

## Chapter 2

# Fundamentals of the Analytic Hierarchy Process

Thomas L. Saaty

*University of Pittsburgh, Pittsburgh PA USA*

**Key words:** Analytic hierarchy process, ratio scale, subjective judgement, group decision making.

**Abstract:** The seven pillars of the analytic hierarchy process (AHP) are presented. These include: (1) ratio scales derived from reciprocal paired comparisons; (2) paired comparisons and the psychophysical origin of the fundamental scale used to make the comparisons; (3) conditions for sensitivity of the eigenvector to changes in judgements; (4) homogeneity and clustering to extend the scale from 1-9 to  $1-\infty$ ; (5) additive synthesis of priorities, leading to a vector of multi-linear forms as applied within the decision structure of a hierarchy or the more general feedback network to reduce multi-dimensional measurements to a uni-dimensional ratio scale; (6) allowing rank preservation (ideal mode) or allowing rank reversal (distributive mode); and (7) group decision making using a mathematically justifiable way for synthesising individual judgements which allows the construction of a cardinal group decision compatible with individual preferences. These properties of the AHP give it both theoretical support and broad application.

## 1. INTRODUCTION

The analytic hierarchy process (AHP) provides the objective mathematics to process the inescapably subjective and personal preferences of an individual or a group in making a decision. With the AHP and its generalisation, the analytic network process (ANP), one constructs hierarchies or feedback networks that describe the decision environment structure. The decision maker then makes judgements or performs measurements on pairs of elements with respect to a controlling element to derive ratio scales that are then synthesised throughout the structure to select the best alternative.

Fundamentally, the AHP works by developing priorities for alternatives and the criteria used to judge the alternatives. Criteria are selected by a decision maker (irrelevant criteria are those that are not included in the hierarchy). Selected criteria may be measured on different scales, such as weight and length, or may even be intangible for which no scales yet exist. Measurements on different scales, of course, cannot be directly combined. First, priorities are derived for the criteria in terms of their importance to achieve the goal, then priorities are derived for the performance of the alternatives on each criterion. These priorities are derived based on pairwise assessments using judgement or ratios of measurements from a scale if one exists. The process of prioritisation solves the problem of having to deal with different types of scales, by interpreting their significance to the values of the user or users. Finally, a weighting and adding process is used to obtain overall priorities for the alternatives as to how they contribute to the goal. This weighting and adding parallels what one would have done arithmetically prior to the AHP to combine alternatives measured under several criteria having the *same* scale to obtain an overall result (a scale that is often common to several criteria is money). With the AHP a multidimensional scaling problem is thus transformed to a uni-dimensional scaling problem.

The AHP can be viewed as a formal method for rational and explicit decision making. It possesses the seven fundamental properties, below. Subsequent sections examine each in greater detail.

*Normalised ratio scales* are central to the generation and synthesis of priorities, whether in the AHP or in any multicriteria method that needs to integrate existing ratio scale measurements with its own derived scales.

*Reciprocal paired comparisons* are used to express judgements semantically, and to automatically link them to a numerical and fundamental scale of absolute numbers (derived from stimulus-response relations). The principal right eigenvector of priorities is then derived; the eigenvector shows the dominance of each element with respect to the other elements. Inconsistency in judgement is allowed and a measure for it is provided which can direct the decision maker in both improving judgement and arriving at a better understanding of the problem. The AHP has at least three modes for arriving at a ranking of the alternatives: *relative*, which *ranks* a few alternatives by comparing them in pairs (particularly useful in new and exploratory decisions), *absolute*, which *rates* an unlimited number of alternatives one at a time on intensity scales constructed separately for each covering criterion (particularly useful in decisions where there is considerable knowledge to judge the relative importance of the intensities), and *benchmarking*, which ranks alternatives by including a known alternative in the group and comparing the others against it.

*Sensitivity of the principal right eigenvector* to perturbation in judgements limits the number of elements in each set of comparisons to a few and requires that they be homogeneous.

*Homogeneity and clustering* are used to extend the fundamental scale gradually from cluster to adjacent cluster, eventually enlarging the scale from 1-9 to 1- $\infty$ .

*Synthesis that can be extended to dependence and feedback* is applied to the derived ratio scales to create a uni-dimensional ratio scale for representing the overall outcome. Synthesis of the scales derived in the decision structure can only be made to yield correct outcomes on known scales by additive weighting.

*Rank preservation and reversal* can be shown to occur without adding or deleting criteria, such as by simply introducing enough copies of an alternative. This leaves no doubt that rank reversal is as intrinsic to decision making as rank preservation also is.

*Group judgements* must be integrated one at a time carefully and mathematically, taking into consideration, when desired, the experience, knowledge, and power of each person involved in the decision. The AHP's *cardinal* ratio scale preferences allow one the *possibility* of constructing a social utility function—an impossibility when using *ordinal* preferences. To deal with a large group requires the use of questionnaires and statistical procedures for large samples.

## 2. RATIO SCALES

A *ratio* is the relative value or quotient  $a/b$  of two quantities  $a$  and  $b$  of the same kind; it is called *commensurate* if it is a rational number, otherwise it is *incommensurate*. A statement of the equality of two ratios  $a/b$  and  $c/d$  is called *proportionality*. A ratio scale is a set of numbers that is invariant under a similarity transformation (multiplication by a positive constant). The constant cancels when the ratio of any two numbers is formed. Either pounds or kilograms can be used to measure weight, but the ratio of the weight of two objects is the same for both scales. An extension of this idea is that the weights of an entire set of objects, whether in pounds or in kilograms, can be standardised to read the same by normalising. In general if the readings from a ratio scale are  $aw_i^*$ ,  $i=1, \dots, n$ , the standard form is given by  $w_i = aw_i^* / aw_i^* = w_i^* / w_i^*$  as a result of which we have  $\sum w_i = 1$ , and the  $w_i$ ,  $i=1, \dots, n$ , are said to be normalised. We no longer need to specify whether weight for example is given in pounds or in kilograms or in another kind of unit. The weights (2.21, 4.42) in pounds and (1, 2) in kilograms, are both given by (1/3, 2/3) in the standard ratio scale form.

The relative ratio scale derived from a pairwise comparison reciprocal matrix of judgements is derived by solving:

$$\begin{cases} \sum_{j=1}^n a_{ij} w_j = \lambda_{max} w_i \\ \sum_{i=1}^n w_i = 1 \end{cases} \quad (2.1)$$

with  $a_{ji}=1/a_{ij}$  or  $a_{ij}a_{ji}=1$  (the reciprocal property),  $a_{ij}>0$  (thus  $\mathbf{A}$  is known as a positive matrix) whose solution, known as the principal right eigenvector, is normalised. A relative ratio scale does not need a unit of measurement.

When  $a_{ij}a_{jk} = a_{ik}$ , the matrix  $\mathbf{A}=(a_{ij})$  is said to be consistent and its principal eigenvalue is equal to  $n$ . Otherwise, it is simply reciprocal. The general eigenvalue formulation given in (2.1) is obtained by perturbation of the following consistent formulation:

$$\mathbf{A} \mathbf{w} = \begin{matrix} \mathbf{A}_1 & \cdots & \mathbf{A}_n \\ \vdots & & \\ \mathbf{A}_n & \cdots & \mathbf{A}_n \end{matrix} \begin{bmatrix} \frac{w_1}{w_1} & \cdots & \frac{w_1}{w_n} \\ \frac{w_1}{w_1} & \cdots & \frac{w_1}{w_n} \\ \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \cdots & \frac{w_n}{w_n} \\ \frac{w_n}{w_1} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = n \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = n \mathbf{w} \quad (2.2)$$

where  $\mathbf{A}$  has been multiplied on the right by the transpose of the vector of weights  $\mathbf{w}=(w_1, \dots, w_n)$ . The result of this multiplication is  $n\mathbf{w}$ . Thus, to recover the scale from the matrix of ratios, one must solve the problem  $\mathbf{A}\mathbf{w}=n\mathbf{w}$  or  $(\mathbf{A}-n\mathbf{I})\mathbf{w} = \mathbf{0}$ . This is a system of homogeneous linear equations. It has a nontrivial solution if and only if the determinant of  $\mathbf{A}-n\mathbf{I}$  vanishes, that is,  $n$  is an eigenvalue of  $\mathbf{A}$ . Now  $\mathbf{A}$  has unit rank since every row is a constant multiple of the first row. Thus, all its eigenvalues except one are zero. The sum of the eigenvalues of a matrix is equal to its trace, that is, the sum of its diagonal elements. In this case, the trace of  $\mathbf{A}$  is equal to  $n$ . Thus  $n$  is an eigenvalue of  $\mathbf{A}$ , and one has a nontrivial solution. The solution consists of positive entries and is unique to within a multiplicative constant.

The discrete formulation given in (2.1) above generalises to the continuous case through Fredholm's integral equation of the second kind and is given by:

$$\int_a^b K(s, t)w(t)dt = \lambda_{max} w(s), \quad \int_a^b w(s)ds = 1 \quad (2.3)$$

where instead of the matrix  $\mathbf{A}$  we have as a positive kernel,  $K(s,t) > 0$ . Note that the entries in a matrix  $\mathbf{A}$  depend on the two variables  $i$  and  $j$  which assume discrete values. Thus, the matrix itself depends on these discrete variables, and its generalisation, the kernel function, also depends on two (continuous) variables. The reason for calling it a kernel is the role it plays in the integral, where we cannot determine the exact form of the solution without knowing the kernel. The standard reciprocal form of (2.3) is written by moving the eigenvalue to the left hand side. As in the finite case, we have the reciprocal property and the consistency relation (2.4).

$$\begin{cases} K(s,t)K(t,s) = 1 \\ K(s,t)K(t,u) = K(s,u) \quad \forall s,t,u \end{cases} \tag{2.4}$$

An example of this type of kernel is  $K(s,t) = e^{-s-t} = e^s/e^t$ . It follows by putting  $s=t=u$ , that  $K(s,s) = 1$  for all  $s$  which is analogous to having ones down the diagonal of the matrix in the discrete case. A value of  $\lambda$  for which Fredholm's equation has a nonzero solution  $w(t)$  is called a characteristic value (or its reciprocal is called an eigenvalue) and the corresponding solution is called an eigenfunction. An eigenfunction is determined to within a multiplicative constant. If  $w(t)$  is an eigenfunction corresponding to the characteristic value  $\lambda$  and if  $C$  is an arbitrary constant, we can easily see by substituting in the equation that  $Cw(t)$  is also an eigenfunction corresponding to the same  $\lambda$ . The value  $\lambda=0$  is not a characteristic value because we have the corresponding solution  $w(t)=0$  for every value of  $t$ , which is the trivial case, excluded in our discussion.

A matrix is consistent if and only if it has the form  $\mathbf{A}=(w_i/w_j)$  which is equivalent to multiplying a column vector that is the transpose of  $(w_1, \dots, w_n)$  by the row vector  $(1/w_1, \dots, 1/w_n)$ . As we see below, the kernel  $K(s,t)$  is separable and can be written as

$$K(s,t) = k_1(s)k_2(t) \tag{2.5}$$

**Theorem**  $K(s,t)$  is consistent if and only if it is separable of the form:

$$K(s,t) = k(s) / k(t) \tag{2.6}$$

**Theorem** If  $K(s,t)$  is consistent, the solution of (2.3) is given by

$$w(s) = \frac{k(s)}{\int_s k(s)ds} \tag{2.7}$$

In the discrete case, the normalised eigenvector is independent of whether all the elements of the pairwise comparison matrix  $\mathbf{A}$  are multiplied by the same constant  $a$  or not, and thus we can replace  $\mathbf{A}$  by  $a\mathbf{A}$  and obtain the same eigenvector. Generalising this result we have:

$$K(as, at) = aK(s, t) = k(as) / k(at) = ak(s) / k(t) \quad (2.8)$$

which means that  $K$  is a homogeneous function of order one. In general, when  $f(ax_1, \dots, ax_n) = a^n f(x_1, \dots, x_n)$  holds,  $f$  is said to be homogeneous of order  $n$ . Because  $K$  is a degenerate kernel, we can replace  $k(s)$  above by  $k(as)$  and obtain  $w(as)$ . We have now derived from considerations of ratio scales the following condition to be satisfied by a ratio scale:

**Theorem** *A necessary and sufficient condition for  $w(s)$  to be an eigenfunction solution of Fredholm's equation of the second kind, with a consistent kernel that is homogeneous of order one, is that it satisfy the functional equation*

$$w(as) = bw(s), \text{ where } b = \alpha a. \quad (2.9)$$

We have for the general damped periodic response function  $w(s)$ ,

$$w(s) = Ce^{\log b \left( \frac{\log s}{\log a} \right)} P \left( \frac{\log s}{\log a} \right) \quad (2.10)$$

where  $P$  is periodic of period 1 and  $P(0) = 1$ .

We can write this solution as

$$v(u) = C_1 e^{-\beta u} P(u) \quad (2.11)$$

where  $P(u)$  is periodic of period 1,  $u = \log s / \log a$  and  $\log ab = -\beta$ ,  $\beta > 0$ . It is interesting to observe the logarithmic function appear as part of the solution. It gives greater confirmation to the Weber-Fechner law developed in the next section.

### 3. PAIRED COMPARISONS AND THE FUNDAMENTAL SCALE

Instead of assigning two numbers  $w_i$  and  $w_j$  and forming the ratio  $w_i/w_j$  we assign a single number drawn from the fundamental 1-9 scale of absolute

numbers to represent the ratio  $(w_i/w_j)/1$ . It is a nearest integer approximation to the ratio  $w_i/w_j$ . The derived scale will reveal what the  $w_i$  and  $w_j$  are. This is a central fact about the relative measurement approach of the AHP and the need for a fundamental scale.

In 1846, Weber found, for example, that people holding different weights in their hand, could distinguish between a weight of 20 g and a weight of 21 g, but could not if the second weight is only 20.5 g. On the other hand, while they could not distinguish between 40 g and 41 g, they could between 40 g and 42 g, and so on at higher levels. We need to increase a stimulus  $s$  by a minimum amount  $\Delta s$  to reach a point where our senses can first discriminate between  $s$  and  $s+\Delta s$ . The amount  $\Delta s$  is called the just noticeable difference (jnd). The ratio  $r=\Delta s/s$  does not depend on  $s$ . Weber's law states that change in sensation is noticed when the stimulus is increased by a constant percentage of the stimulus itself. This law holds in ranges where  $\Delta s$  is small when compared with  $s$ , and hence in practice it fails to hold when  $s$  is either too small or too large. Aggregating or decomposing stimuli as needed into clusters or hierarchy levels is an effective way to extend the use of this law.

In 1860, Fechner considered a sequence of just noticeable increasing stimuli. He denotes the first one by  $s_0$ . The next just noticeable stimulus is given by

$$s_1 = s_0 + \Delta s_0 = s_0 + \frac{\Delta s_0}{s_0} s_0 = s_0(1+r) \quad (3.1)$$

based on Weber's law. Similarly,

$$s_2 = s_1 + \Delta s_1 = s_1(1+r) = s_0(1+r)^2 \equiv s_0\alpha^2. \quad (3.2)$$

In general,

$$s_n = s_{n-1}\alpha = s_0 \alpha^n \quad (n=0, 1, 2, \dots). \quad (3.3)$$

Thus, stimuli of noticeable differences follow sequentially in a geometric progression. Fechner noted that the corresponding sensations should follow each other in an arithmetic sequence at the discrete points at which just noticeable differences occur. However, the latter are obtained when we solve for  $n$ . We have

$$n = \frac{(\log s_n - \log s_0)}{\log \alpha} \quad (3.4)$$

and sensation is a linear function of the logarithm of the stimulus. Thus, if  $M$  denotes the sensation and  $s$  the stimulus, the psychophysical law of Weber-Fechner is given by

$$M = a \log s + b, \quad a \neq 0. \quad (3.5)$$

We assume that the stimuli arise in making pairwise comparisons of relatively comparable activities. We are interested in responses whose numerical values are in the form of ratios. Thus  $b=0$ , from which we must have  $\log s_0=0$  or  $s_0=1$ , which is possible by calibrating a unit stimulus. Here the unit stimulus is  $s_0$ . The next noticeable stimulus is  $s_1 = s_0\alpha = \alpha$  which yields the second noticeable response  $a(\log\alpha)$ . The third noticeable stimulus is  $s_2=s_0\alpha^2$  which yields a response of  $2a(\log\alpha)$ . Thus, we have for the different responses:

$$M_0 = a \log s_0, \quad M_1 = a \log \alpha, \quad M_2 = 2a \log \alpha, \quad \dots, \quad M_n = na \log \alpha. \quad (3.6)$$

While the noticeable ratio stimulus increases geometrically, the response to that stimulus increases arithmetically. Note that  $M_0=0$  and there is no response. By dividing each  $M_i$  by  $M_1$  we obtain the sequence of absolute numbers 1, 2, 3, ... of the fundamental 1-9 scale. Paired comparisons are made by identifying the less dominant of two elements and using it as the unit of measurement. One then determines, using the scale 1-9 or its verbal equivalent, how many times more the dominant member of the pair is than this unit. In making paired comparisons, we use the nearest integer approximation from the scale, relying on the insensitivity of the eigenvector to small perturbations (discussed below). The reciprocal value is then automatically used for the comparison of the less dominant element with the more dominant one. Despite the foregoing derivation of the scale in the form of integers, someone might think that other scale values would be better, for example using 1.3 in the place of 2. Imagine comparing the magnitude of two people with respect to the magnitude of one person and using 1.3 for how many instead of 2.

We note that there may be elements that are closer than 2 on the 1-9 scale, and we need a variant of the foregoing. Among the elements that are close, we select the smallest. Observe the incremental increases between that smallest one and the rest of the elements in the close group. We now consider these increments to be new elements and pairwise compare them on the scale 1-9. If two of the increments are themselves closer than 2 we treat them as identical, assigning a 1 (we could carry this on ad infinitum). In the end, each component of the eigenvector of comparisons for the increments is added to unity to yield the un-normalised priorities of the close elements for



that criterion. Note that only the least of these close elements is used in comparisons with the other elements that can be compared directly using the normal 1-9 scale. Its priority is used to multiply the priorities of these close elements and finally the priorities of all the elements are re-normalised.

How large should the upper value of the scale be? Qualitatively, people have a capacity to divide their response to stimuli into three categories: high, medium and low. They also have the capacity to refine this division by further subdividing each of these intensities of responses into high, medium and low, thus yielding in all nine subdivisions. It turns out, from the requirement of homogeneity developed below, that to maintain stability (and limit inconsistency), our minds work with a few elements at a time.

#### 4. SENSITIVITY OF THE PRINCIPAL EIGENVECTOR

To a first order approximation, perturbation  $\Delta \mathbf{w}_1$  in the principal eigenvector  $\mathbf{w}_1$  from perturbation  $\Delta \mathbf{A}$  in the consistent matrix  $\mathbf{A}$  is given by:

$$\Delta \mathbf{w}_1 = \sum_{j=2}^n (\mathbf{v}_j^T \Delta \mathbf{A} \mathbf{w}_1 / (\lambda_1 - \lambda_j) \mathbf{v}_j^T \mathbf{w}_j) \mathbf{w}_j \quad (4.1)$$

The eigenvector  $\mathbf{w}_1$  is insensitive to perturbation in  $\mathbf{A}$ , if the principal eigenvalue  $\lambda_1$  is separated from the other eigenvalues  $\lambda_j$ , here assumed to be distinct, and none of the products  $\mathbf{v}_j^T \mathbf{w}_j$  of left and right eigenvectors is small. We should recall that the nonprincipal eigenvectors need not be positive in all components, and they may be complex. One can show that all the  $\mathbf{v}_j^T \mathbf{w}_j$  are of the same order, and that  $\mathbf{v}_j^T \mathbf{w}_j$ , the product of the normalised left and right principal eigenvectors, is equal to  $n$ . If  $n$  is relatively small and the elements being compared are homogeneous, none of the components of  $\mathbf{w}_1$  is arbitrarily small and correspondingly, none of the components of  $\mathbf{v}_1^T$  is arbitrarily small. Their product cannot be arbitrarily small, and thus  $\mathbf{w}$  is insensitive to small perturbations of the consistent matrix  $\mathbf{A}$ . The conclusion is that  $n$  must be small, and one must compare homogeneous elements.

#### 5. CLUSTERING TO EXTEND THE SCALE FROM 1-9 TO 1- $\infty$

In Figure 1, an unripe cherry tomato is eventually and indirectly compared with a large watermelon by first comparing it with a small tomato

and a lime, the lime is then used again in a second cluster with a grapefruit and a honey dew where we then divide by the weight of the lime and then multiply by its weight in the first cluster, and then use the honey dew again in a third cluster and so on. In the end we have a comparison of the unripe cherry tomato with the large watermelon and would accordingly extended the scale from 1-9 to 1-721.








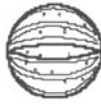
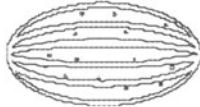
 .07 Unripe Cherry Tomato	 .28 Small Green Tomato	 .65 Lime
 .08 Lime $\frac{.08}{.08} = 1$ $.65 \times 1 = .65$	 .22 Grapefruit $\frac{.22}{.08} = 2.75$ $.65 \times 2.75 = 1.79$	 .70 Honeydew $\frac{.70}{.08} = 8.75$ $.65 \times 8.75 = 5.69$
 .10 Honeydew $\frac{.10}{.10} = 1$ $5.69 \times 1 = 5.69$	 .30 Sugar Baby Watermelon $\frac{.30}{.10} = 3$ $5.69 \times 3 = 17.07$	 .60 Oblong Watermelon $\frac{.60}{.10} = 6$ $5.69 \times 6 = 34.14$
This means that $34.14 / .07 = 487.7$ unripe cherry tomatoes are equal to the oblong watermelon.		

Figure 1. Comparisons according to volume.

Such clustering is essential, and must be done separately for each criterion. We should note that in most decision problems, there may be one or two levels of clusters and conceivably it may go up to three or four adjacent ranges of homogeneous elements (Maslow put them in seven groupings). Very roughly we have in decreasing order of importance: (1) survival, health, family, friends and basic religious beliefs some people were known to die for; (2) career, education, productivity and lifestyle; (3) political and social beliefs and contributions; (4) beliefs, ideas, and things that are flexible and it does not matter exactly how one advocates or uses them. These categories can be generalised to a group, a corporation, or a government. For very important decisions, two categories may need to be considered. Note that the priorities in two adjacent categories would be sufficiently different, one being an order of magnitude smaller than the other, that in the synthesis, the priorities of the elements in the smaller set have little effect on the decision. We do not have space to show how some *undesirable* elements can be compared among themselves and gradually

extended to compare them with *desirable* ones as above. Thus one can go from negatives to positives but keep the measurement of the two types positive, by eventually clustering them separately.

**6. SYNTHESIS: HOW TO COMBINE TANGIBLES WITH INTANGIBLES – ADDITIVE VS MULTIPLICATIVE**

Let  $H$  be a complete hierarchy with  $h$  levels. Let  $\mathbf{B}_k$  be the priority matrix of the  $k$ th level,  $k=2, \dots, h$ . If  $\mathbf{W}'$  is the global priority vector of the  $p$ th level with respect to some element  $z$  in the  $(p-1)$ st level, then the priority vector  $\mathbf{W}$  of the  $q$ th level ( $p < q$ ) with respect to  $z$  is given by the multilinear (and thus nonlinear) form,

$$\mathbf{W} = \mathbf{B}_q \mathbf{B}_{q-1} \cdots \mathbf{B}_{p+1} \mathbf{W}' \tag{6.1}$$

The global priority vector of the lowest level with respect to the goal is given by,

$$\mathbf{W} = \mathbf{B}_h \mathbf{B}_{h-1} \cdots \mathbf{B}_2 \mathbf{W}' \tag{6.2}$$

In general,  $\mathbf{W}'$  equals 1. The sensitivity of the bottom level alternatives with respect to changes in the weights of elements in any level can be studied by means of this multilinear form.

Assume that a family is considering buying a house and there are three houses to consider A, B, and C. Four factors dominate their thinking: house price, remodeling costs, house size as reflected by its footage, and style of the house, which is an intangible. They have looked at three houses with numerical data shown below on the quantifiables (Figure 2).

**Choosing the Best House**

	Price (\$1000)	Remodeling Costs (\$300)	Size (sq. ft.)	Style
<b>A</b>	200	150	3000	Colonial
<b>B</b>	300	50	2000	Ranch
<b>C</b>	500	100	5500	Split Level

Figure 2. Ranking houses on four criteria.

If we add the costs on price and modelling and normalise we obtain respectively (A,B,C)=(.269,.269,.462). Now let us see what is needed for normalisation to yield the same result. First, we normalise for each of the quantifiable factors. Then we must normalise the factors measured with respect to a single scale (Figure 3).

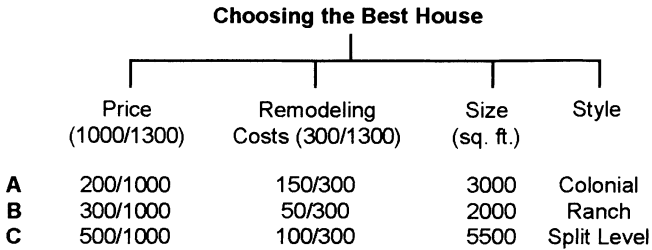


Figure 3. Normalising the measurements.

Here we learn two important lessons to be used in the general approach. Normalising the alternatives for the two criteria involving money in terms of the money involved on both criteria leads to relative weights of importance for the criteria. Here for example Price is in the ratio of about three to one when compared with Remodelling Cost and when compared with the latter with respect to the goal of choosing the best house, it is likely to be assigned the value “moderate” which is nearly three times more as indicated by the measurements. Here the criteria Price and Remodelling Cost derive their priorities only from the alternatives because they are equally important factors, although they can also acquire priorities from higher level criteria as to their functional importance with respect to the ease and availability of different amounts of money. We now combine the two factors with a common scale by weighting and adding (Figure 4).

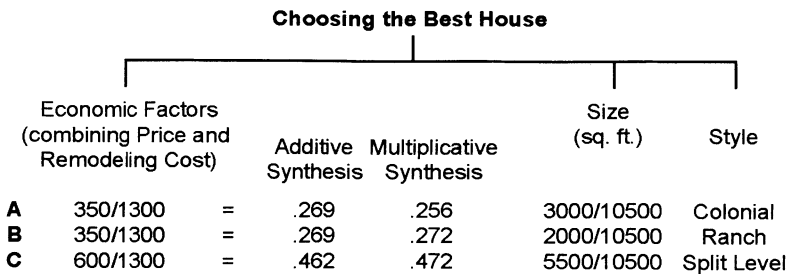


Figure 4. Combining the two costs through additive or multiplicative syntheses.

The left column and its decimal values in the second column give the exact value of the normalised dollars spent on each house obtained by additive synthesis (weighting and adding). By aggregating the two factor measured with dollars into a single factor, one then makes the decision as to which house to buy by comparing the three criteria as to their importance with respect to the goal.

The second lesson is that when the criteria have different measurements, their importance cannot be determined from the bottom up through measurement of the alternatives, but from the top down, in terms of the goal. The same process of comparison of the criteria with respect to the goal is applied to all criteria if, despite the presence of a physical scale, they are assumed to be measurable on different scales as they might when actual values are unavailable or when it is thought that such measurement does not reflect the relative importance of the alternatives with respect to the given criterion. Imagine that no physical scale of any kind is known! We might note in passing that the outcome of this process of comparison with respect to higher level criteria yields meaningful (not arbitrary) results as noted by two distinguished proponents of multi-attribute value theory (MAVT) Buede and Maxwell (1995), who wrote about their own experiments in decision making:

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.

Multiplicative synthesis, as in the third column of numbers above, done by raising each number in the two columns in the previous table to the power of its criterion measured in the relative total dollars under it, multiplying the two outcomes for each alternative and normalising, does not yield the exact answer obtained by adding dollars! In addition, A and B should have the same value, but they do not with multiplicative synthesis. The multiplicative “solution” devised for the fallacy of always preserving rank and avoiding inconsistency fails, because it violates the most basic of several requirements mentioned in the introduction to this chapter.

Multiplicative and additive syntheses are related analytically through approximation. If we denote by  $a_i$  the priority of the  $i$ th criterion,  $i=1, \dots, n$ , and by  $x_i$ , the priority of alternative  $x$  with respect to the  $i$ th criterion, then

$$\begin{aligned} \prod x_i^{a_i} &= \exp \log \prod x_i^{a_i} = \exp (\sum \log x_i^{a_i}) = \exp (\sum a_i \log x_i) \\ &\approx 1 + (\sum a_i \log x_i) \\ &\approx 1 + \sum (a_i x_i - a_i) = \sum a_i x_i \end{aligned} \tag{6.3}$$

If desired, one can include a remainder term to estimate the error. With regard to additive and multiplicative syntheses being close, one may think that in the end it does not matter which one is used, but it does. Saaty and Hu (1998) have shown that despite such closeness on every matrix of consistent judgements in a decision, the synthesised outcomes by the two methods not only lead to different final priorities (which can cause a faulty allocation of resources) but more significantly to *different rankings* of the alternatives. For all these problems, but more significantly because it does not generalise to dependence and feedback even with consistency guaranteed, and because of the additive nature of matrix multiplication needed to compute feedback in network circuits to extend the AHP to the ANP, I do not recommend ever using multiplicative synthesis. It can lead to an undesirable ranking of the alternatives of a decision.

## 7. RANK PRESERVATION AND REVERSAL

### 7.1 Theoretical and Practical Issues

Given the assumption that the alternatives of a decision are completely independent of one another, can and should the introduction (deletion) of new (old) alternatives change the rank of some alternatives without introducing new (deleting old) criteria, so that a less preferred alternative becomes most preferred? Incidentally, how one prioritises the criteria and subcriteria is even more important than how one does the alternatives which are themselves composites of criteria. Can rank reverse among the criteria themselves if new criteria are introduced? Why should that not be as critical a concern? The answer is simple. In its original form utility theory assumed that criteria could not be weighted and the only important elements in a decision were the alternatives and their utilities under the various criteria. Today, utility theorists imitate the AHP by rating, and some even by comparing the criteria, somehow. There was no concern then about what would happen to the ranks of the alternatives should the criteria weights themselves change as there were none. The tendency, even today, is to be unconcerned about the theory of rank preservation and reversal among the criteria themselves.

The house example of the previous section teaches us an important lesson. If we add a fourth house to the collection, the priority weights of the criteria Price and Remodelling Cost would change accordingly. Thus the measurements of the alternatives and their number which we call structural factors, always affect the importance of the criteria. When the criteria are incommensurate and their functional priorities are determined in terms of yet

higher level criteria or goals, one must still weight such functional importance of the criteria by the structural effect of the alternatives. What is significant in all this is that the importance of the criteria always depends on the measurements of the alternatives. If we assume that the alternatives are measured on a different scale for each criterion, it becomes obvious that normalisation is the instrument that provides the structural effect to update the importance of the criteria in terms of what alternatives there are. Finally, the priorities of the alternatives are weighted by the priorities of the criteria that depend on the measurements of the alternatives. This implies that the overall ranking of any alternative depends on the measurement and number of all the alternatives. To always preserve rank means that the priorities of the criteria should not depend on the measurements of the alternatives but should only derive from their own functional importance with respect to higher goals. This implies that the alternatives should not depend on the measurements of other alternatives. Thus, one way to always preserve rank is to *rate* the alternatives one at a time. In the AHP, this is done through absolute measurement with respect to a complete set of intensity ranges with the largest value intensity value equal to one. It is also possible to preserve rank in relative measurement by using an ideal alternative with full value of one for each criterion.

The logic about what can or should happen to rank when the alternatives *depend* on each other has always been that *anything* can happen. Thus, when the criteria functionally depend on the alternatives, which implies that the alternatives, which of course depend on the criteria, would then depend on the alternatives themselves, rank may be allowed to reverse. The Analytic Network Process (ANP) is the generalisation of the AHP to deal with ranking alternatives when there is functional dependence and feedback of any kind. Even here, one can have a decision problem with dependence among the criteria, but with no dependence of criteria on alternatives and rank may still need to be preserved. The ANP takes care of functional dependence, but if the criteria do not depend on the alternatives, the latter are kept out of the supermatrix and ranked precisely as in a hierarchy (Saaty 1996).

Examples of rank reversal abound in practice, and they do not occur because new criteria are introduced. The requirement that rank always be preserved or that it should be preserved with respect to irrelevant alternatives is not universally accepted. To every rule or generalisation that one may wish to set down about rank, it is possible to find a counterexample that violates that rule. Here is the last and most extreme form of four variants of an attempt to qualify what should happen to rank given by Luce and Raiffa (1957), each of which is followed by a counterexample. They state it but and then reject it. *The addition of new acts to a decision problem under*

*uncertainty never changes old, originally non-optimal acts into optimal ones. The all-or-none feature of the last form may seem a bit too stringent ... a severe criticism is that it yields unreasonable results.* The AHP has a theory and implementation procedures and guidelines for when to preserve rank and when to allow it to reverse. One mode of the AHP allows an irrelevant alternative to cause reversal among the ranks of the original alternatives.

## 7.2 Selecting the Distributive or Ideal Mode

The distributive mode of the AHP produces preference scores by normalising the performance scores; it takes the performance score received by each alternative and divides it by the sum of performance scores of all alternatives under that criterion. This means that with the Distributive mode the preference for any given alternative would go up if we reduce the performance score of another alternative or remove some alternatives. The Ideal mode compares each performance score to a fixed benchmark such as the performance of the best alternative under that criterion. This means that with the Ideal mode the preference for any given alternative is independent of the performance of other alternatives, except for the alternative selected as a benchmark. Saaty and Vargas (1993) have shown by using simulation, that there are only minor differences produced by the two synthesis modes. This means that the decision should select one or the other if the results diverge beyond a given set of acceptable data.

The following guidelines were developed by Millet and Saaty (1999) to reflect the core differences in translating performance measures to preference measures of alternatives. *The Distributive (dominance) synthesis mode should be used when the decision maker is concerned with the extent to which each alternative dominates all other alternatives under the criterion. The Ideal (performance) synthesis mode should be used when the decision maker is concerned with how well each alternative performs relative to a fixed benchmark.* In order for dominance to be an issue, the decision maker should regard inferior alternatives as relevant even after the ranking process is completed. This suggests a simple test for the use of the Distributive mode: *if the decision maker indicates that the preference for a top ranked alternative under a given criterion would improve if the performance of any lower ranked alternative was adjusted downward, then one should use the Distributive synthesis mode.* To make this test more actionable we can ask the decision maker to imagine the amount of money he or she would be willing to pay for the top ranked alternative. If the decision maker would be willing to pay more for a top ranked alternative



after learning that the performance of one of the lower-ranked alternatives was adjusted downward, then the Distributive mode should be used.

Consider selecting a car: Two different decision makers may approach the same problem from two different points of views even if the criteria and standards are the same. The one who is interested in "getting a well performing car" should use the Ideal mode. The one who is interested in "getting a car that stands out" among the alternatives purchased by co-workers or neighbours, should use the Distributive mode.

## 8. GROUP DECISION MAKING

Here we consider two issues in group decision making. The first is how to aggregate individual judgements, and the second is how to construct a group choice from individual choices.

### 8.1 How to Aggregate Individual Judgements

Let the function  $f(x_1, x_2, \dots, x_n)$  for synthesising the judgements given by  $n$  judges, satisfy the following conditions:

1. *Separability condition (S)*:  $f(x_1, x_2, \dots, x_n) = g(x_1)g(x_2)\dots g(x_n)$  for all  $x_1, x_2, \dots, x_n$  in an interval  $P$  of positive numbers, where  $g$  is a function mapping  $P$  onto a proper interval  $J$  and is a continuous, associative and cancellative operation. [(S) means that the influences of the individual judgements can be separated as above.]
2. *Unanimity condition (U)*:  $f(x, x, \dots, x) = x$  for all  $x$  in  $P$ . [(U) means that if all individuals give the same judgement  $x$ , that judgement should also be the synthesised judgement.]
3. *Homogeneity condition (H)*:  $f(ux_1, ux_2, \dots, ux_n) = uf(x_1, x_2, \dots, x_n)$  where  $u > 0$  and  $x_k, ux_k$  ( $k=1, 2, \dots, n$ ) are all in  $P$ . [For ratio judgements (H) means that if all individuals judge a ratio  $u$  times as large as another ratio, then the synthesised judgement should also be  $u$  times as large.]
4. *Power conditions (P<sub>1</sub>)*:  $f(x_1^t, x_2^t, \dots, x_n^t) = f^t(x_1, x_2, \dots, x_n)$ . [(P<sub>2</sub>), for example, means that if the  $k$ th individual judges the length of a side of a square to be  $x_k$ , the synthesised judgement on the area of that square will be given by the square of the synthesised judgement on the length of its side.]

Special case (R=P<sub>1</sub>):  $f(1/x_1, 1/x_2, \dots, 1/x_n) = 1/f(x_1, x_2, \dots, x_n)$ . [(R) is of particular importance in ratio judgements. It means that the synthesised

value of the reciprocal of the individual judgements should be the reciprocal of the synthesised value of the original judgements.]

Aczél and Saaty (see Saaty 1990 and Saaty 1994) proved the following theorem:

**Theorem** *The general separable (S) synthesising functions satisfying the unanimity (U) and homogeneity (H) conditions are the geometric mean and the root-mean-power. Moreover, if the reciprocal property (R) is assumed even for a single  $n$ -tuple  $(x_1, x_2, \dots, x_n)$  of the judgements of  $n$  individuals, where not all  $x_k$  are equal, then only the geometric mean satisfies all the above conditions.*

In any rational consensus, those who know more should, accordingly, influence the consensus more strongly than those who are less knowledgeable. Some people are clearly wiser and more sensible in such matters than others, others may be more powerful and their opinions should be given appropriately greater weight. For such unequal importance of voters, not all  $g$ 's in (S) are the same function. In place of (S), the weighted separability property (WS) is now:  $f(x_1, x_2, \dots, x_n) = g_1(x_1)g_2(x_2)\dots g_n(x_n)$ . [(WS) implies that not all judging individuals have the same weight when the judgements are synthesised and the different influences are reflected in the different functions  $(g_1, g_2, \dots, g_n)$ .]

In this situation, Aczél and Alsina (see Saaty 1994) proved the following theorem:

**Theorem** *The general weighted-separable (WS) synthesising functions with the unanimity (U) and homogeneity (H) properties are the weighted geometric mean*

$$f(x_1, x_2, \dots, x_n) = x_1^{q_1} x_2^{q_2} \dots x_n^{q_n} \quad (8.1)$$

and the weighted root-mean-powers

$$f(x_1, x_2, \dots, x_n) = \sqrt[\gamma]{q_1 x_1^\gamma + q_2 x_2^\gamma + \dots + q_n x_n^\gamma} \quad (8.2)$$

where  $q_1 + q_2 + \dots + q_n = 1$ ,  $q_k > 0$  ( $k = 1, 2, \dots, n$ ),  $\gamma > 0$ , but otherwise  $q_1, q_2, \dots, q_n, \gamma$  are arbitrary constants.

If  $f$  also has the reciprocal property (R) and for a single set of entries  $(x_1, x_2, \dots, x_n)$  of judgements of  $n$  individuals, where not all  $x_k$  are equal, then only the weighted geometric mean applies. We give the following theorem which is an explicit statement of the synthesis problem that follows from the

previous results, and applies to the second and third cases of the deterministic approach:

**Theorem** *If  $x_1^{(1)}, \dots, x_n^{(1)}$   $i=1, \dots, m$  are rankings of  $n$  alternatives by  $m$  independent judges and if  $a_i$  is the importance of judge  $i$  developed from a hierarchy for evaluating the judges, and hence*

$$\sum_{i=1}^m a_i = 1, \tag{8.3}$$

then

$$\left( \prod_{i=1}^m x_1^{a_i} \right)^{1/m}, \dots, \left( \prod_{i=1}^m x_n^{a_i} \right)^{1/m} \tag{8.4}$$

are the combined ranks of the alternatives for the  $m$  judges.

The power or priority of judge  $i$  is simply a replication of the judgement of that judge (as if there are as many other judges as indicated by his/her power  $a_i$ ), which implies multiplying his/her ratio by itself  $a_i$  times, and the result follows.

The first requires knowledge how well a particular alternative performs and how well it compares with a standard or benchmark. The second requires comparison with the other alternatives to determine its importance.

## 8.2 On the Construction of Group Choice from Individual Choices

Given a group of individuals, a set of alternatives (with cardinality greater than 2), and individual ordinal preferences for the alternatives, Arrow proved with his Impossibility Theorem that it is impossible to derive a rational group choice (construct a social choice function that aggregates individual preferences) from ordinal preferences of the individuals that satisfy the following four conditions, i.e., at least one of them is violated:

1. *Decisiveness*: the aggregation procedure must generally produce a group order.
2. *Unanimity*: if all individuals prefer alternative A to alternative B, then the aggregation procedure must produce a group order indicating that the group prefers A to B.

3. *Independence of irrelevant alternatives*: given two sets of alternatives which both include A and B, if all individuals prefer A to B in both sets, then the aggregation procedure must produce a group order indicating that the group, given any of the two sets of alternatives, prefers A to B.
4. *No dictator*: no single individual preferences determine the group order.

The main conclusion about group decision making, using the ratio scale approach of the AHP, is that it can be shown that because now individual preferences are cardinal rather than ordinal, it is possible to derive a rational group choice satisfying the above four conditions. It is possible because: (1) individual priority scales can always be derived from a set of pairwise cardinal preference judgements as long as they form at least a minimal spanning tree in the completely connected graph of the elements being compared and (2) the cardinal preference judgements associated with group choice belong to a ratio scale that represents the relative intensity of the group preferences.

## 9. CONCLUSIONS

The seven fundamental properties discussed above provide philosophical, mathematical, and practical bases for the AHP and its application. Of primary importance is the capability of the AHP to transform a multidimensional, multi-scale problem into one that is uni-dimensional over a single scale. This allows decision makers to combine vastly different criteria in a rational, context-preserving, and meaningful way. The use of paired comparisons in judgement matrices is intuitively understandable and is easily done in practice. Although calculating priority vectors from these matrices limits the number of elements that can be compared, this difficulty can be easily remedied by absolute rating. In addition, incommensurate element comparisons can be handled by hierarchical clustering that effectively expands the original 1-9 scale to  $1-\infty$ . Either rank preservation or rank reversal can be accommodated, depending on the desires of the decision maker and the needs of the decision problem. Finally, cardinal ratio scale preferences permit one to include multiple decision makers in the process and to incorporate their individual judgements in a fair manner that also reconciles their specialised knowledge, experience, and authority.

Any formal decision process (e.g., the AHP, MAVT) tries to capture often ill-formed and complex problems using rational frameworks that appeal to our sense of intelligent decision making. Along the way, assumptions and simplification are made (both implicit and explicit) that make formal decision making practical and manageable. The fundamental

properties of the AHP are based on stimulus-response theory, rigorous mathematics, and practical necessities. By doing so, this process mitigates many of the limitations of less "grounded" methods while maintaining broad applicability.

## REFERENCES

- BUEDE, D. AND MAXWELL, D.T. 1995. Rank disagreement: A comparison of multi-criteria methodologies. *Journal of Multi-Criteria Decision Analysis* 4: 1-21.
- LUCE, R. D. AND RAIFFA, H. 1957. *Games and Decisions*. Wiley, New York.
- MILLET, I. AND SAATY, T. L. 1999. On the relativity of relative measures—accommodating both rank preservation and rank reversal in the AHP. *European Journal of Operational Research* 121: 205-212.
- SAATY, T. L. 1990. *Multicriteria Decision Making: The Analytic Hierarchy Process*. RWS Publications, Pittsburgh, PA.
- SAATY, T. L. 1994. *Fundamentals of Decision Making and Priority Theory*. RWS Publications, Pittsburgh, PA.
- SAATY, T. L. 1996. *Decision Making with Dependence and Feedback: The Analytic Network Process*. RWS Publications, Pittsburgh, PA.
- SAATY, T. L. AND HU, G. 1998. Ranking by eigenvector versus other methods in the Analytic Hierarchy Process. *Applied Mathematics Letters* 11: 121-125.
- SAATY, T. L. AND VARGAS, L. G. 1993. Experiments on rank preservation and reversal in relative measurement. *Mathematical and Computer Modeling* 17: 13-18.