.

Table 2
Consistent matrix of comparisons of the ratios of the numbers 12, 6, 1.33, 1

	12	6	1.33	1	PRIORITIES
12	1	2	9	12	0.590
6	0.500	1	4.5	6	0.295
1.33	0.111	0.222	1	1.33	0.066
1	0.083	0.167	0.75	1	0.049

Table 3
Consistent matrix ($\lambda_{max} = n, or\ a_{ij}a_{jk} = a_{ik}$) of judgments within the 1-9 Scale

Number	12	6	1.33	1	PRIORITIES
12	1	2	9	9	0.581
6	0.500	1	4.5	4.5	0.290
1.33	0.111	0.222	1	1	0.065
1	0.111	0.222	1	1	0.065

This matrix is perfectly consistent (check this by forming the ratios $a_{ik} = a_{ij}/a_{jk}$), but gives only approximate values for the priorities. Table 4 presents the inconsistent matrix ($\lambda_{max} > n$) of judgments within the 1-9 scale.

Table 4
Inconsistent Matrix ($\lambda_{max} > n$) of Judgments within the 1-9 Scale.

Number	12	6	1.33	1	PRIORITIES
12	1	3	9	8	0.590
6	0.333	1	6	6.75	0.295
1.33	0.111	0.148	1	2	0.066
1	0.125	0.167	0.5	1	0.049

This matrix gives the identical priorities to within three decimals, though it is not consistent. Thus, at least in this example, consistency does not ensure validity and inconsistency can yield more valid results. This implies that improving consistency cannot be relied on to improve validity and can be misleading particularly when fuzziness is used on judgments that are already within the tolerable range of inconsistency. In fuzzy what people do is improve consistency regardless of the consequences.

Stability of the principal eigenvector also imposes a limit on channel capacity and also highlights the importance of homogeneity. To a first order approximation, perturbation Δw_1 in the principal right eigenvector w_1 due to a perturbation ΔA in the matrix A where A is consistent is given by Wilkinson [9]:

$$\Delta w_1 = \sum_{j=2}^{n} (v_j^T \, \Delta A w_1 / (\lambda_i - \lambda_j) \, v_i^T w_j) w_j$$

Here T indicates transposition. The eigenvector w_1 is insensitive to perturbation in A, if 1) the number of terms n is small, 2) if the principal eigenvalue λ_1 is separated from the other eigenvalues λ_j, here assumed to be distinct (otherwise a slightly more complicated argument given below can be made) and, 3) if none of the products $v_j^T w_j$ of left and right eigenvectors is small but if one of them is small, they are all small. However, $v_1^T w_1$, the product of the normalized left and right principal eigenvectors of a consistent matrix is equal to n that as an integer is never very small. If n is relatively small and the elements being compared are homogeneous, none of the components of w_1 is arbitrarily small and correspondingly, none of the components of v_1^T is arbitrarily small. Their product cannot be arbitrarily small, and thus w is insensitive to small perturbations of the consistent matrix A. The conclusion is that n *must be small*, and one must compare *homogeneous* elements.

When the eigenvalues have greater multiplicity than one, the corresponding left and right eigenvectors will not be unique. In that case the cosine of the angle between them which is given by $v_i^T w_i$ corresponds to a particular choice of w_i and v_i. Even when w_i and v_i correspond to a simple λ_i they are arbitrary to

within a multiplicative complex constant of unit modulus, but in that case $|v_i^T w_i|$ is fully determined. Because both vectors are normalized, we always have $|v_i^T w_i| < 1$.

5. Fuzzy AHP Guarantees nothing and can Foul Up the Outcome of a Decision

We now illustrate with examples in the AHP whose numerical answers are known which on fuzzifying do not yield better answers, so why bother. Uncertainty in the AHP is successfully remedied by using intermediate values in the 1-9 scale combined with the verbal scale and that seems to work better to obtain accurate results than using fuzziness to change the numbers for convenience and rather arbitrarily.

Fuzzy set theory has been introduced by fuzzy neophytes into AHP mainly to deal with uncertainty associated with pairwise comparison judgments. Since their only justification to use fuzziness is the claim of uncertainty they work hard to apply their technique and publish a paper by arguing that a user might have different confidence levels on different (pairwise comparison) judgments due to various reasons, incomplete information, variation of data, dynamics of the problem under study and so on. But they never ask a user about how certain he is nor whether changing the 1-9 scale value up or down to match the user's confidence improves the validity of the outcome. There is no mathematical proof that making the judgments fuzzy can be depended on to improve the practical validity of the results. Fuzzy thinking advocates have determined that the judgments are always crisp numbers, whatever that may mean in metric topology so they can play with them to make them fuzzy. To justify their manipulations, they say that interval or fuzzy numbers might be an alternative. Interval is suitable when no reference is given to any value on the possible range. In contrast, it is thought that fuzzy numbers are appropriate when some values are preferred to others in the possible range. Note that, if all of the judgments are at the same confidence level and can be represented by crisp numbers, there is no point to using interval or fuzzy numbers. The point is to use fuzzy numbers as a suitable tool to represent uncertain judgments, rather than to "fuzzify" certain and crisp judgments. Of course we know from various publications that people blindly use fuzzy numbers in their work without verifying the uncertainty and that may be one reason why they are unable to ensure the validity of the fuzzy outcome.

There are many different ways to construct fuzzy numbers from raw information (see [2,4,5,6,8,10] for more details on fuzzy numbers and their construction). One wonders if more than one of these methods can be applied to the same problem they lead to the same outcome. It is very likely they do not, and then what?

A simple way is to construct a triangle fuzzy number for a judgment is by using its most possible value and its possible range (i.e., its minimum and maximum values). To illustrate, assume the most possible, minimum, and maximum values of a judgment in comparing A and B are 8, 6, and 9, respectively. Then the judgment can be represented by the triangular fuzzy number (6, 8, 9) in Figure 3. If more information is given, other types of fuzzy numbers (e.g., nonlinear fuzzy membership functions) can be constructed to best describe the situation.
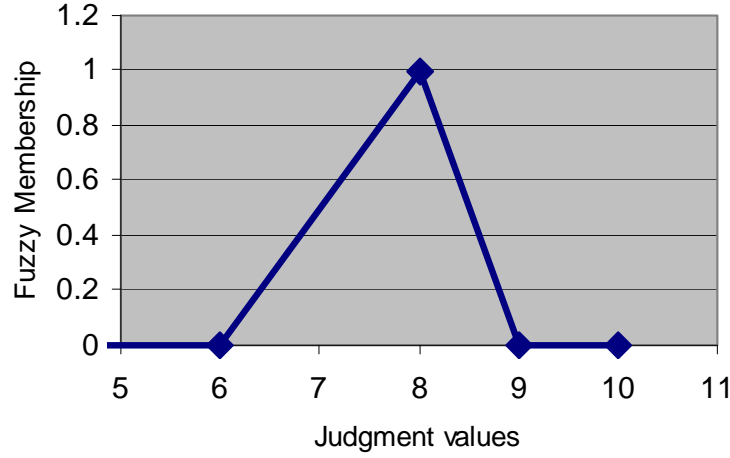
Fig. 3. Fuzzy triangle number $(6, 8, 9)_T$.

Finding the eigenvector/eigenvalue in the case of fuzzy judgments

When there is one or more fuzzy judgments in the model, there is an infinite set of eigenvector/ eigenvalue outputs. The problem turns out to be one of finding the "best" (maximum) eigenvalue in a multi-objective setting: small inconsistency index (<0.1), high fuzzy membership values (i.e., the closer the chosen judgments to the most possible values are, the better the result is). There are several different ways to set up such multi-objective optimization problems, such as:

❖ Minimize the maximum eigenvalue (i.e., minimize the inconsistency index) while setting constraints on fuzzy membership values (e.g., acceptable value(s) > 0.5).

❖ Maximize fuzzy membership values while setting constraint on the inconsistency index (e.g., <0.1).

❖ Run simulation to determine the entire range of maximum eigenvalues according to all possible judgment values at various fuzzy membership values, then ask the user to make a decision.

In this paper the third method is used.

6. Examples

We learn three things from the examples shown below:

❖ When the original judgments are good, by Wilkinson's theorem, small perturbations of them are close and some components of the derived vector may be a little better, other components a little worse. There is no gain in fuzzifying the matrix.

❖ If the judgments are partly good partly bad in the sense that they are somewhat far from the ratio's of the actual answer, fuzzifying the matrix does not improve the closeness of the vector to the actual value.

❖ When the judgments are poor both in capturing the cardinal ratios and reverse ordinal rank, fuzzifying can make the values of some components even farther away from the actual values.

The conclusion is, as we said in the opening discussion, one should never use fuzzy arithmetic on AHP judgment matrices. We await some convincing theory with proofs what one should apply a process that we are convinced with numerous counter-examples is shown to give untrustworthy answers in decision making. It appears that the observations made by Buede and Maxwell [1] are fully justified.

An Example where good judgments already give good outcomes and fuzzy methods do not improve it and can make it worse for some alternatives: Relative Consumption of Drinks in the United States.

Table 5 shows how an audience of about 30 people, using consensus (instead of the geometric mean to combine judgments in the AHP) to arrive at each judgment, provided judgments to estimate the *dominance* of

the consumption of drinks in the United States (which drink is consumed more in the US and how much more than another drink?). The derived vector of relative consumption and the actual vector, obtained by normalizing the consumption given in official statistical data sources, are at the bottom of the Table 5. Note that for several of the drinks the fuzzy outcome is not as close to the actual value as is the AHP eigenvector.

Table 5
Which drink is consumed more in the U.S.?

An Example of Estimation Using Judgments

Drink Consumption in the U.S.	Coffee	Wine	Tea	Beer	Sodas	Milk	Water
Coffee	1	9	5	2	1	1	1/2
Wine	1/9	1	1/3	1/9	1/9	1/9	1/9
Tea	1/5	2	1	1/3	1/4	1/3	1/9
Beer	1/2	9	3	1	1/2	1	1/3
Sodas	1	9	4	2	1	2	1/2
Milk	1	9	3	1	1/2	1	1/3
Water	2	9	9	3	2	3	1

The derived scale based on the judgments in the matrix is:

Coffee	Wine	Tea	Beer	Sodas	Milk	Water
.177	.019	.042	.116	.190	.129	.327

with a consistency ratio of .022.

The fuzzy outcome obtained from the matrix below is:

.198	.016	.033	.107	.165	.123	.358

with a consistency ratio of .017

The actual consumption (from statistical sources) is:

.180	.010	.040	.120	.180	.140	.330

Those who did the example did not know the answers at all until the results confirmed their judgments. The fuzzy algorithm found the best results with the following inputs shown in Table 6.

Table 6
Inputs that provide the best results with the fuzzy algorithm for the drink example

	Coffee	Wine	Tea	Beer	Sodas	Milk	Water
Coffee	1.000	10.501	5.179	2.430	1.473	1.486	0.514
Wine	0.100	1.000	0.500	0.167	0.099	0.111	0.036
Tea	0.200	2.043	1.000	0.277	0.257	0.263	0.073
Beer	0.400	5.636	3.600	1.000	0.500	1.000	0.400
Sodas	0.700	10.053	3.889	1.875	1.000	1.792	0.465
Milk	0.700	9.000	3.800	1.000	0.556	1.000	0.427
Water	1.900	27.824	13.708	2.407	2.152	2.386	1.000

A similar Outcome with a second Example: Relative Amount of Protein in Seven Foods.
Table 7 shows pairwise comparisons of seven foods with respect to their protein content. The question is which food has more protein. Because the AHP combines the objective problem of measurement with the subjective problem of meaning, it uses a scientifically justifiable process of combining group judgments to give greater objectivity to its findings even by using priorities for the judges themselves to raise their judgments to that power of importance. In this example and in several others below group judgments were

combined using the geometric mean but without weighting the judges. This avoids the issue that the AHP outcome somehow approximates to the fuzziness that may be associated with the many outcomes of the different judges treated as statistics.

Table 7
Which food has more protein?

Protein in Food	A	B	C	D	E	F	G
A: Steak	1	9	9	6	4	5	1
B: Potatoes	1/9	1	1	1/2	1/4	1/3	1/4
C: Apples	1/9	1	1	1/3	1/3	1/5	1/9
D: Soybean	1/6	2	3	1	1/2	1	1/6
E: Whole Wheat Bread	1/4	4	3	2	1	3	1/3
F: Tasty Cake	1/5	3	5	2	1/3	1	1/5
G: Fish	1	8	9	1	3	5	1

The derived scale and actual values are:

Steak	Potatoes	Apples	Soybean	W. Bread	T. Cake	Fish
.345	.031	.030	.065	.124	.078	.328

with a consistency ratio of .028. The fuzzy outcome obtained from the matrix below is;

.297	.038	.025	.087	.147	.081	.324

With inconsistency ratio of .018. The actual values are:

.370	.040	.000	.070	.110	.090	.320

The apple should not be included in the comparisons in any case because it has no protein.
The fuzzy algorithm found the best results with the following inputs shown in Table 8.

Table 8
Inputs that provide the best results with the fuzzy algorithm for the protein example.

	A	B	C	D	E	F	G
A: Steak	1.000	7.601	9.729	3.754	1.955	4.022	1.047
B: Potatoes	0.100	1.000	1.250	0.500	0.333	0.375	0.143
C: Apples	0.100	0.802	1.000	0.248	0.131	0.418	0.053
D: Soybean	0.300	2.117	4.000	1.000	0.500	1.000	0.375
E: Whole Wheat Bread	0.500	2.979	7.667	1.855	1.000	1.922	0.379
F: Tasty Cake	0.200	2.714	2.400	1.000	0.500	1.000	0.250
G: Fish	1.000	7.453	19.022	2.710	2.642	3.851	1.000

Example of Poor Outcome with Poor Judgments: Relative sizes of Areas

Fig. 4 shows five areas. The object is to compare them in pairs to reproduce their relative weights. The reader can apply the paired comparison process using the 1-9 scale and find its principal eigenvector and compute the compatibility index with the actual values given below to test the validity of the procedure.
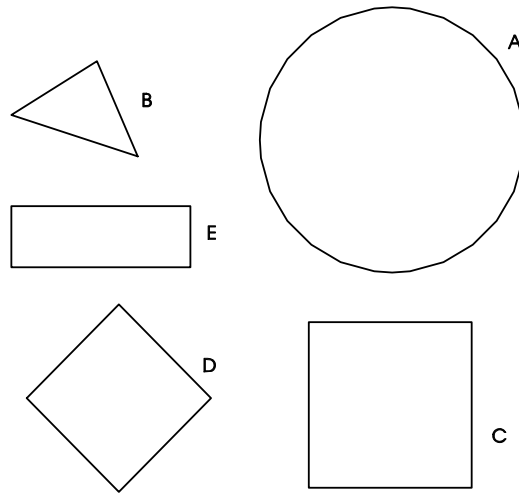
Fig. 4. Five Geometric Shapes

The inputs shown in Table 9 indicate that answers from poor judgments are not improved by fuzzy thinking.

Table 9
Judgments, outcomes, and actual relative sizes of the five geometric shapes.

Figure	Circle	Triangle	Square	Diamond	Rectangle	Eigen Vector Inc .631	Fuzzy vector Inc .467, .361	Actual Relative Size
Circle	1	12	2/3	5	12	.482	.495, .532	.471
Triangle	1/12	1	1/6	3	3/2	.085	.107, .099	.050
Square	3/2	6	1	2/3	1/4	.196	.196, .168	.234
Diamond	1/5	1/3	3/2	1	2	.101	.095, .093	.149
Rectangle	1/12	2/3	4	1/2	1	.136	.106, .108	.096

Fuzzy inputs for the first person's comparisons had two runs with two outcomes. Table 10 presents the fuzzy matrix for the first run.

Table 10
Fuzzy matrix for the first run of the five-geometric shapes example.

	Circle	Triangle	Square	Diamond	Rectangle

	Circle	Triangle	Square	Diamond	Rectangle
Circle	1	A=(11, 12, 13)	B=(0.001, 0.667, 1.167)	C=(4.5, 5, 5.5)	D=(11.5, 12, 12.5)
Triangle	1/A	1	E=(0.001, 0.167, 1.167)	F=(2.5, 3, 3.5)	G=(1.0, 1.5, 2.0)
Square	1/B	1/E	1	H=(0.167, 0.667, 1.167)	I=(0.001, 0.25, 0.75)
Diamond	1/C	1/F	1/H	1	J=(0.001, 0.5, 1.0)
Rectangle	1/D	1/G	1/I	1/J	1

The fuzzy algorithm found the best results for the first run with the inputs presented in Table 11.

Table 11
Inputs that provide the best results with the fuzzy algorithm for the first run of the five-geometric shape example.

	Circle	Triangle	Square	Diamond	Rectangle
Circle	1.000	11.489	1.337	4.915	11.557
Triangle	0.087	1.000	0.519	2.626	1.038
Square	0.748	1.927	1.000	1.122	0.679
Diamond	0.204	0.381	0.891	1.000	0.771
Rectangle	0.086	0.963	1.472	1.297	1.000

Table 12 presents the fuzzy matrix for the second run.

Table 12
Fuzzy matrix for the second run of the five-geometric shape example

	Circle	Triangle	Square	Diamond	Rectangle
Circle	1	A=(9, 10, 11)	B=(0.667, 1.167, 2.667)	C=(4.0, 5.0, 6.0)	D=(10.0, 11.0, 12.0)
Triangle	1/A	1	E=(0.001, 0.8, 1.8)	F=(1.5, 2.5, 3.5)	G=(0.5, 1.5, 2.5)
Square	1/B	1/E	1	H=(0.5, 1.5, 2.5)	I=(0.001, 0.75, 1.75)
Diamond	1/C	1/F	1/H	1	J=(0.001, 1.0, 2.0)
Rectangle	1/D	1/G	1/I	1/J	1

The fuzzy algorithm found the best results for the second run with the inputs presented in Table 13.

Table 13
Inputs that provide the best results with the fuzzy algorithm for the second run of the five-geometric shape example.

	Circle	Triangle	Square	Diamond	Rectangle
Circle	1.000	9.238	2.625	5.617	10.056
Triangle	0.108	1.000	0.670	2.408	1.230
Square	0.381	1.491	1.000	1.833	1.442
Diamond	0.178	0.415	0.546	1.000	0.984

Rectangle	0.099	0.813	0.694	1.012	1.000

As previously mentioned, there is one line of argument somewhere that fuzzy numbers are justified as they provide a valid way to aggregate judgments from a group of experts in the AHP. Note that the AHP has a formal way of combining group judgments and talks about spreading fuzziness. Aczel and Saaty [1] proved a theorem that group judgments can be synthesized suitably with the geometric mean. So far there has been no work proving that aggregating group judgments with fuzzy numbers provides better results than those in [1].

6. Conclusions

A judge, whether certain or uncertain of the accuracy of his/her judgments may be totally uninformed and ignorant. In that case there is no way to capture reality through his judgments. No fuzzifying can make his judgments more valid in practice. They could make them even less accurate. If the judge is well informed, the AHP has been demonstrated to obtain valid results from his/her judgments and perturbation theory shows that changing the numbers to cope with uncertainty would not change the answers substantially nor would they direct them towards the actual answer when that answer is known. Fuzzy set practitioners have been leading a parasitic existence (a parasite according to Webster's dictionary is one depending on another and not making adequate return- "we might add most of the time") by looking at all numbers as if they are subject to uncertainty and fuzzifying them purportedly to improve consistency without either giving good reasons for doing it because we know that good consistency does not imply greater validity, or proving that the results thus obtained are more valid than is obtained directly from the judgments.

In sum, we note that making poor judgments leads to poor outcomes and fuzzifying poor judgments still leads to poor outcomes. Making good judgments gives good (valid) answers with the AHP and fuzzifying these judgments is simply a perturbation that leaves the results where they are without producing uniformly better outcomes. Good judgments that are inconsistent can give better outcomes than consistent ones, thus fuzzifying these inconsistent judgments can lead to worse rather than better outcomes. The minimum one would expect is more valid outcomes which we have shown through examples are not obtained. For greater validity of a decision with the AHP, is there evidence that it is better to fuzzify the judgments? The answer is an unqualified **no.** It is not advisable to use fuzzy numbers. Proofs are needed to invalidate our conclusion that fuzzy thinking is an artifice that should not be used at all in the AHP.

References

[1] J. Aczel and T. Saaty, 1983. "Procedures for Synthesizing Ratio Judgements," Journal of Mathematical Psychology, Vol. 27, pp. 93-102.
[2] T. Allahviranloo and Hajari, T 2005. Numerical methods for approximation of fuzzy data. Applied Mathematics and Computation 169:16–33
[3] D. Buede and D.T. Maxwell, 1995. Rank Disagreement: A comparison of multi-criteria methodologies. *Journal of Multi-Criteria Decision Analysis* 4:1-21.
[4] C.-B. Cheng, 2004. Group Opinion Aggregation Based on a Grading Process: A Method for Constructing Triangular Fuzzy Numbers. Computers and Mathematics with Applications 48:1619-1632.
[5] D. Dubois and H. Prade. 1980. *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York, 393 pp.
[6] D. Dubois and H. Prade (Eds.) 2000. Fundamentals of Fuzzy Sets. Dordrecht: Kluwer Academic Publishers.
[7] C. Elkan, 1994. The Paradoxical Success of Fuzzy Logic. *IEEE Expert: Intelligent Systems and Their Applications*, 9(4):3-8.
[8] A. Kaufman, and M.M. Gupta. 1991. *Introduction to Fuzzy Arithmetic, Theory and Application.* Van Nostrand Reinhold Company, New York, 351 pp.
[9] J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

[10]   H.-J. Zimmermann, 2001. Fuzzy set theory and its applications (4th edition). Dordrecht; Boston: Kluwer Academic Publishers.

Reading List

[11]   Cengiz Kahraman, Tijen Ertay, Gülc¸in Büyüközkan 2006. A fuzzy optimization mode for QFD planning process using analytic network approach. European Journal of Operational Research 171:390–411.

[12]   Cengiz Kahraman, Ufuk Cebeci, and Da Ruan 2004. Multi-attribute comparison of catering service companies using fuzzy AHP: The case of Turkey. Int. J. Production Economics 87:171–184.

[13]   Chian-Son Yu 2001. A GP-AHP method for solving group decision-making fuzzy AHP problems. Computers & Operations Research 29(14):1969-2001.

[14]   Ching-Hsue Cheng 1997. Evaluating naval tactical missile systems by fuzzy AHP based on the grade value of membership function. European Journal of Operational Research 96(2):343-350.

[15]   Ching-Hsue Cheng and Don-Lin Mon 1994. Evaluating weapon system by Analytical Hierarchy Process based on fuzzy scales. Fuzzy Sets and Systems 63(1):1-10.

[16]   Da-Yong Chang 1996. Applications of the extent analysis method on fuzzy AHP European Journal of Operational Research 95:649-655.

[17]   Da-Yong Chang 1996. Applications of the extent analysis method on fuzzy AHP. European Journal of Operational Research 95(3):649-655.

[18]   Don-Lin Mon, Ching-Hsue Cheng and Jiann-Chern Lin 1994. Evaluating weapon system using fuzzy analytic hierarchy process based on entropy weight. Fuzzy Sets and Systems 62(2):127-134.

[19]   Ethem Tolga, Murat Levent Demircan and Cengiz Kahraman 2005. Operating system selection using fuzzy replacement analysis and analytic hierarchy process. International Journal of Production Economics 97(1):89-117.

[20]   F. Tunc¸ Bozbura, Ahmet Beskese, Cengiz Kahraman 2006. Prioritization of human capital measurement indicators using fuzzy AHP. Expert Systems with Applications xxx: xxx–xxx (on-line for now).

[21]   F.A. Lootsma 1997. Fuzzy Logic for Planning and Decision Making. Kluwer Academic Publishers, Dordrecht, 195 pp.

[22]   Gwo-Hshiung Tzeng, Cheng-Hsin Chiang, Chung-Wei Li 2006. Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL. Expert Systems with Applications xxx: xxx–xxx (on-line for now)

[23]   Heung-Suk Hwang 2004. Web-based multi-attribute analysis model for engineering project evaluation. Computers & Industrial Engineering 46(4):669-678.

[24]   He-Yau Kang, Amy H.I. Lee 2006. Priority mix planning for semiconductor fabrication by fuzzy AHP ranking. Expert Systems with Applications xxx: xxx–xxx (on-line for now)

[25]   Hua-Kai Chiou, Gwo-Hshiung Tzeng and Ding-Chou Cheng 2005. Evaluating sustainable fishing development strategies using fuzzy MCDM approach. Omega 33(3):223-234.

[26]   Jiuh-Biing Sheu 2004. A hybrid fuzzy-based approach for identifying global logistics strategies. Transportation Research Part E: Logistics and Transportation Review 40(1):39-61.

[27]   Ke-Jun Zhu, Yu Jing and Da-Yong Chang 1999. A discussion on Extent Analysis Method and applications of fuzzy AHP. European Journal of Operational Research 116(2):450-456.

[28]   L. Mikhailov 2004. Group prioritization in the AHP by fuzzy preference programming method. Computers & Operations Research 31:293–301.

[29]   L.C. Leung , D. Cao 2000. On consistency and ranking of alternatives in fuzzy AHP. European Journal of Operational Research 124:102–113.

[30]   F.A. Lootsma, 1999. Multi-criteria Decision Analysis via Ratio and Difference Judgement. Kluwer Academic Publishers, Dordrecht, 283 pp.

[31]   M. Weck, F. Klocke, H. Schell and E. Rüenauver 1997. Evaluating alternative production cycles using the extended fuzzy AHP method. European Journal of Operational Research 100(2):351-366.

[32]   Mei-Fang Chen and Gwo-Hshiung Tzeng 2004. Combining grey relation and TOPSIS concepts for selecting an expatriate host country. Mathematical and Computer Modelling 40(13):1473-1490.

[33]   Ming-Lung Hung, Wan-Fa Yang, Hwong-Wen Ma, Ya-Mei Yang 2006. A novel multiobjective programming approach dealing with qualitative and quantitative objectives for environmental management. Ecological Economics 56:584– 593.

[34]   Osman Kulak and Cengiz Kahraman 2005. Fuzzy multi-attribute selection among transportation companies using axiomatic design and analytic hierarchy process. Information Sciences 170(2-4):191-210.

[35]   P.J.M. Van Laarhoven, W. Pedrycz 1983. A fuzzy extention of Saaty's priority theory. Fuzzy Sets and Systems 11 (3):229–241.

[36]   R. J. Kuo, S. C. Chi and S. S. Kao 1999. A decision support system for locating convenience store through fuzzy AHP. Computers & Industrial Engineering, 37(1-2):323-326.

[37]   Rachung Yu, Gwo-Hshiung Tzeng 2006. A soft computing method for multi-criteria decision making with dependence and feedback. Applied Mathematics and Computation xxx: xxx–xxx (on-line for now)

[38]   Xu Ruoning and Zhai Xiaoyan 1992. Extensions of the analytic hierarchy process in fuzzy environment. Fuzzy Sets and Systems 52(3):251-257.

[39]   Ying-Feng Kuoa, Pang-Cheng Chen 2006. Selection of mobile value-added services for system operators using fuzzy synthetic evaluation. Expert Systems with Applications 30:612–620.