RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **102** (2), 2008, pp. 251–318
Estadística e Investigación Operativa / Statistics and Operations Research
Artículo panorámico / Survey

Relative Measurement and Its Generalization in Decision Making Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors The Analytic Hierarchy/Network Process

(To the Memory of my Beloved Friend Professor Sixto Rios Garcia)

Thomas L. Saaty*

Abstract According to the great mathematician Henri Lebesgue, making direct comparisons of objects with regard to a property is a fundamental mathematical process for deriving measurements. Measuring objects by using a known scale first then comparing the measurements works well for properties for which scales of measurement exist. The theme of this paper is that direct comparisons are necessary to establish measurements for intangible properties that have no scales of measurement. In that case the value derived for each element depends on what other elements it is compared with. We show how relative scales can be derived by making pairwise comparisons using numerical judgments from an absolute scale of numbers. Such measurements, when used to represent comparisons can be related and combined to define a cardinal scale of absolute numbers that is stronger than a ratio scale. They are necessary to use when intangible factors need to be added and multiplied among themselves and with tangible factors. To derive and synthesize relative scales systematically, the factors are arranged in a hierarchic or a network structure and measured according to the criteria represented within these structures. The process of making comparisons to derive scales of measurement is illustrated in two types of practical real life decisions, the Iran nuclear show-down with the West in this decade and building a Disney park in Hong Kong in 2005. It is then generalized to the case of making a continuum of comparisons by using Fredholm's equation of the second kind whose solution gives rise to a functional equation. The Fourier transform of the solution of this equation in the complex domain is a sum of Dirac distributions demonstrating that proportionate response to stimuli is a process of firing and synthesis of firings as neurons in the brain do. The Fourier transform of the solution of the equation in the real domain leads to nearly inverse square responses to natural influences. Various generalizations and critiques of the approach are included.

Presentado por / Submitted by Francisco Javier Girón.
Recibido / Received: 6 de agosto de 2008. *Aceptado / Accepted*: 15 de octubre de 2008.
Palabras clave / Keywords: comparisons, conflict resolution, decision, eigenvalue, functional equation, hierarchy, intangibles, judgment, measurement, network, neural firing, sensitivity, synthesis.
Mathematics Subject Classifications: Primary: 62C25, 90B50, 91B20. Secondary: 91A35, 91B08, 91C05

*The author has been awarded with the *2008 Informs Impact Prize* by the Institute for Operations Research and the Management Sciences for his seminal work on the Analytic Hierarchy Process, and for its deployment and extraordinary impact.

Thomas L. Saaty

Medidas relativas y su generalización en la toma de decisiones
Porqué las comparaciones dos a dos son fundamentales en matemáticas para la medida de factores intangibles
El proceso analítico jerarquía/red

Resumen. En primer lugar se argumenta, tal como hizo Henri Lebesgue, que el establecer comparaciones directas entre objetos en relación con alguna propiedad es un proceso matemático fundamental para deducir medidas. Podemos comprobar que esta idea funciona bien para aquellas propiedades para las que se pueden construir escalas de medida. La idea fundamental de este artículo es demostrar que también se puede realizar un proceso de comparación para establecer medidas de propiedades intangibles. Se demuestra que se pueden deducir escalas relativas haciendo comparaciones por parejas, es decir dos a dos, utilizando estimaciones numéricas a partir de otra escala numérica como son los autovectores principales. Estas medidas son del todo necesarias cuando hay que determinar la tasa de intercambio entre los factores intangibles y tangibles. Para deducir y sintetizar de un modo sistemático estas escalas relativas es necesario organizar los factores en una estructura jerárquica o de red. Este proceso se ilustra con dos casos reales, la carrera armamentística nuclear de Irán que le ha enfrentado con el occidente durante la última década y la construcción de un parque Disney en Hong Kong en 2005. A continuación se generaliza el proceso al caso de considerar un continuo de comparaciones usando la ecuación de Fredholm de segunda especie cuya solución da lugar a una ecuación funcional. La transformada de Fourier de la solución de esta ecuación en el plano complejo es una suma de distribuciones de Dirac, lo que demuestra que la respuesta proporcionada por los estímulos es un proceso de activación y síntesis de activaciones del mismo tipo como el que las neuronas lo hacen en el cerebro. La transformada de Fourier de la solución de la ecuación en la recta real conduce a respuestas que aproximadamente son como las recíprocas de los cuadrados a las influencias naturales. Por último, se incluyen varias generalizaciones y críticas al enfoque propuesto.

1 Introduction

We beg the interested, but perhaps impatient reader, to stay with our story to the end to see how applicable the mathematical ideas presented here are to domains of human thinking that have been considered to be forbidding and outside the realm of mathematics.

The subject of this paper, decision making, should be of interest to every living person and most of all to mathematicians because while order and priority have been studied extensively in mathematics they have not been studied in a way to make them applicable in people's lives. Everyone has to make decisions all the time and the complexity of our world of 6.8 billion people requires that more and more we need to consider multiple perspectives in making our decisions. Conflicts of interest will arise and decision making processes help us to resolve them. We cannot escape such considerations in our lives because complexity forces them on us.

Many people including mathematicians whose thinking is grounded in the use of Cartesian axes based on scales of measurement believe that there is only one way to measure things, and it needs a physical measurement scale with a zero and a unit to apply to objects. But, that is not true. Surprisingly, we can also derive accurate and reliable relative scales that do not have a zero or a unit by using our understanding and judgments that are the most fundamental determinants of why we want to measure anything. In reality we do that all the time and we do it subconsciously without thinking about it. Physical scales help with our understanding and use of things that we already know how to measure. After we obtain readings from a physical scale, because they have an arbitrary unit, they still need to be interpreted, by someone experienced, according to what the reading means and thus how adequate or inadequate the object is to satisfy some need we have. But the number of things we don't know how to measure is much larger than the things we know how to measure, and it is highly unlikely that we will ever find ways to measure everything on a physical scale with a unit because, unlike physical things, most of our ideas, feelings, behavior and actions are not fixed once and for all, but change from moment to moment and from one situation to another. Scales of measurement are inventions of a technological mind. Our minds and ways of understanding have always been with us and will always be with us. The mind is an electrical device of neurons whose firings and synthesis give us all the meaning and understanding that we need to survive in a complex world. Can we rely on our minds to be accurate guides with their judgments? The answer depends how well we know the phenomena to which we apply our judgments and how good these judgments are to represent our understanding. In our own personal affairs we are the best judges of what may be good for us. In situations involving many people, we need the judgments from all the participants. In general we think that there are people who are more expert in some areas and their judgments should have precedence over the judgments of those who know less and we often do this.

In mathematics we have two fundamentally different kinds of topology: metric topology and order topology. The first is concerned with how much of a certain attribute an element has as measured on a scale with an arbitrary unit and an origin that is applied uniformly to measure all objects with respect to the given property. The arbitrariness of the unit requires that one must use judgment by an expert to determine the meaning of the numerical outcomes with respect to observables and to compare them with what was known before. Generally one forms, sums and differences of such measurements when they are meaningful. For example, it is meaningful to do arithmetic operations on weights and lengths measured on ratio scales, but not on interval scale readings of temperature on a Fahrenheit or a Celsius scale. Readings that use different metrics with respect to different attributes are combined in physics by using an appropriate formula. Often metric properties belong to the measurement of attributes of the physical world as studied in physics, astronomy, engineering and economics.

The second kind of topology is concerned with measurement of the dominance of one element over others with respect to a common attribute. Order properties belong to the mental world with regard to the importance of its happenings according to human values, preferences and estimation of likelihoods and thus always need judgment before the measurements are made, and not after, as with metric properties. The outcome of such numerical measurements is known as priorities. Unlike physics, all measurements are reduced to priorities.

In this paper we intend to introduce an essentially new paradigm of measurement that has numerous practical implications because it makes it possible for us to deal with intangible factors alongside tangibles used in science and mathematics in a realistic and justifiable way. We need to let the reader know where and how this method of decision making has been used. Among the many applications made by companies and governments, now perhaps numbering in the thousands, the Analytic Hierarchy Process was used by IBM as part of its quality improvement strategy to design its AS/400 computer and win the prestigious Malcolm Baldrige National Quality Award [4]. In (2001) it was used to determine the best site to relocate the earthquake devastated Turkish city Adapazari. British Airways used it in 1998 to choose the entertainment system vendor for its entire fleet of airplanes. A company used it in 1987 to choose the best type of platform to build to drill for oil in the North Atlantic. A platform costs around 3 billion dollars to build, but the demolition cost was an even more significant factor in the decision. The process was applied to the U.S. versus China conflict in the intellectual property rights battle of 1995 over Chinese individuals copying music, video, and software tapes and CD's. An AHP analysis involving three hierarchies for benefits, costs, and risks showed that it was much better for the U.S. not to sanction China. Shortly after the study was complete, the U.S. awarded China the most-favored nation trading status and did not sanction it. Xerox Corporation has used the AHP to allocate close to a billion dollars to its research projects. In 1999, the Ford Motor Company used the AHP to establish priorities for criteria that improve customer satisfaction. Ford gave Expert Choice Inc, an Award for Excellence for helping them achieve greater success with its clients. In 1986 the Institute of Strategic Studies in Pretoria, a government-backed organization, used the AHP to analyze the conflict in South Africa and recommended actions ranging from the release of Nelson Mandela to the removal of apartheid and the granting of full citizenship and equal rights to the black majority. All of these recommended actions were quickly implemented. The AHP has been used in student admissions, military personnel promotions, and hiring decisions. In sports it was used in 1995 to predict which football team would go to the Superbowl and win (correct outcome, Dallas won over my hometown, Pittsburgh).

The AHP was applied in baseball to analyze which of the San Diego Padres players should be retained. There are a number of uses of the ideas by the military that cannot be listed here. Interestingly, it has been used in China dozens of times to determine sights for building dams and other engineering applications. We have not kept up with all the applications but there is an international society on this subject that meets every two years to report on research and applications of the subject. They are referred to as ISAHP (International Symposium on the Analytic Hierarchy Process) meetings. The last one (the ninth) was in Santiago Chile and the tenth will be in Pittsburgh, Pennsylvania.

2 Measurement

In science, measurements of factors with different ratio scales are combined by means of formulas. The formulas apply within structures involving variables and their relations. Each scale has a zero as an origin and an arbitrary unit applied uniformly in all measurements on that scale but as we said before, the meaning of the unit remains elusive and becomes better understood through much practice according to the judgment of experts as to how well it meets understanding and experience or satisfies laws of nature that are always there. Science measures objects objectively, but interprets the significance of the measurements subjectively. Because of the diversity of influences with which decision making is concerned, there are no set laws to characterize in fine detail structures that apply to every decision. Understanding is needed to structure a problem and then make judgments. In decision making the priority scales are derived objectively after subjective judgments are made. The process is the opposite of what we do in science.

Until decision making arrived on the scene as a science late in the 20th century, metric properties were the paradigm that people learned, lived under and were conditioned by in their knowledge of science, engineering, economics, business and operations research. Not much was known about ordering objects on many attributes. People saw the world in terms of different metrics and order was generated whimsically according to metric properties. Still even the many people with a background in engineering and operations research who study decision making try to develop their theories in terms of metric properties. Measurements need human judgments to estimate their importance and from that importance one develops order and measures priorities. The measurement of priorities must stay away from metric properties and be done in terms of order properties. The only possible scale to use that straddles both measurement and order is an absolute scale (invariant under the identity transformation). People have tended to resist the new paradigm and many have tried unsuccessfully to imitate it within the old paradigm of metrics. But order topology is not metric topology and is a subject that stands entirely on its own. Order cannot be logically derived metrically in terms of how close things are. For example minimizing differences does not guarantee a best order as many simple examples show, but the transitivity of order does. Note that decisions with interdependence can lead to transitivity along cycles (like a whirlpool or a cyclone whose tail dominates or is stronger than its head). Such transitivity in theory should not be possible and inevitably leads to inconsistency.

Number is an abstraction often used in three ways: to name things, to count things and to measure things. Naming things with numbers is the least useful indicator of the idea of a number. Counting is a systematic way of tallying things so we know how many things there are. Counting numbers indicate magnitude or quantity. They do not need a unit to define them. The number 5 is the set of all sets of five elements. With counting numbers we can say how many times one set of elements is a multiple of another. The result is an absolute number. That number cannot be changed to some other meaningful number to indicate the multiplicity of that element over the other. So when we make comparisons judgments we use absolute numbers.

We now give a brief exposition about the basic need for using comparisons and especially paired comparisons to derive relative scales of measurement needed to measure intangibles that occur abundantly in many areas of science and in multicriteria decision making. Our purpose is to give some of the highlights of the process of comparisons, its fecundity and what it leads to when discrete judgments are used and then again when generalized to the continuous case where the senses (continuous judgments) are involved.

Being a relative process, comparisons give rise to results that go against the grain of direct measurement. In the latter measurements are applied directly to each object independently of other objects there are. Not so in relative measurement where the numbers derived depend on one another.

3 The Compelling Need to Make Comparisons [11, 12, 13]

Long before measurement scales were invented, people had no direct way to measure because they had no scales and had to compare things with each other or against a standard to determine their relative order. We still have that ability, and it is still critically necessary to be able to make comparisons much of the time, especially when we cannot measure things. One reason may be that we do not have the instrument or scale to do it. Another reason is that we may believe that the outcome of comparisons using our judgment would be calibrated better to our values than using a scale of measurement that was not devised particularly for the use we are putting it to. A third reason may be that there is no way known to measure something: political effectiveness, happiness, aesthetic appeal. Ancient people used their judgment to order things. The way they did it was to compare two things at a time to determine which was the larger or more preferred. By repeating the process they obtained a total ordering of the objects without assigning them numerical values. After being ordered they could rank them: first, second, and so on. But when many criteria are involved it is not so easy to combine the orders obtained with respect to the different criteria to obtain a total order unless there is associated with each partial order a set of numbers that are in some sense commensurate so they can be combined using the numbers (weights or priorities?) associated with the criteria.

How is it that Rene Descartes the genius who invented systems of axes with units for representing functions derived in science and engineering, could not capture meaningfully on his axes the importance of influences like politics, social values, love and hate and the myriad attributes we deal with daily? It is because the measurement of these concepts changes from one situation to another and they depend on a value system which varies from one person to another. Thus their importance cannot be measured once and for all and has to be determined in terms of our values individually. The only way to measure such attributes is by comparing their relative importance with respect to a higher goal. Comparisons are relative and cannot ever be replaced by measurements on unchanging scales concrete scales like meters and kilograms. From the comparisons one derives a scale of priorities which are in relative values. Judgments come first, priorities second.

Our theories and knowledge in science are based on numbers obtained from measurement with respect to properties like length, mass, temperature and time; criteria that we refer to as tangible. But there is a myriad of intangibles we have no scales to measure. Some have attempted to create measurements for them in terms of some common unit like money, but they do not all relate to economic value except by a great and unreliable stretch of the imagination. How are our minds equipped to deal with intangibles and then relate them in a meaningful way to the tangibles that we know how to measure?

The great mathematician Henri Lebesgue, who was concerned with questions of measure theory and measurement, wrote [8]:

> "It would seem that the principle of economy would always require that we evaluate ratios directly and not as ratios of measurements. However, in practice, all lengths are measured in meters, all angles in degrees, etc.; that is we employ auxiliary units and, as it seems, with only the disadvantage of having two measurements to make instead of one. Sometimes, this is because of experimental difficulties or impossibilities that prevent the direct comparison of lengths or angles. But there is also another reason.
>
> In geometrical problems, one needs to compare two lengths, for example, and only those two. It is quite different in practice when one encounters a hundred lengths and may expect to have to compare these lengths two at a time in all possible manners. Thus it is desirable and economical procedure to measure each new length. One single measurement for each length,

> made as precisely as possible, gives the ratio of the length in question to each other length. This explains the fact that in practice, comparisons are never, or almost never, made directly, but through comparisons with a standard scale."

Lebesgue did not go far enough in examining why we have to compare. When we deal with intangible factors, which by definition have no scales of measurement, we can compare them in pairs. Making comparisons is a talent we all have. Not only can we indicate the preferred object, but we can also discriminate among intensities of preference.

For a very long time people believed and vehemently argued that it is impossible to express the intensity of people's feelings with numbers. The epitome of such a belief was expressed by A. F. MacKay who writes [9] that pursuing the cardinal approaches is like chasing what cannot be caught. It was also expressed by Davis and Hersh [6], "If you are more of a human being, you will be aware there are such things as emotions, beliefs, attitudes, dreams, intentions, jealousy, envy, yearning, regret, longing, anger, compassion and many others. These things —the inner world of human life— can never be mathematized." In their book Einstein's space and Van Gogh's sky: Physical reality and beyond, Macmillan, 1982, Lawrence LeShan and Henry Margenau write: "We cannot as we have indicated before, quantify the observables in the domain of consciousness. There are no rules of correspondence possible that would enable us to quantify our feelings. We can make statements of the relative intensity of feelings, but we cannot go beyond this. I can say: I feel angrier at him today than I did yesterday. We cannot, however, make meaningful statements such as, I feel three and one half times angrier than I did yesterday."

The Nobel Laureate, Henri Bergson in "The Intensity of Psychic States". Chapter 1 in,*Time and Free Will: An Essay on the Immediate Data of Consciousness*, translated by F. L. Pogson, M. A. London: George Allen and Unwin (1910): 1–74, writes: But even the opponents of psychophysics do not see any harm in speaking of one sensation as being more intense than another, of one effort as being greater than another, and in thus setting up differences of quantity between purely internal states. Common sense, moreover, has not the slightest hesitation in giving its verdict on this point ; people say they are more or less warm, or more or less sad, and this distinction of more and less, even when it is carried over to the region of subjective facts and unextended objects, surprises nobody.

Much of quantitative thinking about values in economics is based on utility theory, which is grounded in lotteries and wagers, thus implicitly subsuming benefits, opportunities, and costs within a single framework of risks. But lotteries are deeply rooted in the material exchange of money, and it does not make very much sense to equate intangibles such as love and happiness with money. Besides, the real value of money varies among people and is a utility. Utilities are measured on interval scales like temperature. They cannot be added or multiplied and fall quite short of what we need to do to make a decision on objects involving many attributes, and sub-attributes, actors known as stakeholders, people affected by the decision and so on represented richly with hierarchic and network structures. To structure a decision that represents all that is on our minds is at least as important as providing measurements for the important attributes and outcomes.

How can one measure things in a way that captures their influence on one another? We need numbers that do not need a unit and an origin in their definition. What are such numbers? The absolute numbers previously discussed do not need to be defined in terms of a unit or an origin. An absolute number means how many times more. So for a pair let the lesser element be the unit and use an absolute number to indicate how many times more is the larger element.

The numbers derived from making comparisons are specific to each situation that represents the dominance of one element relative to another and cannot be used in general like our standardized scales with a zero and a unit that can be applied in every situation. They are known as priorities. In the next section we give an example of how priorities are derived for tangibles so we can compare the outcome with actual measurements (in terms of relative values). In the next section we introduce and justify the use of numbers to represent judgments. As mentioned above these numbers belong to an absolute scale and can be used in every decision making situation.

4 The Fundamental Scale

When we use judgment to estimate dominance in making comparisons, and in particular when the criterion of the comparisons is an intangible, instead of using two numbers w_i and w_j from a scale (if we must rather than interpreting the significance of their ratio w_i/w_j) we assign a single number drawn from the fundamental 1–9 scale of absolute numbers shown in Table 1 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 and the need for a fundamental scale. This scale is derived from basic principles involving the generalization of comparisons to the continuous case, obtaining a functional equation as a necessary condition and then solving that equation in the real and complex domains (see later).

Table 1. Fundamental Scale of Absolute Numbers

Intensity of Importance	*Definition*	*Explanation*
1	Equal Importance	Two activities contribute equally to the objective
2	Weak or slight	
3	Moderate importance	Experience and judgment slightly favor one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgment strongly favor one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favored very strongly over another; its dominance demonstrated in practice
8	Very,very strong	
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
1.1–1.9	When activities are very close a decimal is added to 1 to show their difference as appropriate	A better alternative way to assigning the small decimals is to compare two close activities with other widely contrasting ones, favoring the larger one a little over the smaller one when using the 1–9 values.
Reciprocals of above	If activity *i* has one of the above nonzero numbers assigned to it when compared with activity *j*, then *j* has the reciprocal value when compared with *i*	A logical assumption
Measurements from ratio scales		When it is desired to use such numbers in physical applications. Alternatively, often one estimates the ratios of such magnitudes by using judgment

We have assumed that an element with weight zero is eliminated from comparison because zero can be applied to the whole universe of factors not included in the discussion. Reciprocals of all scaled ratios that are ≥ 1 are entered in the transpose positions.

The foregoing integer-valued scale of response used in making paired comparison judgments can be derived mathematically from the well-known psychophysical logarithmic response function of Weber-Fechner [7].

In 1846 Weber found, for example, that people while holding in their hand different weights, could distinguish between a weight of $20\,\mathrm{g}$ and a weight of $21\,\mathrm{g}$, but could not if the second weight is only $20.5\,\mathrm{g}$. On the other hand, while they could not distinguish between $40\,\mathrm{g}$ and $41\,\mathrm{g}$, they could between $40\,\mathrm{g}$ and $42\,\mathrm{g}$, and so on at higher levels. We need to increase a stimulus s by a minimum amount Δs to reach a point where our senses can first discriminate between s and $s + \Delta s$. Δs 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 Δs 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 for extending the uses of this law.

In 1860 Fechner [7] considered a sequence of just noticeable increasing stimuli based on Weber's law: For a given value of the stimulus, the magnitude of response remains the same until the value of the stimulus is increased sufficiently large in proportion to the value of the stimulus, thus preserving the proportionality of relative increase in stimulus for it to be detectable for a new response. This suggests the idea of just noticeable differences (jnd), well known in psychology.

To derive the values in the scale starting with a stimulus s_0 successive magnitudes of the new stimuli take the form:

$$\begin{aligned}
s_1 &= s_0 + \Delta s_0 = s_0 + \frac{\Delta s_0}{s_0} s_0 = s_0(1+r) \\
s_2 &= s_1 + \Delta s_1 = s_1(1+r) = s_0(1+r)^2 \equiv s_0 \alpha^2 \\
&\vdots \\
s_n &= s_{n-1}\alpha = s_0 \alpha^n \qquad (n = 0, 1, 2, \ldots)
\end{aligned}$$

We consider the responses to these stimuli to be measured on a ratio scale ($b = 0$). A typical response has the form $M_i = a \log \alpha^i,\ i = 1, \ldots, n$, or one after another they have the form:

$$M_1 = a \log \alpha, \quad M_2 = 2a \log \alpha, \quad \ldots, \quad M_n = na \log \alpha$$

We take the ratios M_i/M_1, $i = 1, \ldots, n$ of these responses in which the first is the smallest and serves as the unit of comparison, thus obtaining the *integer* values $1, 2, \ldots, n$ of the fundamental scale of the AHP. It appears that the positive integers are intrinsic to our ability to make comparisons, and that they were not an accidental invention by our primitive ancestors. In a less mathematical vein, we note that we are able to distinguish ordinally between high, medium and low at one level and for each of them in a second level below that also distinguish between high, medium and low giving us nine different categories. We assign the value one to (low, low) which is the smallest and the value nine to (high, high) which is the highest, thus covering the spectrum of possibilities between two levels, and giving the value nine for the top of the paired comparisons scale as compared with the lowest value on the scale. We shall see in Section 6, we don't need to keep in mind more than 7 ± 2 elements because of increase in "inconsistency" when we compare more than about 7 elements. Finally, we note that the scale just derived is attached to the importance we assign to judgments. If we have an exact measurement such as 2.375 and want to use it as it is for our judgment without attaching significance to it, we can use its entire value without approximation. It is known that small changes in judgment lead to small changes in the derived priorities (Wilkinson, [20]).

In situations where the scale 1–9 is inadequate to cover the spectrum of comparisons needed, that is, the elements compared are inhomogeneous, as for example in comparing a cherry tomato with a watermelon according to size, one uses a process of clustering with a pivot from one cluster to an adjacent cluster that is one order of magnitude larger or smaller than the given cluster, and continues to use the 1–9 scale within each cluster, and in doing that, the scale is extended as far out as desired. What determines the clusters is the relative size of the priorities of the elements in each one. If a priority differs by an order of magnitude or more, it is moved to the appropriate cluster. Hypothetical elements may have to be introduced to make the transition from cluster to cluster a well-designed operation.

Figure 1 shows five geometric areas to which we can apply the paired comparison process in a matrix to test the validity of the procedure. The object is to compare them in pairs by eyeballing them to reproduce their relative weights. The absolute numbers for each pairwise comparison are shown in the matrix in Table 2. Inverses are automatically entered in the transpose position. We can approximate the priorities from this matrix by normalizing each column and then taking the average of the corresponding entries in the columns. To do it precisely we show in Section 5 that one needs to compute the principal eigenvector as we did in this case to deal with "inconsistency" in the judgments. Table 2 also gives the actual measurements in relative form on the right. An element on the left of the matrix is compared with another element at the top for its dominance with respect to its size. If it is not the reciprocal value is used for the dominance of the other element over it.

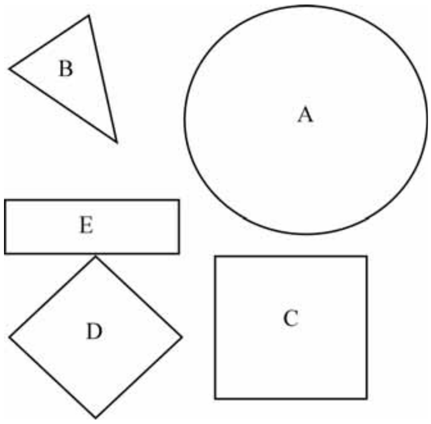

Figure 1. Five figures drawn with appropriate size of area

Table 2. Judgments, outcomes, and actual relative sizes of the five geometric shapes

Figure	Circle	Triangle	Square	Diamond	Rectangle	Priorities (Eigen Vector)	Actual Relative Size
Circle	1	9	2	3	5	0.462	0.471
Triangle	1/9	1	1/5	1/3	1/2	0.049	0.050
Square	1/2	5	1	3/2	3	0.245	0.234
Diamond	1/3	3	2/3	1	3/2	0.151	0.149
Rectangle	1/5	2	1/3	2/3	1	0.093	0.096

Note the closeness of the last two columns in Table 2, the priorities derived from judgment and the actual measurements that were then normalized to give the relative sizes. By including more than two alternatives in a decision problem, one is able to obtain better values for the derived scale because of redundancy in the comparisons, which helps improve the overall accuracy of the judgments.

Unlike the old way of assigning a number from a fixed scale with an arbitrary unit once and for all to each thing or object, in the new paradigm of measurement the measurements are not fixed but depend on each other and on the context of the problem and its objectives. While things may or may not depend on each other according to their function, they are always interdependent when the measurement is relative.

It has been observed that in general comparisons with respect to the dominance of one object over another with respect to a certain attribute or criterion take three forms: importance or significance that includes all kinds of influence —physical measurements in science, engineering and economics fall in this category, preference as in making decisions, and likelihood as in probabilities. If there is adequate knowledge, one can compare anything with anything else that shares a common attribute or criterion. Thus comparisons go beyond ordinary measurement to include intangibles for which there are no scales of measurement.

A more abstract form of comparisons would involve elements with tangible properties that one must think about but cannot be perceived through the senses. See the judgments in Table 3 for estimating the relative amount of protein in seven food items. Reciprocal values are used in the transpose positions of the matrix below the diagonal. The answer derived from the matrix and the exact answer from the literature is shown at the bottom of the table.

Table 3. Relative Amount of Protein in Seven Foods
In this example we see that the apple was wrongly chosen with respect to having protein.

Estimating which Food has more Protein

Food Consumption in the U. S.	**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	1	1/3	1	1/5
G: Fish	1	4	9	6	3	5	1

The resulting derived scale and the actual values are shown below

	Steak	Potatoes	Apples	Soybean	W. Bread	T. Cake	Fish
Derived	0.345	0.031	0.030	0.065	0.124	0.078	0.328
Actual	0.370	0.040	0.000	0.070	0.110	0.090	0.320

(Derived scale has a consistency ratio of 0.028)

Thus it can happen that a person mistakenly adopts items for comparison like the apples in the table which upon better knowledge end up being eliminated.

Note in Tables 2 and 3 that while in the first example the eye perceives different size areas, in the second example the mind, through wide experience, derives a feeling for the amount of protein different foods contain. Feelings are ususally distinguished qualitatively and associated with numerical values.

4.1 Probability Example

Here is an interesting example illustrating the fact that priorities derived from comparisons give the correct results in the field of probability. Consider an urn with balls of three colors: two black, one white and three red. The probabilities of drawing a ball of one of these colors are, respectively, $2/6,\ 1/6, 3/6$.

The pairwise comparison matrix in terms of the ratios of the number of balls having the indicated colors is shown in Table 4.

In an AHP pairwise comparison, the smaller element is considered to be the unit, or one, while the larger element is expressed as a multiple of that unit. That is, every judgment contains a 1; it is in the denominator of $(w_i/w_j)/1$, if the row element is dominant, or in the numerator if the column element is

Table 4. Original ratios in terms of color of balls

	Black	White	Red
Black	1	2	$2/3$
White	$1/2$	1	$1/3$
Red	$3/2$	3	1

Table 5. Ratios as pairwise comparisons

	Black	White	Red
Black	1	$2/1$	$1/1.5$
White	$1/2$	1	$1/3$
Red	$1.5/1$	$3/1$	1

dominant. The original ratios in Table 4 are shown in Table 5 after being translated into the form usually used for judgments.

Since this matrix is consistent (see Section 6), the entries of any column, after normalizing, give the ratio of balls of each color to the total. The eigenvalue solution described below (Section 5) will of course give the same result. We have, for example, from the first column $1/1.5$, $1/3$, 1 or, changing to a common denominator $2/6$, $1/6$, $3/6$, which, as is to be expected, corresponds to the probabilities that a ball drawn at random has one of the three colors. If the urn has a large number of balls whose relative quantities can be compared, this approach would give an estimate of the relative magnitudes of each color of ball.

The three examples show that an experienced person can provide informed numerical judgments from which relatively good estimates are derived. The same process can be applied to intangibles. The Analytic Hierarchy Process (AHP) and the Analytic Network Process (ANP) are based on the comparisons of both tangibles and intangibles. Before we proceed to deal with intangibles we give a sketch of the theoretical concepts underlying the way scales are derived from paired comparisons.

5 Deriving Scales from Paired Comparisons

Consider n stocks, $A_1, \ldots, A_n$, with known worth $w_1, \ldots, w_n$, respectively, and suppose that a matrix of pairwise ratios is formed whose rows give the ratios of the worth of each stock with respect to all others as follows:

$$\begin{array}{c} \\ A_1 \\ \vdots \\ A_n \end{array} \begin{array}{c} \begin{array}{ccc} A_1 & \ldots & A_n \end{array} \\ \begin{bmatrix} \frac{w_1}{w_1} & \cdots & \frac{w_1}{w_n} \\ \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \end{array}$$

We can recover the scale w using the following equation:

$$\begin{bmatrix} \frac{w_1}{w_1} & \cdots & \frac{w_1}{w_n} \\ \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = n \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

or briefly $Aw = nw$.

Thus, to recover the scale w from the matrix of ratios A, one must solve the eigenvalue problem $Aw = nw$ or $(A - nI)w = 0$. This is a system of homogeneous linear equations. It has a nontrivial solution if and only if the determinant of $A - nI$ vanishes, that is, n is an eigenvalue of A. Now 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, the sum of its diagonal elements, and in this case the trace of A is equal to n. Thus n is the principal eigenvalue of A, and one has a nonzero solution w. The solution consists of positive entries and is unique to within a multiplicative constant. To make w unique, one can normalize its entries by dividing by their sum. Note that if one divides two readings from a ratio scale one obtains an absolute number. Normalization transforms a ratio scale to an absolute scale. Thus, given the comparison matrix, one can recover the original scale in relative terms. In this case, the solution is any column of A normalized. The matrix A is not only reciprocal, but also consistent. Its entries satisfy the condition $a_{jk} = a_{ik}/a_{ij}$. It follows that the entire matrix can be constructed from a set of n elements that form a spanning tree across the rows and columns. We have shown that if values from a standard scale are used to make the comparisons, the principal eigenvector recovers these values in normalized form.

In the general case when only judgment but not the numbers themselves are available, the precise value of w_i/w_j which is a dimensionless number and thus belongs to an absolute scale that is invariant under the identity transformation, is not known, but instead only an estimate of it can be given as a numerical judgment. For the moment, consider an estimate of these values by an expert who is assumed to make small perturbations of the ratio w_i/w_j. This implies small perturbations of the eigenvalues. The problem now becomes $A'w' = \lambda_{\max}w'$ where $\lambda_{\max}$ is the largest eigenvalue of A'. To determine how good the derived estimate w is we note that if w is obtained by solving $A'w' = \lambda_{\max}w'$, the matrix A whose entries are w'_i/w'_j is a consistent matrix. It is a consistent estimate of the matrix A' which need not be consistent. In fact, the entries of A' need not even be transitive; that is, A_1 may be preferred to A_2 and A_2 to A_3 but A_3 may be preferred to A_1. What we would like is a measure of the error due to inconsistency. It turns out that A' is consistent if and only if $\lambda_{\max} = n$ and that we always have $\lambda_{\max} \geq n$. The existence of the vector w' with positive components and its uniqueness to within a multiplicative constant in the inconsistent case, is proven by using the concept of dominance and the limiting matrix of powers of A'. Thus w' belongs to a ratio scale or when normalized, to an absolute scale that needs no origin or arbitrary unit because dividing two numbers from the same ratio scale yields a number that is belongs to an absolute scale.

By way of further elaboration on the need for the eigenvector to derive the priority scale of values from a matrix of comparisons we note that a major property of a consistent matrix is that it satisfies the condition $A^k = n^{k-1}A$ where n is the order of A, so all powers of A are essentially equal to A. Now dominance of an inconsistent matrix no longer satisfies this condition and one must consider priorities derived from direct dominance as in the matrix itself, second order dominance obtained from the square of the matrix and so on. The total dominance of each element is obtained as the normalized sum of its rows. The result is an infinite number of priority vectors each representing a different order of dominance. The Cesaro sum of these vectors is equal to the priority vector obtained from the limiting powers of the matrix. That limit matrix is again consistent. By raising the matrix to powers, the principal eigenvector has captured transitivity of all order. Here is some detail.

Let a_{ij} be the relative dominance of A_i over A_j. To simplify the notation, let the matrix corresponding to the reciprocal pairwise relation be denoted by $A = (a_{ij})$. The relative dominance of A_i over A_j along paths of length k is given by

$$\frac{\sum_{j=1}^{n} a_{ij}^{(k)}}{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{(k)}}$$

where $a_{ij}^{(k)}$ is the (i, j) entry of the kth power of the matrix (a_{ij}). The total dominance $w(A_i)$, of alternative

i over all other alternatives along paths of all lengths is given by the infinite series

$$w(A_i) = \sum_{k=1}^{\infty} \frac{\sum_{j=1}^{n} a_{ij}^{(k)}}{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^{(k)}}$$

which coincides with the Cesaro sum

$$\lim_{M \to \infty} \frac{1}{M} \sum_{k=1}^{M} \frac{\sum_{j=1}^{n} a_{ij}^{(k)}}{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^{(k)}}$$

6 When is a Positive Reciprocal Matrix Consistent?

In light of the foregoing, for the validity of the vector of priorities to describe response, we need greater redundancy and therefore also a large number of comparisons. We now show that for consistency we need to make a small number of comparisons. So where is the optimum number?

We now relate the psychological idea of the consistency of judgments and its measurement, to a central concept in matrix theory and also to the size of our channel capacity to process information. It is the principal eigenvalue of a matrix of paired comparisons.

Let $A = [a_{ij}]$ be an n-by-n positive reciprocal matrix, so all $a_{ii} = 1$ and $a_{ij} = 1/a_{ji}$ for all $i, j = 1, \ldots,$ n. Let $w = [w_i]$ be the principal right eigenvector of A, let $D = \mathrm{diag}(w_1, \ldots, w_n)$ be the n-by-n diagonal matrix whose main diagonal entries are the entries of w, and set $E \equiv D^{-1}AD = [a_{ij}w_j/w_i] = [\varepsilon_{ij}]$. Then E is similar to A and is a positive reciprocal matrix since $\varepsilon_{ij} = a_{ji}w_i/w_j = (a_{ij}w_j/w_i)^{-1} = 1/\varepsilon_{ij}$. Moreover, all the row sums of E are equal to the principal eigenvalue of A:

$$\sum_{j=1}^{n} \varepsilon_{ij} = \sum_{j=1}^{n} a_{ij} \frac{w_j}{w_i} = \frac{[Aw]_i}{w_i} = \lambda_{\max} \frac{w_i}{w_i} = \lambda_{\max}$$

The computation

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

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 $\varepsilon_{ij} = 1$, 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. We can say that as $\mu \to 0$, $a_{ij} \to w_i/w_j$, or $\varepsilon_{ij} = a_{ij}w_j/w_i \to 1$. These two desirable properties explain the term "n" in the numerator of μ; what about the term "$n - 1$" in the denominator? Since $\mathrm{trace}(A) = n$ is the sum of all the eigenvalues of A, if we denote the eigenvalues of A that are different from $\lambda_{\max}$ by $\lambda_2, \ldots, \lambda_n$, we see that $n = \lambda_{\max} + \sum_{i=2}^{n} \lambda_i$, so $n - \lambda_{\max} = \sum_{i=2}^{n} \lambda_i$ and $\mu = \frac{1}{n-1} \sum_{i=2}^{n} \lambda_i$ is the average of the non-principal eigenvalues of A.

It is an easy, but instructive, computation to show that $\lambda_{\max} = 2$ for every 2-by-2 positive reciprocal matrix:

$$\begin{bmatrix} 1 & \alpha \\ \alpha^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 + \alpha \\ (1 + \alpha)\alpha^{-1} \end{bmatrix} = 2 \begin{bmatrix} 1 + \alpha \\ (1 + \alpha)\alpha^{-1} \end{bmatrix}$$

Thus, every 2-by-2 positive reciprocal matrix is consistent.

Not every 3-by-3 positive reciprocal matrix is consistent, but in this case we are fortunate to have again explicit formulas for the principal eigenvalue and eigenvector. For

$$A = \begin{bmatrix} 1 & a & b \\ 1/a & 1 & c \\ 1/b & 1/c & 1 \end{bmatrix}$$

we have $\lambda_{\max} = 1 + d + d^{-1},\ d = (ac/b)^{1/3}$ and

$$w_1 = \frac{bd}{1 + bd + c/d}, \qquad w_2 = \frac{c}{d(1 + bd + c/d)}, \qquad w_3 = \frac{1}{1 + bd + c/d}$$

Note that $\lambda_{\max} = 3$ when $d = 1$ or $c = b/a$, which is true if and only if A is consistent.

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 $\{1/9, 1/8, \ldots, 1, 2, \ldots, 8, 9\}$. 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 6 shows the values obtained from one set of such simulations and also their first order differences, for matrices of size 1, 2, …, 15.

Table 6. Random index

Order	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
R. I.	0	0	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	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

Figure 2 below is a plot of the first two rows of Table 6. It shows the asymptotic nature of random inconsistency.

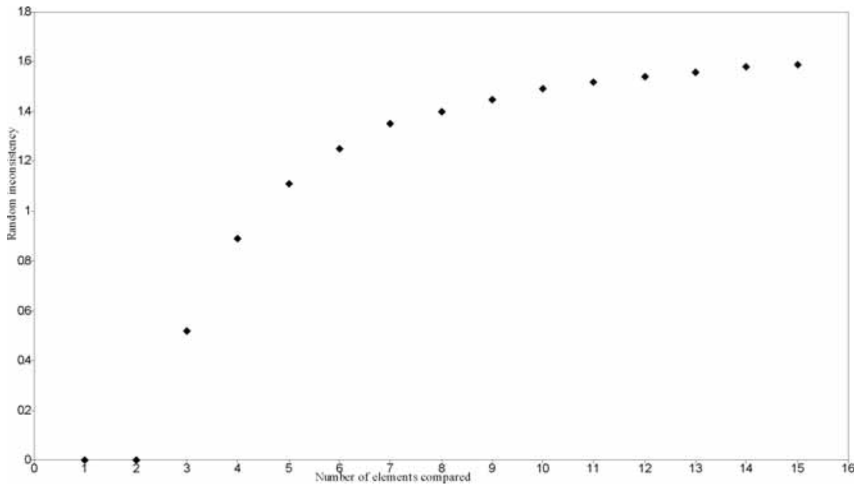

Figure 2. Plot of Random Inconsistency

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 6. The *consistency ratio* (C. R.) of a pairwise comparison matrix is the ratio of its consistency index μ to the corresponding random index value in Table 5. 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 wants 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 0.10. The requirement of 10% cannot be made smaller such as 1% or 0.1% 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 $\varepsilon_{ij} = a_{ij}w_j/w_i$ 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. Three methods are plausible for this purpose. All require theoretical investigation of convergence and efficiency. The first uses an explicit formula for the partial derivatives of the Perron eigenvalue with respect to the matrix entries.

For a given positive reciprocal matrix $A = [a_{ij}]$ and a given pair of distinct indices $k > l$, define $A(t) = [a_{ij}(t)]$ by $a_{kl}(t) \equiv a_{kl} + t$, $a_{lk}(t) \equiv (a_{lk} + t)^{-1}$, and $a_{ij}(t) \equiv a_{ij}$ for all $i \neq k$, $j \neq l$, so $A(0) = A$. Let $\lambda_{\max}(t)$ denote the Perron eigenvalue of $A(t)$ for all t in a neighborhood of $t = 0$ that is small enough to ensure that all entries of the reciprocal matrix $A(t)$ are positive there. Finally, let $v = [v_i]$ be the unique positive eigenvector of the positive matrix A^T that is normalized so that $v^T w = 1$. Then a classical perturbation formulatells us that

$$\left.\frac{\mathrm{d}\lambda_{\max}(t)}{\mathrm{d}t}\right|_{t=0} = \frac{v^T A'(0) w}{v^T w} = v^T A'(0) w = v_k w_l - \frac{1}{a_{kl}^2} v_l w_k.$$

We conclude that

$$\frac{\partial \lambda_{\max}}{\partial a_{ij}} = v_i w_j - a_{ji}^2 v_j w_i, \qquad \text{for all } i, j = 1, \ldots, n.$$

Because we are operating within the set of positive reciprocal matrices, $\frac{\partial \lambda_{\max}}{\partial a_{ji}} = -\frac{\partial \lambda_{\max}}{\partial a_{ij}}$ for all i and j. Thus, to identify an entry of A whose adjustment within the class of reciprocal matrices would result in the largest rate of change in $\lambda_{\max}$ we should examine the $n(n-1)/2$ values $\{v_i w_j - a_{ji}^2 v_j w_i\}$, $i > j$ and select (any) one of largest absolute value. If the judge is unwilling to change that judgment at all, we try with the second most inconsistent judgment and so on. If no judgment is changed the decision is postponed until a 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 judge's knowledge does not permit one to improve his or her consistency and more information is required to improve the consistency of judgments.

Before proceeding further, the following observations may be useful for a better understanding of the importance of the concept of a limit on our ability to process information and also change information.

The quality of response to stimuli is determined by three factors. Accuracy or validity, consistency, and efficiency or amount of information generated. Our judgment is much more sensitive and responsive to large perturbations. 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 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 6 gives the differences between successive numbers in the second row. Figure 3 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. Thus in general, one should only compare a few elements (about seven), and when their number is larger, one should put them into groups with a common element from one group to the next so that its weight can be used as a pivot to combine the final weights.

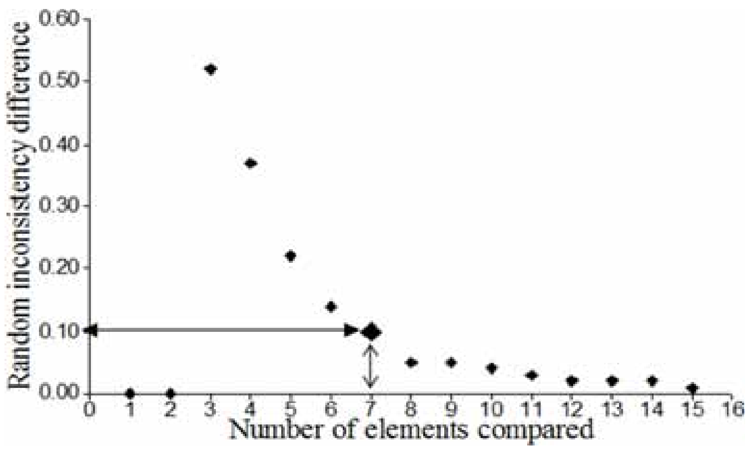

Figure 3. Plot of First Differences in Random Inconsistency

7 Second Affirmation through the Eigenvector

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 eigenvector w_1 due to a perturbation ΔA in the matrix A where A is inconsistent is given by Wilkinson [20]:

$$\Delta w_1 = \sum_{j=2}^{n} (v_j^T \Delta A w_1 / (\lambda_1 - \lambda_j) v_j^T w_j) w_j$$

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 $|v_i^T w_i| < 1$.

8 Axioms

For completeness we have borrowed the axioms section from Chapter 10 of my book *Fundamentals of Decision Making* [12]. The reader might consult that book for a full description and results derived from the axioms, but the editors thought it would be useful to include them here for ready reference.

Let $\mathfrak{A}$ be a finite set of n elements called alternatives. Let $\mathfrak{C}$ be a set of properties or attributes with respect to which elements in $\mathfrak{A}$ are compared. A *property* is a feature that an object or individual possesses even if we are ignorant of this fact, whereas an *attribute* is a feature we assign to some object: it is a concept. Here we assume that properties and attributes are interchangeable, and we generally refer to them as criteria. A *criterion* is a primitive concept.

When two objects or elements in $\mathfrak{A}$ are compared according to a criterion C in $\mathfrak{C}$, we say that we are performing binary comparisons. Let $>_C$ be a binary relation on $\mathfrak{A}$ representing "more preferred than" or "dominates" with respect to a criterion C in $\mathfrak{C}$. Let $\sim_C$ be the binary relation "indifferent to" with respect to a criterion C in $\mathfrak{C}$. Hence, given two elements, $A_i, A_j \in \mathfrak{A}$, either $A_i >_C A_j$ or $A_j >_C A_i$ or $A_i \sim_C A_j$ for all $C \in \mathfrak{C}$. We use $A_i >\sim_C A_j$ to indicate more preferred or indifferent. A given *family of binary relations* $>_C$ with respect to a criterion C in $\mathfrak{C}$ is a primitive concept. We shall use this relation to derive the notion of priority or importance both with respect to one criterion and also with respect to several.

Let $\mathfrak{P}$ be the set of mappings from $\mathfrak{A}\times\mathfrak{A}$ to $\mathbb{R}^+$ (the set of positive reals). Let $f : \mathfrak{C} \to \mathfrak{P}$. Let $P_C \in f(C)$ for $C \in \mathfrak{C}$. P_C assigns a positive real number to every pair $(A_i, A_j) \in \mathfrak{A} \times \mathfrak{A}$. Let $P_C(A_i, A_j)/a_{ij} \in \mathbb{R}^+$, $A_i, A_j \in \mathfrak{A}$. For each $C \in \mathfrak{C}$, the triple $(\mathfrak{A}\times\mathfrak{A}, \mathbb{R}^+, P_C)$ is a *fundamental* or *primitive scale*. A fundamental scale is a mapping of objects to a numerical system.

Definition 1. *For all* A_i, $A_j \in \mathfrak{A}$ *and* $C \in \mathfrak{C}$

$$A_i >_C A_j \qquad \textit{if and only if } P_C(A_i, A_j) > 1,$$

$$A_i \sim_C A_j \qquad \textit{if and only if } P_C(A_i, A_j) = 1.$$

If $A_i >_C A_j$, we say that A_i dominates A_j with respect to $C \in \mathfrak{C}$. Thus P_C represents the intensity or strength of preference for one alternative over another.

8.1 Reciprocal Axiom

Axiom 1. *For all* A_i, $A_j \in \mathfrak{A}$ *and* $C \in \mathfrak{C}$

$$P_C(A_i, A_j) = \frac{1}{P_C(A_j, A_i)}$$

Whenever we make paired comparisons, we need to consider both members of the pair to judge the relative value. The smaller or lesser one is first identified and used as the unit for the criterion in question. The other is then estimated as a not necessarily integer multiple of that unit. Thus, for example, if one stone is judged to be five times heavier than another, then the other is automatically one fifth as heavy as the first because it participated in making the first judgment. The comparison matrices that we consider are formed by making paired reciprocal comparisons. It is this simple yet powerful means of resolving multicriteria problems that is the basis of the AHP.

Let $A = (a_{ij})/(P_C(A_i, A_j))$ be the set of paired comparisons of the alternatives with respect to a criterion $C \in \mathfrak{C}$. By Axiom 1, A is a positive reciprocal matrix. The object is to obtain a *scale of relative dominance* (or *rank order*) of the alternatives from the paired comparisons given in A.

There is a natural way to derive the relative dominance of a set of alternatives from a pairwise comparison matrix A.

Definition 2. *Let $R_{M(n)}$ be the set of $(n \times n)$ positive reciprocal matrices $A = (a_{ij})/(P_C(A_i, A_j))$ for all $C \in \mathfrak{C}$. Let $[0, 1]^n$ be the n-fold cartesian product of $[0, 1]$ and let $\Psi(A)\colon R_{M(n)} \to [0, 1]^n$ for $A \in R_{M(n)}$, $\Psi(A)$ is an n-dimensional vector whose components lie in the interval $[0, 1]$. The triple $(R_{M(n)}, [0, 1]^n, \Psi)$ is a* derived scale. *A derived scale is a mapping between two numerical relational systems.*

It is important to point out that the rank order implied by the derived scale Ψ may not coincide with the order represented by the pairwise comparisons. Let $\Psi_i(A)$ be the ith component of $\Psi(A)$. It denotes the relative dominance of the ith alternative. By definition, for A_i, $A_j \in \mathfrak{A}$, $A_i >_C A_j$ implies $P_C(A_i, A_j) > 1$. However, if $P_C(A_i, A_j) > 1$, the derived scale could imply that $\Psi_j(A) > \Psi_i(A)$. This occurs if row dominance does not hold, i.e., for A_i, $A_j \in \mathfrak{A}$, and $C \in \mathfrak{C}$, $P_C(A_i, A_j) \geq P_C(A_j, A_k)$ does not hold for all $A_k \in \mathfrak{A}$. In other words, it may happen that $P_C(A_i, A_j) > 1$, and for some $A_k \in \mathfrak{A}$ we have

$$P_C(A_i, A_k) < P_C(A_j, A_k)$$

A more restrictive condition is the following:

Definition 3. *The mapping P_C is said to be* consistent *if and only if $P_C(A_i, A_j)P_C(A_j, A_k) = P_C(A_i, A_k)$ for all i, j, and k. Similarly the matrix A is consistent if and only if $a_{ij}a_{jk} = a_{ik}$ for all i, j, and k.*

If P_C is consistent, then Axiom 1 automatically follows and the rank order induced by Ψ coincides with pairwise comparisons.

Luis Vargas has proposed, through personal communication with the author, that the following "behavioral" independence axiom could be used instead of the more mathematical reciprocal axiom that would then follow as a theorem. However, the reciprocal relation does not imply independence as defined by him, and unless one wishes to assume independence, one should retain the reciprocal axiom.

Two alternatives A_i and A_j are said to be mutually independent with respect to a criterion $C \in \mathfrak{C}$ if and only if, for any A_k the paired comparison of the component $\{A_i, A_j\}$ with respect to A_k satisfies

$$P_C[\{A_i, A_j\}, A_k] = P_C(A_i, A_k)P_C(A_j, A_k)$$

and

$$P_C[A_k, \{A_i, A_j\}] = P_C(A_k, A_i)P_C(A_k, A_j)$$

A set of alternatives is said to be independent if they are mutually independent.

Axiom 1. *All the alternatives in $\mathfrak{A}$ are independent.*

8.2 Hierarchic Axioms

Definition 4. *A* partially ordered set *is a set S with a binary relation $\leq$ which satisfies the following conditions:*

a. Reflexive*: For all $x \in S$, $x \leq x$,*

b. Transitive*: For all x, y, $z \in S$, if $x \leq y$ and $y \leq z$ then $x \leq z$,*

c. Antisymmetric*: For all x, $y \in S$, if $x \leq y$ and $y \leq x$ then $x = y$ (x and y coincide).*

Definition 5. *For any relation $x \leq y$ (read, y includes x) we define $x < y$ to mean that $x \leq y$ and $x \neq y$. y is said to* cover *(*dominate*) x if $x < y$ and if $x < t < y$ is possible for no t.*

Partially ordered sets with a finite number of elements can be conveniently represented by a directed graph. Each element of the set is represented by a vertex so that an arc is directed from y to x if $x < y$.

Definition 6. *A subset E of a partially ordered set S is said to be* bounded *from above (below) if there is an element $s \in S$ such that $x \leq s$ ($\geq s$) for every $x \in E$. The element s is called an upper (lower) bound of E. We say that E has a supremum (infimum) if it has upper (lower) bounds and if the set of upper (lower) bounds U (L) has an element u_1 (l_1) such that $u_1 \leq u$ for all $u \in U$ ($l_1 \geq l$ for all $l \in L$).*

Definition 7. *Let $\mathfrak{H}$ be a finite partially ordered set with largest element b. The set $\mathfrak{H}$ is a* hierarchy *if it satisfies the following conditions:*

1. *There is a partition of $\mathfrak{H}$ into sets called levels $\{L_k,\ k = 1, 2, \ldots, h\}$, where $L_1 = \{b\}$.*

2. *$x \in L_k$ implies $x^- \subseteq L_{k+1}$, where $x^- = \{y \mid x \text{ covers } y\}$, $k = 1, 2, \ldots, h-1$.*

3. *$x \in L_k$ implies $x^+ \subseteq L_{k-1}$, where $x^+ = \{y \mid y \text{ covers } x\}$, $k = 2, 3, \ldots, h$.*

Definition 8. *Given a positive real number $\rho \geq 1$, a nonempty set $x^- \subseteq L_{k+1}$ is said to be ρ-homogeneous with respect to $x \in L_k$ if for every pair of elements y_1, $y_2 \in x^-$, $1/\rho \leq P_C(y_1, y_2) \leq \rho$. In particular the reciprocal axiom implies that $P_C(y_i, y_i) = 1$.*

Axiom 2. *Given a hierarchy $\mathfrak{H}$, $x \in \mathfrak{H}$ and $x \in L_k$, $x^- \subseteq L_{k+1}$ is ρ-homogeneous for $k = 1, \ldots, h-1$.*

Homogeneity is essential for comparing similar things, as the mind tends to make large errors in comparing widely disparate elements. For example, we cannot compare a grain of sand with an orange according to size. When the disparity is great, the elements are placed in separate components of comparable size, giving rise to the idea of levels and their decomposition. This axiom is closely related to the well-known Archimedean property which says that forming two real numbers x and y with $x < y$, there is an integer n such that $nx \geq y$, or $n \geq y/x$.

The notions of fundamental and derived scales can be extended to $x \in L_k$, $x^- \subseteq L_{k+1}$ replacing $\mathfrak{C}$ and $\mathfrak{A}$, respectively. The derived scale resulting from comparing the elements in x^- with respect to x is called a *local derived scale* or *local priorities*. Here no irrelevant alternative is included in the comparisons, and such alternatives are assumed to receive the value of zero in the derived scale.

Given L_k, $L_{k+1} \subseteq \mathfrak{H}$, let us denote the local derived scale for $y \in x^-$ and $x \in L_k$ by $\Psi_{k+1}(y|x)$, $k = 2, 3, \ldots, h-1$. Without loss of generality we may assume that $\sum_{y0x^-} \Psi_{k+1}(y|x) = 1$. Consider the matrix $\Psi_k(L_k|L_{k-1})$ whose columns are local derived scales of elements in L_k with respect to elements in L_{k-1}.

Definition 9. *A set $\mathfrak{A}$ is said to be* outer dependent *on a set $\mathfrak{C}$ if a fundamental scale can be defined on $\mathfrak{A}$ with respect to every $c \in \mathfrak{C}$.*

Decomposition implies containment of the small elements by the large components or levels. In turn, this means that the smaller elements depend on the outer parent elements to which they belong, which themselves fall in a large component of the hierarchy. The process of relating elements (e.g., alternatives) in one level of the hierarchy according to the elements of the next higher level (e.g., criteria) expresses the outer dependence of the lower elements on the higher elements. This way comparisons can be made between them. The steps are repeated upward in the hierarchy through each pair of adjacent levels to the top element, the focus or goal.

The elements in a level may depend on one another with respect to a property in another level. Input-output dependence of industries (e.g., manufacturing) demonstrates the idea of inner dependence. This may be formalized as follows:

Definition 10. *Let $\mathfrak{A}$ be outer dependent on $\mathfrak{C}$. The elements in $\mathfrak{A}$ are said to be* inner dependent *with respect to $C \in \mathfrak{C}$ if for some $A \in \mathfrak{A}$, $\mathfrak{A}$ is outer dependent on A.*

Axiom 3. *Let $\mathfrak{H}$ be a hierarchy with levels $L_1, L_2, \ldots, L_h$. For each L_k, $k = 1, 2, \ldots, h-1$,*

1. *L_{k+1} is outer dependent on L_k*
2. *L_k is not outer dependent on L_{k+1}*
3. *L_{k+1} is not inner dependent with respect to any $x \in L_k$.*

8.3 Principle of Hierarchic Composition

If Axiom 3 holds, the global derived scale (rank order) of any element in $\mathfrak{H}$ is obtained from its component in the corresponding vector of the following:

$$\begin{aligned} \Psi_1(b) &= 1 \\ \Psi_2(L_2) &= \Psi_2(b^-|b) \\ \vdots \quad & \quad \vdots \\ \Psi_k(L_k) &= \Psi_k(L_k|L_{k-1}), \quad \Psi_{k-1}(L_{k-1}), \qquad k = 3, \ldots, h \end{aligned}$$

Were one to omit Axiom 3, the Principle of Hierarchic Composition would no longer apply because of outer and inner dependence among levels or components which need not form a hierarchy. The appropriate composition principle is derived from the supermatrix approach explained below, of which the Principle of Hierarchic Composition is a special case.

A hierarchy is a special case of a system, defined as follows:

Definition 11. *Let $\mathfrak{H}$ be a family of nonempty sets $\mathfrak{C}_1, \mathfrak{C}_2, \ldots, \mathfrak{C}_n$ where $\mathfrak{C}_i$ consists of the elements $\{e_{ij}, j = 1, \ldots, m_i\}$, $i = 1, 2, \ldots, n$. $\mathfrak{H}$ is a system if it is a directed graph whose vertices are $\mathfrak{C}_i$ and whose arcs are defined through the concept of outer dependence; thus, given two components $\mathfrak{C}_i$ and $\mathfrak{C}_j \in \mathfrak{H}$, there is an arc from $\mathfrak{C}_i$ to $\mathfrak{C}_j$ if $\mathfrak{C}_j$ is outer dependent on $\mathfrak{C}_i$.*

Axiom 3'. *Let $\mathfrak{H}$ be a system consisting of the subsets $\mathfrak{C}_1, \mathfrak{C}_2, \ldots, \mathfrak{C}_n$. For each $\mathfrak{C}_i$ there is some $\mathfrak{C}_j$ so that either $\mathfrak{C}_i$ is outer dependent on $\mathfrak{C}_j$ or $\mathfrak{C}_j$ is outer dependent on $\mathfrak{C}_i$, or both.*

Note that $\mathfrak{C}_i$ may be outer dependent on $\mathfrak{C}_i$, which is equivalent to inner dependence in a hierarchy. Actually Axiom 3' would by itself be adequate without Axiom 3. We have separated them because of the importance of hierarchic structures, which are more widespread at the time of this writing than are systems with feedback.

Many of the concepts derived for hierarchies also relate to systems with feedback. Here one needs to characterize dependence among the elements. We now give a criterion for this purpose.

Let $D_A \subseteq \mathfrak{A}$ be the set of elements of $\mathfrak{A}$ outer dependent on $A \in \mathfrak{A}$. Let $\Psi_{A_i,C}(A_j),\ A_j \in \mathfrak{A}$ be the derived scale of the elements of $\mathfrak{A}$ with respect to $A_i \in \mathfrak{A}$, for a criterion $C \in \mathfrak{C}$. Let $\Psi_C(A_j)$, $A_j \in \mathfrak{A}$ be the derived scale of the elements of $\mathfrak{A}$ with respect to a criterion $C \in \mathfrak{C}$. We define the dependence weight

$$\Phi_C(A_j) = \sum_{A_i \in D_{A_j}} \Psi_{A_i C}(A_j)\Psi_C(A_i)$$

If the elements of $\mathfrak{A}$ are inner dependent with respect to $C \in \mathfrak{C}$, then $\Psi_C(A_i) \neq \Psi_C(A_j)$ for some $A_j \in \mathfrak{A}$.

Hierarchic composition yields multilinear forms which are of course nonlinear and have the form

$$\sum_{i_i,\dots,i_p} x_1^{i_1} x_2^{i_2} \cdots x_p^{i_p}$$

The richer the structure of a hierarchy in breadth and depth the more complex are the derived multilinear forms from it. There seems to be a good opportunity to investigate the relationship obtained by composition to covariant tensors and their algebraic properties. More concretely we have the following covariant tensor for the priority of the ith element in the hth level of the hierarchy.

$$w_i^h = \sum_{i_2,\dots,i_{h-1}=1}^{N_{h-1},\dots,N_1} w_{i_1,i_2}^{h-1} \cdots w_{i_{h-2}i_{h-1}}^2 w_{i_{h-1}}^1 \qquad i_i \equiv i$$

The composite vector for the entire hth level is represented by the vector with covariant tensorial components. Similarly, the left eigenvector approach to a hierarchy gives rise to a vector with contravariant tensor components. Tensors, are generalizations of scalars (which have no indices), vectors (which have a single index), and matrices or arrays (which have two indices) to an arbitrary number of indices. They are widely known and used in physics and engineering.

8.4 Expectations

Expectations are beliefs about the rank of alternatives derived from prior knowledge. Assume that a decision maker has a ranking, arrived at intuitively, of a finite set of alternatives $\mathfrak{A}$ with respect to prior knowledge of criteria $\mathfrak{C}$.

Axiom 4.

1. *Completeness:* $\mathfrak{C} \subset \mathfrak{H} - L_h,\ \mathfrak{A} = L_h$.
2. *Rank: To preserve rank independently of what and how many other alternatives there may be. Alternatively, to allow rank to be influenced by the number and the measurements of alternatives that are added to or deleted from the set.*

This axiom simply says that those thoughtful individuals who have reasons for their beliefs should make sure that their ideas are adequately represented for the outcome to match these expectations; i.e., all criteria are represented in the hierarchy. It assumes neither that the process is rational nor that it can accommodate only a rational outlook. People could have expectations that are branded irrational in someone else's framework. It also says that the rank of alternatives depends both on the expectations of the decision maker and on the nature of a decision problem.

Now we illustrate how the foregoing theory can be used in practice. The following example uses hierarchies to model the Iran crisis in the context of the Middle East conflict and the fear of Israel and the West about Iran having nuclear weapons by refining fissionable material in quantities beyond what is required for satisfying Iran's energy needs.

9 Example: AHP Analysis of Strategies towards Iran

The threat of war in Iran is a complex and controversial issue, involving many actors in different regions and several possible courses of action. Nearly 40 people were involved in the exercise done in October 2007. They were divided into groups of 4 or 5 and each of these groups worked out the model and derived results for a designated merit: benefits, opportunities, costs or risks. In the end there were two outcomes for each merit which were combined using the geometric mean as described in the section on group decision making and then the four resulting outcomes were combined into a single overall outcome as described below. It should be understood that this is only an exercise to illustrate use of the method and no real life conclusions should be drawn from it primarily because it did not involve political expert and negotiators from all the interested parties. Its conclusions should be taken as hypotheses to be further tested. In the summer of 2008 many people, including authors who write from Israel who want to prevent Iran from acquiring nuclear power the most, consider that attacking Iran would lead to great harm to the economies of the world because nearly 45% of the world's oil flows out of the Persian Gulf, and Iran would then make sure of its disruption.

9.1 Creating the Model

A model for determining the policy to pursue towards Iran seeking to obtain weapon grade nuclear material was designed using a benefits (B), opportunities (O), costs (C), and risks (R) or combined (BOCR) model. The benefits model shows which alternative would be most beneficial, the opportunities model shows which alternative has the greatest potential for benefits, the costs model (costs may include monetary, human and intangible costs) shows which alternative would be most costly and the risks model shows which alternative has the highest potential costs.

9.2 Strategic Criteria

Strategic Criteria are used to evaluate the BOCR merits of all decisions by a decision maker. They are the overriding criteria that individuals, corporations or governments use to determine which decision to make first, and what are the relative advantages and disadvantages of that decision.

For policy towards Iran, the BOCR model structured by the group is evaluated using the strategic criteria of *World Peace* (0.361), *Regional Stability* (0.356), *Reduce Volatility* (0.087) and *Reduce Escalation of Middle East Problem* (0.196). The priorities of the strategic criteria indicated in parentheses next to each, are obtained from a pairwise comparisons matrix with respect to the goal of long term peace in the world.

9.3 Control Criteria

The BOCR model is evaluated using the control criteria (focusing thought to answer the question in making pairwise comparisons): *Economic, Political, Rule of Law* and *Security*. They are the criteria for which we are able to represent the different kinds of influences that we are able to perceive which later need to be combined into an overall influence using AHP/ANP calculations.

9.4 Actors

The countries mainly concerned with this problem are: the *US*, *Iran*, *Russia* & *China*, *Middle East* countries and *Israel*.

9.5 Alternatives

The group identified six Alternatives:

(1) It is reasonable to undertake *Aerial Strikes* towards Iran

(2) *Economic Sanctions* should be applied against Iran

(3) The Actors should carry out *Ground Invasion* of Iran

(4) *Israeli Action* towards Iran

(5) To do *Nothing*, leaving everything so as it is

(6) To make efforts to make a *Regime Change*

9.6 BOCR Models

With a view to saving space we do not give all the hierarchies and their matrices of judgments. In this exercise it was determined to keep the structure simple by using the same structure for all four merits (see Figure 4) albeit with different judgments. In particular for the costs and risks one asks the question which is more (not less) costly or risky, and in the end subtract the corresponding values from those of benefits and opportunities. The analysis derives four rankings of the alternatives, one for each of the BOCR merits. Following that one must obtain priorities for the BOCRs themselves in terms of the strategic criteria and use the top ranked alternative for each merit in order to think about that merit and then use those priorities to weigh and synthesize the alternatives. The priorities of the alternatives are proportional to the priority of the top ranked alternative, thus they would all be multiplied by the same number that is the priority of the merit.

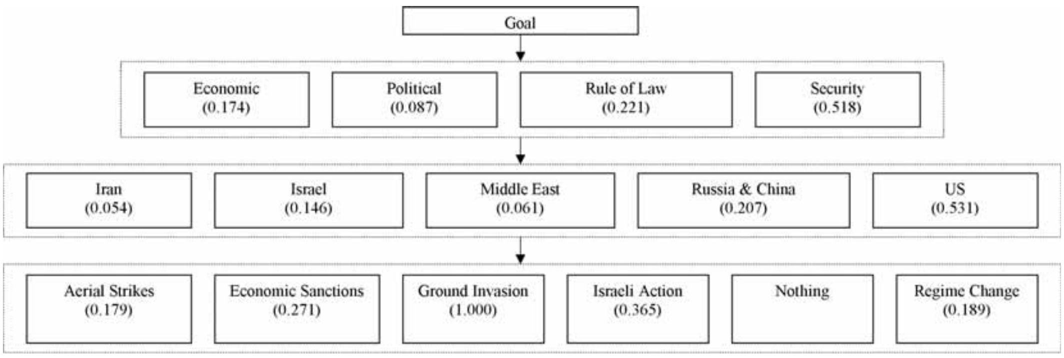

Figure 4. Costs Hierarchy to Choose the Best Strategy towards Iran

It is important to note again that usually for a general decision problem each merit would have a different structure than the other merits. However, for the sake of expediency in this decision, the group decided to use the same structure with the appropriate formulation of the questions to provide the judgments.

9.7 Judgments and Comparisons

As previously mentioned, a judgment is an expression of opinion about the dominance (importance, preference or likelihood) of one element over another. It is done every day through verbal expression that has some quantitative significance that we need to use to combine the many dominance judgments involved in a decision. The set of all such judgments in making comparisons with respect to a single property or goal can be represented by means of a square matrix in which the set of elements is compared. It is a way of organizing all the judgments with respect to some property to be processed and synthesized along with other matrices of comparisons involved in the decision. Each judgment represents the dominance of an element in the left column of the matrix over an element in the row on top. It reflects the answers to two questions: which of the two elements is more important with respect to a higher level criterion, and how strongly.

As usual with the AHP, in the models of benefits, opportunities, cost, and risks the group compared the criteria and subcriteria according to their relative importance with respect to the parent element in the adjacent upper level. For example, the entries in the matrix shown in Table 7 are responses to the question: which control criterion is more important with respect to choosing the best strategy towards Iran and how strongly? Here economic costs are moderately more important than political costs and are assigned the absolute number 3 in the $(1, 2)$ or first-row second-column position. Three signifies three times more. The reciprocal value is automatically entered in the $(2, 1)$ position, where political costs on the left are compared with economic costs at the top. Similarly a 5, corresponding to strong dominance or importance, is assigned to security costs over political costs in the $(4, 2)$ position, and a 2, corresponding to weak or slight dominance, is assigned to the costs of rule of law over political costs in the $(3, 2)$ position with corresponding reciprocals in the transpose positions of the matrix.

Judgments in a matrix may not be consistent. In eliciting judgments, one makes redundant comparisons to improve the validity of the answer, given that respondents may be uncertain or may make poor judgments in comparing some of the elements. Redundancy gives rise to multiple comparisons of an element with other elements and hence to numerical inconsistencies. The group first made all the comparisons using semantic terms from the fundamental scale and then translated them to the corresponding numbers.

Table 7. Judgment Matrix for the Control Criteria of the Costs Hierarchy

Choosing best strategy towards Iran (costs)	Economic	Political	Rule of Law	Security	Normalized Priorities
Economic	1	3	1/2	1/3	0.173
Political	1/3	1	1/2	1/5	0.087
Rule of Law	2	2	1	1/3	0.222
Security	3	5	3	1	0.518

C. R. $= 0.049$

For example, where we compare security costs with economic costs and with political costs, we have the respective judgments 3 and 5. Now if $x = 3y$ and $x = 5z$ then $3y = 5z$ or $y = 5/3z$. If the judges were consistent, economic costs would be assigned the value $5/3$ instead of the 3 given in the matrix. Thus the judgments are inconsistent. In fact, we are not sure which judgments are the accurate ones and which are the cause of the inconsistency. However, these can be determined in a systematic way and improved by interrogating the decision maker.

The process is repeated for all the matrices by asking the appropriate dominance or importance question. For example, the entries in the judgment matrix shown in Table 8 are responses to the question: which party is more committed to ensure security?

Here US security costs are regarded as extremely more important than the security costs for Iran, and 9 is entered in the $(5, 1)$ position and $1/9$ in the $(1, 5)$ position.

In comparing the six strategies (alternatives) towards Iran, we asked members of the decision group which strategy in their opinion would be more costly for each of the actors. For example, for the USA, we obtained a matrix of paired comparisons in Table 9 in which Ground Invasion is the most expensive strategy. On the contrary, regime change and doing nothing are the least costly ones. In this example the criteria (here the different countries) are assumed to be independent of the alternatives and hence the priorities of alternatives are given in ideal form by dividing by the largest priority among them. Here ground invasion would be the most costly.

Table 8. Judgment Matrix of Subcriteria with Respect to Security Costs

Security Costs	Iran	Israel	Middle East	Russia & China	U.S.	Normalized Priorities
Iran	1	1/8	1/3	1/6	1/9	0.029
Israel	8	1	4	1/2	1/7	0.138
Middle East	3	1/4	1	1/5	1/7	0.054
Russia & China	6	2	5	1	1/6	0.182
U. S.	9	7	7	6	1	0.597

C. R. = 0.1

Table 9. Relative Costs of the Strategies for the U. S.

Strategies costs for the U.S.	Aerial Strikes	Economic Sanctions	Ground Invasion	Israeli Action	Nothing	Normalized Priorities	Idealized Priorities
Aerial Strikes	1	1/3	1/7	1/2	3	0.087	0.164
Economic Sanctions	3	1	1/6	2	3	0.160	0.301
Ground Invasionv	7	6	1	6	7	0.533	1.000
Israeli Action	2	1/2	1/6	1	3	0.122	0.229
Nothing	1/3	1/3	1/7	1/3	1	0.058	0.108
Regime Change	1/3	1/3	1/6	1/4	1/3	0.040	0.075

C. R. = 0.08

Tables 10, 11 and 12 give the priorities obtained from all the comparisons for the BOCR. Each column in Table 10 gives in bold face the priorities of the control criteria with respect to which the comparisons are made. For example economic has the value 0.047 under opportunities obtained by comparing it in a matrix with Political, Rule of Law and security whose priorities are also shown in bold face. These priorities sum to one. Similarly under opportunities costs and risks. The priorities of the actors are given under the priority of each of the control criteria in the same column. At the bottom of Table 10 are given the overall priorities of the actors with respect to each of the BOCR obtained by weighting by the priority of the control criteria and adding in the same column. We do not yet combine numbers in the same row but only in the same column. Similarly in Table 11 for the actors and the alternatives. At the bottom of Table 11 we have the overall idealized weights of the alternatives for each of the BOCR. In Table 12 we rate the top alternative for each BOCR merit with respect to each of the strategic criteria using a rating scale derived from comparisons. Usually and for greater precision one should develop a different rating scale for each criterion or subcriterion, but we have simplified the analysis here by adopting a single scale for all the strategic criteria. We then weight the ratings by the priorities of the strategic criteria and add to obtain a weight for each BOCR merit. Finally we normalize these four values to obtain the priorities b, o, c, and r. We then use these priorities in Table 13 to synthesize the idealized weights of the alternatives according to the marginal formula that represents short term solution and the total formula that represents long term solution to the problem.

Next, we rate the top outcome for each of the BOCR against the strategic criteria using the five-level ratings scale obtained from paired comparisons: The synthesized Rating Results are shown in Table 13. We want to evaluate or rate the top alternative for benefits and that for opportunities against the strategic criteria as to how they help with respect to each criterion. We also want to rate the top alternatives for the costs and risks as to how much they hurt or damage with respect to each criterion. This yields the priorities of the BOCR before normalization.

Benefits and Opportunities are positive merits, whereas Costs and Risks are negative. The overbalance

Table 10. Priorities of the Actors with respect to Control Criteria of BOCR groups

Control Criteria	Actors	BENEFITS	OPPORTUNITIES	COSTS	RISKS
Economic		**0.047**	**0.626**	**0.174**	**0.209**
	Iran	0.031	0.066	0.129	0.321
	Israel	0.259	0.041	0.036	0.126
	Middle East	0.166	0.233	0.087	0.306
	Russia & China	0.095	0.120	0.425	0.075
	US	0.449	0.540	0.324	0.173
Political		**0.128**	**0.156**	**0.087**	**0.033**
	Iran	0.043	0.201	0.082	0.506
	Israel	0.311	0.125	0.226	0.1454
	Middle East	0.133	0.494	0.067	0.248
	Russia & China	0.079	0.090	0.152	0.045
	US	0.433	0.090	0.483	0.056
Rule of Law		**0.246**	**0.043**	**0.222**	**0.066**
	Iran	0.031	0.114	0.044	0.317
	Israel	0.347	0.415	0.218	0.118
	Middle East	0.132	0.169	0.060	0.443
	Russia & China	0.101	0.051	0.119	0.059
	US	0.389	0.251	0.559	0.063
Security		**0.579**	**0.175**	**0.518**	**0.692**
	Iran	0.068	0.131	0.029	0.200
	Israel	0.115	0.298	0.138	0.200
	Middle East	0.165	0.308	0.054	0.200
	Russia & China	0.407	0.106	0.181	0.200
	US	0.245	0.156	0.597	0.200
OVERALL					
	Iran	**0.031**	**0.100**	**0.054**	**0.243**
	Israel	**0.259**	**0.115**	**0.204**	**0.177**
	Middle East	**0.166**	**0.284**	**0.153**	**0.240**
	Russia & China	**0.095**	**0.110**	**0.275**	**0.159**
	US	**0.449**	**0.390**	**0.314**	**0.180**

of weights is negative for Ground Invasion and Israeli Action and is positive for Aerial Strikes, Economic Sanctions, doing Nothing and Regime Change. As a result, in the current situation doing Nothing turns out to be the best alternative and Ground Invasion is the worst.

9.8 Sensitivity Analysis

There are many ways of doing sensitivity analysis, we show one of them here. Sensitivity graphs for BOCR groups are shown in Figures 5, (a,b,c,d) respectively. From the software program Superdecisions we see that the results obtained by perturbing the priorities of each of the benefits and opportunities, costs and risks are stable. The model is sensitive to changes of priorities in the BOCR merits. As the priority of Costs increases, the alternative 'Israeli Action' becomes more preferred than 'Ground Invasion' and 'Aerial Strikes' becomes more important than 'Economic Sanctions'. On the other hand, as the priority of Risks increases, the last two alternatives 'Israeli Action' and 'Ground Invasion' trade places in the overall order of ranking. Results obtained for Benefits and Opportunities are stable and 'Nothing' remains the best alternative.

Table 11. Priorities of the Alternatives with respect to the Actors in BOCR groups

Actors	Alternatives	BENEFITS	OPPORTUNITIES	COSTS	RISKS
Iran		**0.031**	**0.100**	**0.054**	**0.243**
	Aerial Strikes	1.000	0.078	0.115	0.701
	Economic Sanctions	1.000	0.452	0.260	1.000
	Ground Invasion	1.000	0.057	1.000	0.294
	Israeli Action	1.000	0.142	0.149	0.762
	Nothing	1.000	1.000	0.068	0.239
	Regime Change	1.000	0.220	0.362	0.882
Israel		**0.259**	**0.115**	**0.204**	**0.177**
	Aerial Strikes	0.214	0.359	0.212	0.240
	Economic Sanctions	0.463	0.069	0.120	0.209
	Ground Invasion	0.128	0.228	0.355	0.385
	Israeli Action	0.070	0.104	1.000	1.000
	Nothing	0.177	0.062	0.210	0.153
	Regime Change	1.000	1.000	0.130	0.052
Middle East		**0.166**	**0.284**	**0.153**	**0.240**
	Aerial Strikes	0.095	0.168	0.257	0.186
	Economic Sanctions	0.357	0.676	0.132	0.073
	Ground Invasion	0.159	0.196	1.000	0.328
	Israeli Action	0.131	0.062	0.483	1.000
	Nothing	1.000	0.371	0.111	0.068
	Regime Change	0.702	1.000	0.175	0.045
Russia & China		**0.095**	**0.110**	**0.275**	**0.159**
	Aerial Strikes	0.778	0.141	0.114	0.190
	Economic Sanctions	0.825	0.259	0.215	0.057
	Ground Invasion	0.331	0.174	1.000	0.379
	Israeli Action	1.000	0.122	0.160	1.000
	Nothing	0.559	1.000	0.071	0.072
	Regime Change	0.303	0.154	0.411	0.086
US		**0.449**	**0.390**	**0.314**	**0.180**
	Aerial Strikes	0.167	1.000	0.164	0.133
	Economic Sanctions	0.231	0.094	0.301	0.076
	Ground Invasion	0.130	0.285	1.000	0.388
	Israeli Action	0.165	0.050	0.229	1.000
	Nothing	1.000	0.150	0.108	0.079
	Regime Change	0.130	0.417	0.075	0.038
OVERALL					
	Aerial Strikes	**0.402**	**0.891**	**0.179**	**0.259**
	Economic Sanctions	**0.586**	**0.496**	**0.271**	**0.221**
	Ground Invasion	**0.258**	**0.378**	**1.000**	**0.368**
	Israeli Action	**0.414**	**0.136**	**0.365**	**1.000**
	Nothing	**1.000**	**0.666**	**0.124**	**0.112**
	Regime Change	**0.651**	**1.000**	**0.189**	**0.165**

It has been recently reported in the media that the world economy would suffer greatly were Iran to be attacked militarily. This appears to be in line with current US military thinking who may or may not have

Table 12. Ratings with Respect to Strategic Criteria of Top Alternatives for BOCR merits
Very High (1), High (0.619), Medium (0.381), Low (0.238), Very Low (0.143)

	World peace (0.362)	Regional Stability (0.356)	Reduce Volatility (0.087)	Reduce Escalation of the Middle East Conflict (0.196)	**Priorities**	**Normalized Priorities**
Benefits	Very High	High	High	Medium	**0.710**	**0.254**
Opportunities	Medium	Low	Medium	Medium	**0.330**	**0.118**
Costs	Very High	Very High	Very High	Medium	**0.878**	**0.314**
Risks	Very High	Very High	Very High	Medium	**0.878**	**0.314**

Table 13. Synthesis of the Alternatives' Overall Priorities for the Four BOCR Merits

	BENEFITS	OPPORTUNITIES	COSTS	RISKS	BO/CR	$bB + oO - cC - rR$
	$\mathbf{b = 0.254}$	$\mathbf{o = 0.118}$	$\mathbf{c = 0.314}$	$\mathbf{r = 0.314}$		
Aerial Strikes	0.402	0.891	0.179	0.259	7.711	0.069
Economic Sanctions	0.586	0.496	0.271	0.221	4.841	0.053
Ground Invasion	0.258	0.378	1.000	0.368	0.265	−0.319
Israeli Action	0.414	0.136	0.365	1.000	0.155	−0.308
Nothing	1.000	0.666	0.124	0.112	**48.077**	**0.258**
Regime Change	0.651	1.000	0.189	0.165	20.814	0.172

used this process, of which they are well aware, to arrive at their conclusion.

10 Networks, Dependence and Feedback

In Figure 6, we exhibit a hierarchy and a network [11, 13]. A hierarchy is comprised of a goal, levels of elements and connections between the elements. These connections are oriented only to elements in lower levels. A network has clusters of elements, with the elements in one cluster being connected to elements in another cluster (outer dependence) or the same cluster (inner dependence). A hierarchy is a special case of a network with connections going only in one direction.

There are two kinds of influence: outer and inner. In the first, one compares the influence of elements in a cluster on elements in another cluster with respect to a control criterion. It is assumed that influences are decomposed into recognizable basic influences known as control criteria such as economic, political, technological, environmental and so on in terms of which all the comparisons and ranking of the alternatives are made. In inner influence one compares the influence of elements in a group on each one. For example if one takes a family of father mother and child, and then take them one at a time say the child first, one asks who contributes more to the child's survival, its father or its mother, itself or its father, itself or its mother. In this case the child is not so important in contributing to its survival as its parents are. But if we take the mother and ask the same question on who contributes to her survival more, herself or her husband, herself would be higher, or herself and the child, again herself. Another example of inner dependence is making electricity. To make electricity you need steel to make turbines, and you need fuel. So we have the electric industry, the steel industry and the fuel industry. What does the electric industry depend on more to make electricity, itself or the steel industry, steel is more important, itself or fuel, fuel industry is much more important, steel or fuel, fuel is more important. The electric industry does not need its own electricity to make electricity. It needs fuel. Its electricity is only used to light the rooms, which it may not even need.

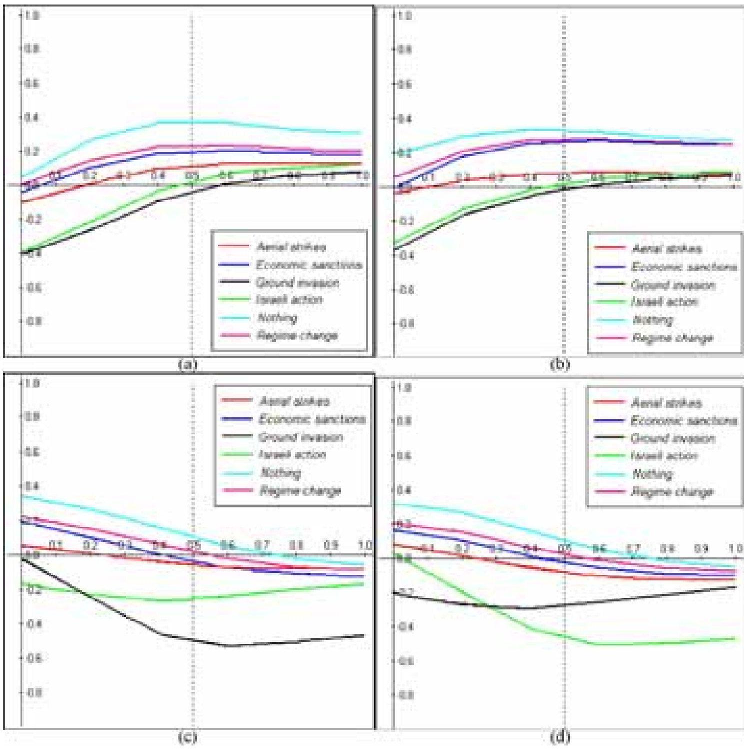

Figure 5. Sensitivity of (a) Benefits, (b) Opportunities, (c) Costs and (d) Risks

If we think about it carefully everything can be seen to influence everything including itself according to many criteria. The world is far more interdependent than we know how to deal with using our existing ways of thinking and acting. The ANP is our logical way to deal with dependence.

The priorities derived from pairwise comparison matrices are entered as parts of the columns of a supermatrix. The supermatrix represents the influence priority of an element on the left of the matrix on an element at the top of the matrix with respect to a particular control criterion. A supermatrix along with an example of one of its general entry matrices is shown in Figure 7. The component C_1 in the supermatrix includes all the priority vectors derived for nodes that are “parent” nodes in the C_1 cluster. Figure 8 gives the supermatrix of a hierarchy and Figure 9 shows the kth power of that supermatrix which is the same as hierarchic composition in the $(k, 1)$ position.

The component C_j in the supermatrix includes all the priority vectors derived for nodes that are “parent” nodes in the C_j cluster. Figure 8 gives the supermatrix of a hierarchy and Figure 9 shows the kth power of that supermatrix which is the same as hierarchic composition in the $(k, 1)$ position.

The $(n, 1)$ entry of the limit supermatrix of a hierarchy as shown in Figure 9 above gives the hierarchic composition principle.

In the ANP we look for steady state priorities from a limit supermatrix. To obtain the limit we must raise the matrix to powers. Each power of the matrix captures all transitivities of an order that is equal to that power. The limit of these powers, according to Cesaro Summability, is equal to the limit of the sum of all the powers of the matrix. All order transitivities are captured by this series of powers of the matrix. The outcome of the ANP is nonlinear and rather complex. The limit may not converge unless the matrix

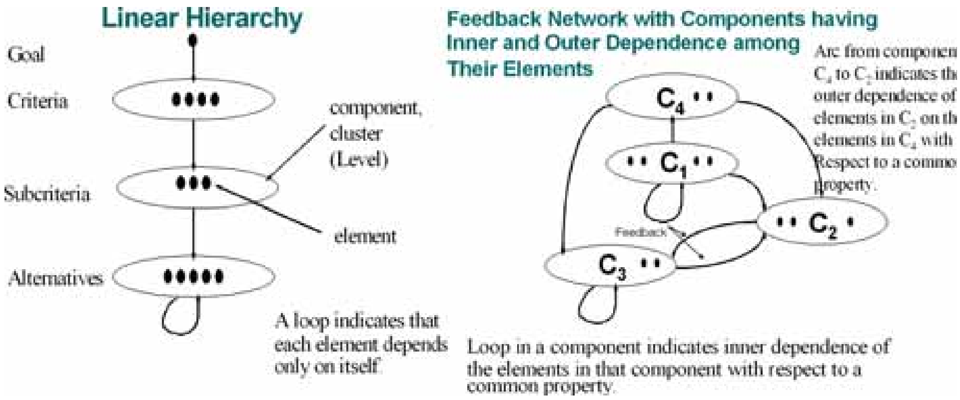

Figure 6. How a Hierarchy Compares with a Network

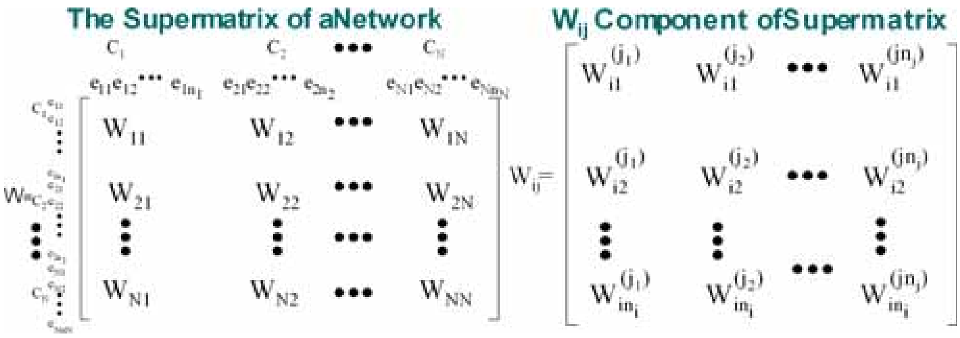

Figure 7. The Supermatrix of a Network and Detail of a Component in it

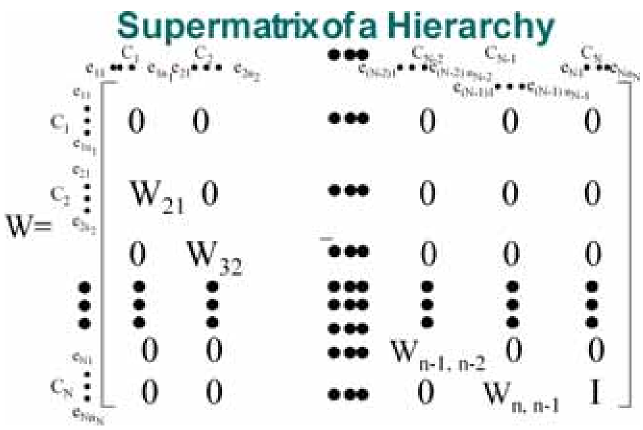

Figure 8. The Supermatrix of a Hierarchy

is column stochastic that is each of its columns sums to one. If the columns sum to one then from the fact that the principal eigenvalue of a matrix lies between its largest and smallest column sums, we know that the principal eigenvalue of a stochastic matrix is equal to one.

$$W^k = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ W_{n,n-1}W_{n-1,n-2}\cdots W_{32}W_{21} & W_{n,n-1}W_{n-1,n-2}\cdots W_{32} & \cdots & W_{n,n-1}W_{n-1,n-2} & W_{n,n-1} & I \end{bmatrix}$$

Figure 9. The Limit Supermatrix of a Hierarchy (Corresponds to Hierarchical Composition)

But for the supermatrix we already know that $\lambda_{\max}(T) = 1$ which follows from:

$$\max \sum_{j=1}^{n} a_{ij} \geq \sum_{j=1}^{n} a_{ij} \frac{w_j}{w_i} = \lambda_{\max} \qquad \text{for } \max w_i$$

$$\min \sum_{j=1}^{n} a_{ij} \leq \sum_{j=1}^{n} a_{ij} \frac{w_j}{w_i} = \lambda_{\max} \qquad \text{for } \min w_i$$

For a row stochastic matrix $1 = \min \sum_{j=1}^{n} a_{ij} \leq \lambda_{\max} \leq \max \sum_{j=1}^{n} a_{ij} = 1$, thus $\lambda_{\max} = 1$.

The same kind of argument applies to a matrix that is column stochastic.

Now we know, for example, from a theorem due to J. J. Sylvester [11] that when the eigenvalues of a matrix W are distinct that an entire function $f(x)$ (power series expansion of $f(x)$ converges for all finite values of x) with x replaced by W, is given by

$$f(W) = \sum_{i=1}^{n} f(\lambda_i) Z(\lambda_i),\ Z(\lambda_i) = \frac{\prod_{j \neq i} (\lambda_j I - A)}{\prod_{j \neq i} (\lambda_j - \lambda_i)},\ \sum_{i=1}^{n} Z(\lambda_i) = I, \quad Z(\lambda_i) Z(\lambda_j) = 0,\ Z^2(\lambda_i) = Z(\lambda_i)$$

where I and 0 are the identity and the null matrices respectively.

A similar expression is also available when some or all of the eigenvalues have multiplicities. We can easily see that if, as we need in our case, $f(W) = W^k$, then $f(\lambda_i) = \lambda_i^k$ and as $k \to \infty$ the only terms that give a finite nonzero value are those for which the modulus of λ_i is equal to one. The fact that W is stochastic ensures this because its largest eigenvalue is equal to one. The priorities of the alternatives (or any set of elements in a component) are obtained by normalizing the corresponding values in the appropriate columns of the limit matrix. When W has zeros and is reducible (its graph is not strongly connected so there is no path from some point to another) the limit can cycle and a Cesaro average over the different limits of the cycles is taken.

11 Why Stochasticity of the Supermatrix is Necessary

Interaction in the supermatrix may be measured according to several different criteria. To display and relate the criteria, we need a separate control hierarchy that includes the criteria and their priorities (see examples below). For each criterion, a different supermatrix of impacts is developed, and in terms of that criterion the components are compared according to their relative impact (or absence of impact) on each other component at the top of the supermatrix, thus developing priorities to weight the block matrices of eigenvector columns under that component in the supermatrix. The resulting stochastic matrix is known as the weighted supermatrix. As we shall see below, it needs to be stochastic to derive meaningful limiting priorities.

Before taking the limit, the supermatrix must first be reduced to a matrix, each of whose columns sums to unity, resulting in what is known as a column stochastic matrix. In general, a supermatrix is not stochastic. This is because its columns are made up of several eigenvectors whose entries in normalized form sum to one and hence that column sums to the number of nonzero eigenvectors. In order to transform it to a stochastic matrix we need to compare its clusters, according to their impact on each other with respect to the general control criterion we have been considering and thus must do it several times for a decision problem, once for each control criterion, and for that criterion several matrices are needed. Each one is used to compare the influence of all the clusters on a given cluster to which they are connected. This yields an eigenvector of influence of all the clusters on each cluster. Such a vector would have zero components when there is no influence. The priority of a component of such an eigenvector is used to weight all the elements in the block of the supermatrix that corresponds to the elements of both the influencing and the influenced cluster. The result is a stochastic supermatrix. This is not a forced way to make the matrix stochastic. It is natural. Why? Because the elements are compared among themselves and one needs information about the importance of the clusters to which they belong, to determine their relative overall weight among all the elements in the other clusters. Here is an example of why it is necessary to weight the priorities of the elements by those of their clusters: If one shouts into a room, “Ladies and Gentlemen, the president”, everyone is alerted and somewhat awed to expect to see the president of the United States because he is in the news so often. But if the announcement is then followed by, "of the garbage collection association", the priority immediately drops according to the importance of the group to which that president belongs. We cannot avoid such a consideration.

The supermatrix of a hierarchy given above is already column stochastic and its clusters have equal weights. As a result, all the blocks of the matrix are multiplied by the same number. Thus the clusters do not have to be weighted. Its limit matrix shown in Figure 9 has a form whose first entry in the bottom row is the well-known hierarchic composition principle which gives the priorties of the alternatives in the bottom level with respect to the goal at the top. In this case, the limit supermatrix is obtained by raising W to powers, but in this case the kth power ($k \geq n-1$) is sufficient to derive the principle of hierarchic composition in its $(k, 1)$ position.

If the supermatrix is stochastic, the limiting priorities depend on its reducibility, primitivity, and cyclicity, with four cases to consider (see Table 14 below). Both acyclic cases are illustrated here. A matrix is reducible if on a permutation of rows and columns it can be put in the form $\left[\begin{smallmatrix} B_1 & 0 \\ B_2 & B_3 \end{smallmatrix}\right]$ where B_1 and B_3 are square submatrices. Otherwise A is irreducible or non-decomposable. It is clear that the supermatrix of a hierarchy is reducible. Its principal eigenvalue $\lambda_{\max}$ is a multiple eigenvalue. A matrix is primitive if some power of it is positive. Otherwise it is called imprimitive. A matrix has to be reducible in order for its powers to cycle which is best illustrated by the following example of successive powers of a matrix and how they shift in an orderly cyclic way the nonzero entries from one power to the next:

$$W = \begin{bmatrix} 0 & W_{12} & 0 \\ 0 & 0 & W_{23} \\ W_{31} & 0 & 0 \end{bmatrix}; \qquad W^2 = \begin{bmatrix} 0 & 0 & W_{12}W_{23} \\ W_{23}W_{31} & 0 & 0 \\ 0 & W_{31}W_{12} & 0 \end{bmatrix}$$

$$W^3 = \begin{bmatrix} W_{12}W_{23}W_{31} & 0 & 0 \\ 0 & W_{23}W_{31}W_{12} & \\ 0 & 0 & W_{31}W_{12}W_{23} \end{bmatrix}$$

$$W^{3k} = \begin{bmatrix} (W_{12}W_{23}W_{31})^k & 0 & 0 \\ 0 & (W_{23}W_{31}W_{12})^k & 0 \\ 0 & 0 & (W_{31}W_{12}W_{23})^k \end{bmatrix}$$

$$W^{3k+1} = \begin{bmatrix} 0 & (W_{12}W_{23}W_{31})^k W_{12} & 0 \\ 0 & 0 & (W_{23}W_{31}W_{12})^k W_{23} \\ (W_{31}W_{12}W_{23})^k W_{31} & 0 & 0 \end{bmatrix}$$

$$W^{3k+2} = \begin{bmatrix} 0 & 0 & (W_{12}W_{23}W_{31})^k W_{12}W_{23} \\ (W_{23}W_{31}W_{12})^k W_{23}W_{31} & 0 & 0 \\ 0 & (W_{31}W_{12}W_{23})^k W_{31}W_{12} & 0 \end{bmatrix}$$

Table 14. Characterization of W^∞ in Terms of Eigenvalue Multiplicity

	Acyclic	Cyclic
Irreducible	$\lambda_{\max} = 1$ is a simple root	C other eigenvalues with modulus $= 1$ (they occur in conjugate pairs)
Reducible	$\lambda_{\max} = 1$ is a multiple root	C other eigenvalues with modulus $= 1$ (they occur in conjugate pairs)

Let W be the stochastic matrix for which we wish to obtain $f(W) = W^\infty$.

What we must do now is find a way to derive priorities for these four cases. We will consider the two cases when $\lambda_{\max} = 1$ is simple and then again when it is a multiple root.

Formally, because the right hand side in the decomposition formula above is a polynomial in W multiplying both sides by W^∞ each term on the right would become a constant multiplied by W^∞ and the final outcome is also a constant multiplied by W^∞. Because we are only interested in the relative values of the entries in W^∞ we can ignore the constant and simply raise W to very large powers which the computer program *SuperDecisions* does in this case of distinct eigenvalues.

Next we consider the case where $\lambda_{\max} = 1$ is a multiple eigenvalue. For that case we have what is known as the confluent form of Sylvester's theorem:

$$f(W) = \sum_{j=1}^{k} T(\lambda_i) = \sum_{j=1}^{k} \frac{1}{(m_i - 1)!} \frac{\mathrm{d}^{m-1}}{\mathrm{d}\lambda^{m-1}} f(\lambda)(\lambda I - W)^{-1} \left. \frac{\prod_{i=1}^{n}(\lambda - \lambda_i)}{\prod_{i=m+1}^{n}(\lambda - \lambda_i)} \right|_{\lambda - \lambda_1}$$

where k is the number of distinct roots and m_i is the multiplicity of the root λ_i. However, as we show below, this too tells us that to obtain the limit priorities it is sufficient to raise W to arbitrarily large power to obtain a satisfactory decimal approximation to W^∞.

The only possible nonzero survivors as we raise the matrix to powers are those λ's that are equal to one or are roots of one. If the multiplicity of the largest real eigenvalue $\lambda_{\max} = 1$ is n_1, then we have

$$W^\infty = n_1 \left. \frac{\frac{\mathrm{d}^{(n_1-1)}}{\mathrm{d}\lambda^{(n_1-1)}} \left[(\lambda I - W)^{-1} \Delta(\lambda) \right]}{\Delta^{(n_1)}(\lambda)} \right|_{\lambda=1}$$

where one takes derivatives of the characteristic polynomial of the matrix W, and $\Delta(\lambda) = \det(\lambda I - W) = \lambda^n + p_1\lambda^{n-1} + \cdots + p_n$. Also $(\lambda I - W)^{-1} = F(\lambda)/\Delta(\lambda)$ and

$$F(\lambda) = W^{n-1} + (\lambda + p_1)W^{n-2} + (\lambda^2 + p\lambda_1 + p_2)W^{n-3} + \cdots + (\lambda^{n-1} + p_1\lambda^{n-2} + \cdots + p_{n-1})I$$

is the adjoint of $(\lambda I - W)$.

Now the right side is a polynomial in W. Again, if we multiply both sides by W^∞, we would have on the right a constant multiplied by W^∞ which means that we can obtain W^∞ by raising W to large powers.

For the cases of roots of one when $\lambda_{\max} = 1$ is a simple or a multiple root let us again formally see what happens to our polynomial expressions on the right in both of Sylvester's formulas as we now multiply both on the left and on the right first by $(W^c)^\infty$ obtaining one equation and then again by $(W^{c+1})^\infty$ obtaining another and so on c times, finally multiplying both sides by $(W^{c+c-1})^\infty$. We then sum these equations and take their average on both sides. The left side of each of the equations reduces to W^∞ and the average is $\frac{1}{c}W^\infty$. On the right side the sum for each eigenvalue that is a root of unity is simply a constant times the sum $(W^c)^\infty + (W^{c+1})^\infty + \cdots + (W^{c+c-1})^\infty$. Also, because this sum is common to all the eigenvalues, it factors out and their different constants sum to a new constant multiplied by $(1/c)$. This is true whether one is a simple or a multiple eigenvalue because the same process applies to accumulating its constants. In the very end we simply have $\frac{1}{c}\left[(W^c)^\infty + (W^{c+1})^\infty + \cdots + (W^{c+c-1})^\infty\right] = \frac{1}{c}(1+W+\cdots+W^{c-1})(W^c)^\infty$, $c \geq 2$, which amounts to averaging over a cycle of length c obtained in raising W to infinite power. The cyclicity c can be determined, among others, by noting the return of the form of the matrix of powers of W to the original form of blocks of zero in W.

Caution: Some interesting things can happen in the limit supermatrix when it is reducible. For example if we have multiple goals in a hierarchy that are not connected to a higher goal, that is if we have multiple sources, we may have several limit vectors for the alternatives and these must be synthesized somehow to give a unique answer. To do that, the sources need to be connected to a higher goal and prioritized with respect to it. Otherwise, the outcome would not be unique and we would obtain nothing that is meaningful in a cooperative decision (but may be useful in a non-cooperative problem where the goals for example, are different ways of facing an opponent). It is significant to note that a hierarchy always has a single source node (the goal) and a single sink cluster (the alternatives), yet its supermatrix is reducible. Only when the supermatrix is irreducible, and thus its graph is strongly connected with a path from any node or cluster to any other node or cluster, that the columns of the supermatrix would be identical. It is rare that the supermatrix of a decision problem is irreducible. If the source clusters do not have sufficient interaction to serve as a single source, one could take the average of the alternatives relating to the several sources as if they are equally important to obtain a single overall outcome. As we have seen with the supermatrix of a hierarchy in Figure 10, for sink nodes or clusters, at which arrows representing influence terminate and no arrows are directed outward from them, there needs to be a loop representing the identity matrix so that when the supermatrix is raised to infinite powers, they receive values that are generally different from zero. For further emphasis of the latter observation, we work out briefly a useful example in the next section.

12 ANP Formulation of the Classic AHP School Example

We show below in Figure 10 the hierarchy, and in Figure 11 its corresponding supermatrix, and limit supermatrix to obtain the priorities of three schools involved in a decision to choose the best one. They are precisely what one obtains by hierarchic composition using the AHP. The priorities of the criteria with respect to the goal and those of the alternatives with respect to each criterion are clearly discernible in the supermatrix itself. Note that there is an identity submatrix for the alternatives with respect to the alternatives in the lower right hand part of the matrix. The level of alternatives in a hierarchy is a sink cluster of nodes that absorbs priorities but does not pass them on. This calls for using an identity submatrix for them in the supermatrix.

13 Market Share Example

13.1 An ANP Network with a Single Control Criterion - Market Share

A market share estimation model is structured as a network of clusters and nodes. The object is to try to determine the relative market share of competitors in a particular business, or endeavor, by considering what affects market share in that business and introduce them as clusters, nodes and influence links in a network. The decision alternatives are the competitors and the synthesized results are their relative dominance. The relative dominance results can then be compared against some outside measure such as dollars. If dollar

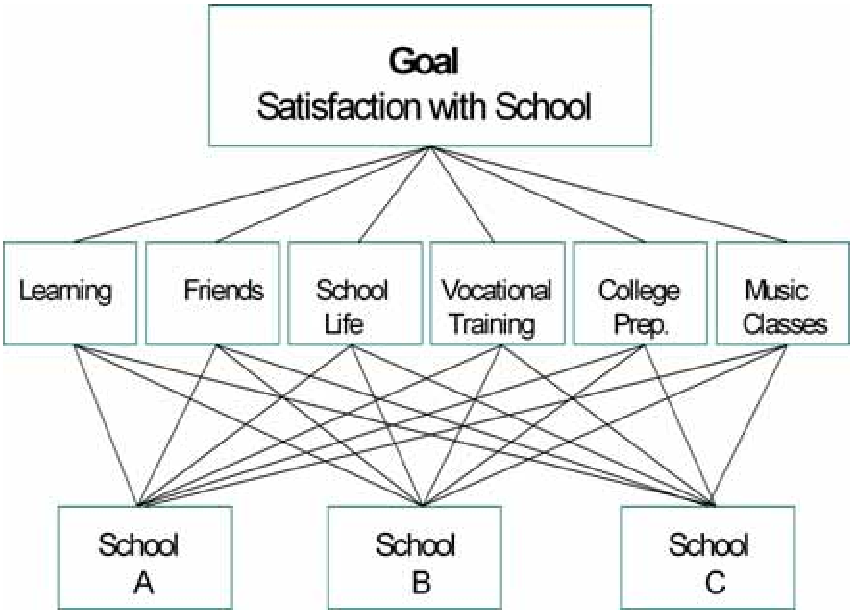

Figure 10. The School Choice Hierarchy

The School Hierarchy as Supermatrix

	Goal	Learning	Friends	School life	Vocational training	College preparation	Music classes	A	B	C
Goal	0	0	0	0	0	0	0	0	0	0
Learning	0.32	0	0	0	0	0	0	0	0	0
Friends	0.14	0	0	0	0	0	0	0	0	0
School life	0.03	0	0	0	0	0	0	0	0	0
Vocational training	0.13	0	0	0	0	0	0	0	0	0
College preparation	0.24	0	0	0	0	0	0	0	0	0
Music classes	0.14	0	0	0	0	0	0	0	0	0
Alternative A	0	0.16	0.33	0.45	0.77	0.25	0.69	1	0	0
Alternative B	0	0.59	0.33	0.09	0.06	0.5	0.09	0	1	0
Alternative C	0	0.25	0.34	0.46	0.17	0.25	0.22	0	0	1

Limiting Supermatrix & Hierarchic Composition

	Goal	Learning	Friends	School life	Vocational training	College preparation	Music classes	A	B	C
Goal	0	0	0	0	0	0	0	0	0	0
Learning	0	0	0	0	0	0	0	0	0	0
Friends	0	0	0	0	0	0	0	0	0	0
School life	0	0	0	0	0	0	0	0	0	0
Vocational training	0	0	0	0	0	0	0	0	0	0
College preparation	0	0	0	0	0	0	0	0	0	0
Music classes	0	0	0	0	0	0	0	0	0	0
Alternative A	0.3676	0.16	0.33	0.45	0.77	0.25	0.69	1	0	0
Alternative B	0.3781	0.59	0.33	0.09	0.06	0.5	0.09	0	1	0
Alternative C	0.2540	0.25	0.34	0.46	0.17	0.25	0.22	0	0	1

Figure 11. The Limit Supermatrix of the School Choice Hierarchy shows same Result as Hierarchic Composition

income is the measure being used, the incomes of the competitors must be normalized to get it in terms of relative market share.

The clusters might include customers, service, economics, advertising, and the quality of goods. The customers' cluster might then include nodes for the age groups of the people that buy from the business: teenagers, 20–33 year olds, 34–55 year olds, 55–70 year olds, and over 70. The advertising cluster might include newspapers, TV, Radio, and Fliers. After all the nodes are created one starts by picking a node and linking it to the other nodes in the model that influence it. The "children" nodes will then be pairwise compared with respect to that node as a "parent" node. An arrow will automatically appear going from the cluster the parent node cluster to the cluster with its children nodes. When a node is linked to nodes in its own cluster, the arrow becomes a loop on that cluster and we say there is inner dependence.

The linked nodes in a given cluster are pairwise compared for their influence on the node they are linked from (the parent node) to determine the priority of their influence on the parent node. Comparisons are made as to which is more important to the parent node in capturing "market share". These priorities are then entered in the supermatrix for the network.

The clusters are also pairwise compared to establish their importance with respect to each cluster they are linked from, and the resulting matrix of numbers is used to weight the corresponding blocks of the original unweighted supermatrix to obtain the weighted supermatrix. This matrix is then raised to powers until it converges to yield the limit supermatrix. The relative values for the companies are obtained from the columns of the limit supermatrix that in this case are all the same because the matrix is irreducible. Normalizing these numbers yields the relative market share.

If comparison data in terms of sales in dollars, or number of members, or some other known measures are available, one can use these relative values to validate the outcome. The AHP/ANP has a compatibility metric to determine how close the ANP result is to the known measure. It involves taking the Hadamard product of the matrix of ratios of the ANP outcome and the transpose of the matrix of ratios of the actual outcome summing all the coefficients and dividing by n^2. The requirement is that the value should be close to 1.

We will give two examples of market share estimation showing details of the process in the first example and showing only the models and results in the second example.

13.2 Estimating The Relative Market Share Of Walmart, Kmart And Target

The network for the ANP model shown in Figure 12 well describes the influences that determine the market share of these companies. We will not dwell on describing the clusters and nodes.

13.3 The Unweighted Supermatrix

The unweighted supermatrix is constructed from the priorities derived from the different pairwise comparisons. The column for a node contains the priorities of all the nodes that have been pairwise compared with respect to it and influence it with respect to the control criterion "market share". The supermatrix for the network in Figure 12 is shown in two parts in Tables 15 and 16.

13.4 The Cluster Matrix

The cluster themselves must be compared to establish their relative importance and use it to weight the corresponding blocks of the supermatrix to make it column stochastic. A cluster impacts another cluster when it is linked from it, that is, when at least one node in the source cluster is linked to nodes in the target cluster. The clusters linked from the source cluster are pairwise compared for the importance of their impact on it with respect to market share, resulting in the column of priorities for that cluster in the cluster matrix. The process is repeated for each cluster in the network to obtain the matrix shown in Table 17. An interpretation of the priorities in the first column is that Merchandise (0.442) and Locations (0.276) have the most impact on Alternatives, the three competitors.

13.5 Weighted Supermatrix

The weighted supermatrix shown in Tables 18 and 19 is obtained by multiplying each entry in a block of the component at the top of the supermatrix by the priority of influence of the component on the left from

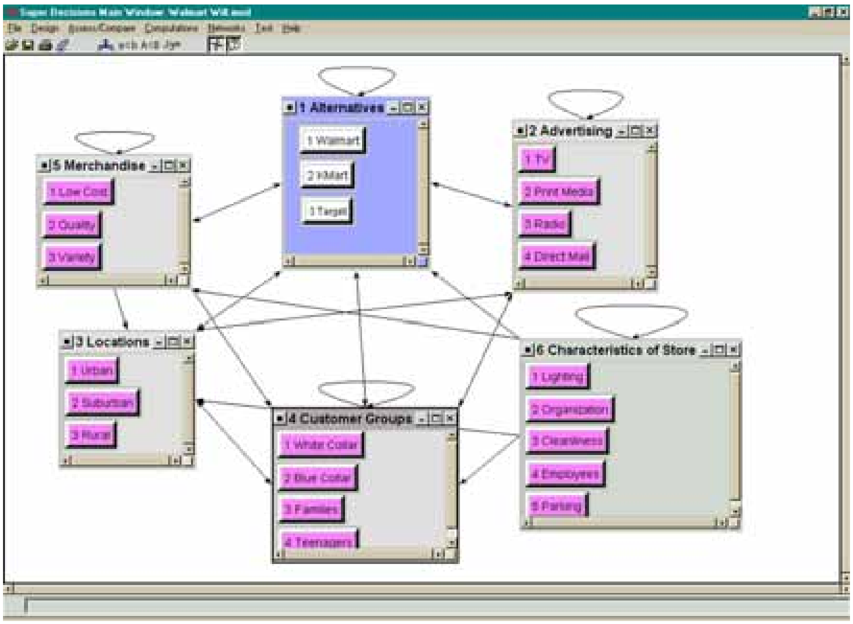

Figure 12. The Clusters and Nodes of a Model to Estimate the Relative Market Share of Walmart, Kmart and Target

the cluster matrix in Table 17. For example, the first entry, 0.137, in Table 17 is used to multiply each of the nine entries in the block (Alternatives, Alternatives) in the unweighted supermatrix shown in Tables 15 and 16. This gives the entries for the Alternatives, Alternatives component in the weighted supermatrix of Table 18. Each column in the weighted supermatrix has a sum of 1, and thus the matrix is stochastic.

13.6 Limit Supermatrix

The limit supermatrix shown in Tables 20 and 21 is obtained from the weighted supermatrix, as we said above. To obtain the final answer we form the Cesaro average of the progression of successive limit vectors.

13.7 Synthesized Results

The relative market shares of the alternatives, 0.599, 0.248 and 0.154 are displayed as synthesized results in the SuperDecisions Program, shown in the middle column of Table 22. They are obtained by normalizing the values for Walmart, Kmart and Target: 0.057, 0.024 and 0.015, taken from the limit supermatrix. The Idealized values are obtained from the Normalized values by dividing each value by the largest value in that column.

In the AHP/ANP the question arises as to how close one priority vector is to another priority vector of relative measurement. When two vectors are close, we say they are *compatible*. The question is how to measure compatibility in a meaningful way. It turns out that consistency and compatibility can be related in a useful way. Our development of a compatibility measure uses the idea of the Hadamard or element-wise product of two matrices.

Table 15. The Unweighted Supermatrix - Part I

		1 Alternatives			2 Advertising				3 Locations		
		1 Walmart	2 Kmart	3 Target	1 TV	2 Print Media	3 Radio	4 Direct Mail	1 Urban	2 Suburban	3 Rural
1 Alternatives	1 Walmart	0.000	0.833	0.833	0.687	0.540	0.634	0.661	0.614	0.652	0.683
	2 Kmart	0.750	0.000	0.167	0.186	0.297	0.174	0.208	0.268	0.235	0.200
	3 Target	0.250	0.167	0.000	0.127	0.163	0.192	0.131	0.117	0.113	0.117
2 Advertising	1 TV	0.553	0.176	0.188	0.000	0.000	0.000	0.000	0.288	0.543	0.558
	2 Print Media	0.202	0.349	0.428	0.750	0.000	0.800	0.000	0.381	0.231	0.175
	3 Radio	0.062	0.056	0.055	0.000	0.000	0.000	0.000	0.059	0.053	0.048
	4 Direct Mail	0.183	0.420	0.330	0.250	0.000	0.200	0.000	0.273	0.173	0.219
3 Locations	1 Urban	0.114	0.084	0.086	0.443	0.126	0.080	0.099	0.000	0.000	0.000
	2 Suburban	0.405	0.444	0.628	0.387	0.416	0.609	0.537	0.000	0.000	0.000
	3 Rural	0.481	0.472	0.285	0.169	0.458	0.311	0.364	0.000	0.000	0.000
4 Cust.Groups	1 White Collar	0.141	0.114	0.208	0.165	0.155	0.116	0.120	0.078	0.198	0.092
	2 Blue Collar	0.217	0.214	0.117	0.165	0.155	0.198	0.203	0.223	0.116	0.224
	3 Families	0.579	0.623	0.620	0.621	0.646	0.641	0.635	0.656	0.641	0.645
	4 Teenagers	0.063	0.049	0.055	0.048	0.043	0.045	0.041	0.043	0.045	0.038
5 Merchandise	1 Low Cost	0.362	0.333	0.168	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2 Quality	0.261	0.140	0.484	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3 Variety	0.377	0.528	0.349	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6 Characteristic	1 Lighting	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2 Organization	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3 Cleanliness	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4 Employees	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5 Parking	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 16. The Unweighted Supermatrix - Part II

		4 Customer Groups				5 Merchandise			6 Characteristics of Store				
		1 White Collar	2 Blue Collar	3 Families	4 Teens	1 Low Cost	2 Quality	3 Variety	1 Light'ng	2 Organization	3 Clean	4 Employees	5 Park
1 Alternatives	1 Walmart	0.637	0.661	0.630	0.691	0.661	0.614	0.648	0.667	0.655	0.570	0.644	0.558
	2 Kmart	0.105	0.208	0.218	0.149	0.208	0.117	0.122	0.111	0.095	0.097	0.085	0.122
	3 Target	0.258	0.131	0.151	0.160	0.131	0.268	0.230	0.222	0.250	0.333	0.271	0.320
2 Advertising	1 TV	0.323	0.510	0.508	0.634	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2 Print Med.	0.214	0.221	0.270	0.170	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3 Radio	0.059	0.063	0.049	0.096	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4 Direct Mail	0.404	0.206	0.173	0.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3 Locations	1 Urban	0.167	0.094	0.096	0.109	0.268	0.105	0.094	0.100	0.091	0.091	0.111	0.067
	2 Suburban	0.833	0.280	0.308	0.309	0.117	0.605	0.627	0.433	0.455	0.455	0.444	0.293
	3 Rural	0.000	0.627	0.596	0.582	0.614	0.291	0.280	0.466	0.455	0.455	0.444	0.641
4 Customers	1 White Collar	0.000	0.000	0.279	0.085	0.051	0.222	0.165	0.383	0.187	0.242	0.165	0.000
	2 Blue Collar	0.000	0.000	0.649	0.177	0.112	0.159	0.165	0.383	0.187	0.208	0.165	0.000
	3 Families	0.857	0.857	0.000	0.737	0.618	0.566	0.621	0.185	0.583	0.494	0.621	0.000
	4 Teenagers	0.143	0.143	0.072	0.000	0.219	0.053	0.048	0.048	0.043	0.056	0.048	0.000
5 Merchandise	1Low Cost	0.000	0.000	0.000	0.000	0.000	0.800	0.800	0.000	0.000	0.000	0.000	0.000
	2 Quality	0.000	0.000	0.000	0.000	0.750	0.000	0.200	0.000	0.000	0.000	0.000	0.000
	3 Variety	0.000	0.000	0.000	0.000	0.250	0.200	0.000	0.000	1.000	0.000	0.000	0.000
6 Characteristics	1 Lighting	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.169	0.121	0.000	0.250
	2 Organization	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.251	0.000	0.575	0.200	0.750
	3 Cleanliness	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.673	0.469	0.000	0.800	0.000
	4 Employee	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.308	0.304	0.000	0.000
	5 Parking	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.075	0.055	0.000	0.000	0.000

Table 17. The Cluster Matrix

	1 Alternatives	2 Advertising	3 Locations	4 Customer Groups	5 Merchandise	6 Characteristics of Store
1 Alternatives	0.137	0.174	0.094	0.057	0.049	0.037
2 Advertising	0.091	0.220	0.280	0.234	0.000	0.000
3 Locations	0.276	0.176	0.000	0.169	0.102	0.112
4 Customer Groups	0.054	0.429	0.627	0.540	0.252	0.441
5 Merchandise	0.442	0.000	0.000	0.000	0.596	0.316
6 Characteristics of Store	0.000	0.000	0.000	0.000	0.000	0.094

13.8 Compatibility Index

Let us show first that the priority vector $w = (w_1, \ldots, w_n)$ is completely compatible with itself. Thus we form the matrix of all possible ratios $W = (w_{ij}) = (w_i/w_j)$ from this vector. This matrix is reciprocal, that is $w_{ji} = 1/w_{ij}$. The Hadamard product of a reciprocal matrix W and its transpose W^T is given by:

$$W \circ W^T = \begin{bmatrix} \frac{w_1}{w_1} & \cdots & \frac{w_1}{w_n} \\ \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \circ \begin{bmatrix} \frac{w_1}{w_1} & \cdots & \frac{w_n}{w_1} \\ \vdots & \ddots & \vdots \\ \frac{w_1}{w_n} & \cdots & \frac{w_n}{w_n} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} = ee^T$$

The sum of the elements of a matrix A can be written as $e^T Ae$. In particular we have $e^T A \circ A^T e = n^2$ for the sum of the elements of the Hadamard product of a matrix and its transpose. The index of compatibility is the sum resulting from the Hadamard product divided by n^2. Thus a vector is completely compatible with itself as $\frac{n^2}{n^2} = 1$. Now we have an idea of how to define a measure of compatibility for two matrices A and B. It is given by $\frac{1}{n^2} e^T A \circ B^T e$. Note that a reciprocal matrix of judgments that is inconsistent is not itself a matrix of ratios from a given vector. However, such a matrix has a principal eigenvector and thus we speak of the compatibility of the matrix of judgments and the matrix formed from ratios of the principal eigenvector. We have the following theorem for a reciprocal matrix of judgments and the matrix W of the ratios of its principal eigenvector:

Theorem 1.

$$\frac{1}{n^2} e^T A \circ W^T e = \frac{\lambda_{\max}}{n}$$

PROOF. From $Aw = \lambda_{\max} w$ we have $\sum_{j=1}^{n} a_{ij} w_j = \lambda_{\max} w_i$ and

$$\frac{1}{n^2} e^T A \circ W^T e = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \frac{w_j}{w_i} = \frac{\lambda_{\max}}{n}$$

■

We want this ratio to be close to one or in general not much more than 1.01 and be less than this value for small size matrices. It is in accord with the idea that a 10% deviation is at the upper end of acceptability.

13.9 Actual Relative Market Share Based on Sales

The object was to estimate the market share of Walmart, Kmart, and Target. The normalized results from the model were compared with sales shown in Table 22 as reported in the Discount Store News of July 13, 1998, p. 77, of \$58, \$27.5 and \$20.3 billions of dollars respectively. Normalizing the dollar amounts shows their actual relative market shares to be 54.8, 25.9 and 19.2. The relative market share from the model

Table 18. The Weighted Supermatrix - Part I

		1 Alternatives			2 Advertising				3 Locations		
		1 Walmart	2 Kmart	3 Target	1 TV	2 Print Media	3 Radio	4 Direct Mail	1 Urban	2 Suburban	3 Rural
1 Alternatives	1 Walmart	0.000	0.114	0.114	0.120	0.121	0.110	0.148	0.058	0.061	0.064
	2 Kmart	0.103	0.000	0.023	0.033	0.066	0.030	0.047	0.025	0.022	0.019
	3 Target	0.034	0.023	0.000	0.022	0.037	0.033	0.029	0.011	0.011	0.011
2 Advertising	1 TV	0.050	0.016	0.017	0.000	0.000	0.000	0.000	0.080	0.152	0.156
	2 Print Media	0.018	0.032	0.039	0.165	0.000	0.176	0.000	0.106	0.064	0.049
	3 Radio	0.006	0.005	0.005	0.000	0.000	0.000	0.000	0.016	0.015	0.014
	4 Direct Mail	0.017	0.038	0.030	0.055	0.000	0.044	0.000	0.076	0.048	0.061
3 Locations	1 Urban	0.031	0.023	0.024	0.078	0.028	0.014	0.022	0.000	0.000	0.000
	2 Suburban	0.112	0.123	0.174	0.068	0.094	0.107	0.121	0.000	0.000	0.000
	3 Rural	0.133	0.130	0.079	0.030	0.103	0.055	0.082	0.000	0.000	0.000
4 Cust.Groups	1 White Collar	0.008	0.006	0.011	0.071	0.086	0.050	0.066	0.049	0.124	0.058
	2 Blue Collar	0.012	0.011	0.006	0.071	0.086	0.085	0.112	0.140	0.073	0.141
	3 Families	0.031	0.033	0.033	0.267	0.356	0.275	0.350	0.411	0.402	0.404
	4 Teenagers	0.003	0.003	0.003	0.021	0.024	0.019	0.023	0.027	0.028	0.024
5 Merchandise	1 Low Cost	0.160	0.147	0.074	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2 Quality	0.115	0.062	0.214	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3 Variety	0.166	0.233	0.154	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6 Characteristic	1 Lighting	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2 Organization	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3 Cleanliness	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4 Employees	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5 Parking	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 19. The Weighted Supermatrix - Part II

		4 Customer Groups				5 Merchandise			6 Characteristics of Store				
		1 White Collar	2 Blue Collar	3 Families	4 Teens	1 Low Cost	2 Quality	3 Variety	1 Light'ng	2 Organization	3 Clean	4 Employees	5 Park
1Alternatives	1 Walmart	0.036	0.038	0.036	0.040	0.033	0.030	0.032	0.036	0.024	0.031	0.035	0.086
	2 Kmart	0.006	0.012	0.012	0.009	0.010	0.006	0.006	0.006	0.004	0.005	0.005	0.019
	3 Target	0.015	0.007	0.009	0.009	0.006	0.013	0.011	0.012	0.009	0.018	0.015	0.049
2 Advertising	1 TV	0.076	0.119	0.119	0.148	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2 Print Med.	0.050	0.052	0.063	0.040	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3 Radio	0.014	0.015	0.012	0.023	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4 Direct Mail	0.095	0.048	0.040	0.023	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3 Locations	1 Urban	0.028	0.016	0.016	0.018	0.027	0.011	0.010	0.016	0.010	0.015	0.018	0.031
	2 Suburban	0.141	0.047	0.052	0.052	0.012	0.062	0.064	0.071	0.051	0.074	0.073	0.135
	3 Rural	0.000	0.106	0.101	0.098	0.063	0.030	0.029	0.076	0.051	0.074	0.073	0.295
4Customers	1 White Collar	0.000	0.000	0.151	0.046	0.013	0.056	0.042	0.247	0.082	0.156	0.107	0.000
	2 Blue Collar	0.000	0.000	0.350	0.096	0.028	0.040	0.042	0.247	0.082	0.134	0.107	0.000
	3 Families	0.463	0.463	0.000	0.398	0.156	0.143	0.157	0.119	0.257	0.318	0.400	0.000
	4 Teenagers	0.077	0.077	0.039	0.000	0.055	0.013	0.012	0.031	0.019	0.036	0.031	0.000
5 Merchandise	1 Low Cost	0.000	0.000	0.000	0.000	0.000	0.477	0.477	0.000	0.000	0.000	0.000	0.000
	2 Quality	0.000	0.000	0.000	0.000	0.447	0.000	0.119	0.000	0.000	0.000	0.000	0.000
	3 Variety	0.000	0.000	0.000	0.000	0.149	0.119	0.000	0.000	0.316	0.000	0.000	0.000
6 Characteristics	1 Lighting	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.017	0.000	0.097
	2 Organization	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.035	0.000	0.079	0.027	0.290
	3 Cleanliness	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.092	0.044	0.000	0.110	0.000
	4 Employee	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.029	0.042	0.000	0.000
	5 Parking	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.005	0.000	0.000	0.000

Table 20. The Limit Supermatrix - Part I

		1 Alternatives			2 Advertising				3 Locations		
		1 Walmart	2 Kmart	3 Target	1 TV	2 Print Media	3 Radio	4 Direct Mail	1 Urban	2 Suburban	3 Rural
1 Alternatives	1 Walmart	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057
	2 Kmart	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024
	3 Target	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015
2 Advertising	1 TV	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079
	2 Print Media	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053
	3 Radio	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009
	4 Direct Mail	0.039	0.039	0.039	0.039	0.039	0.039	0.039	0.039	0.039	0.039
3 Locations	1 Urban	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022
	2 Suburban	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062
	3 Rural	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069
4 Cust.Groups	1 White Collar	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068
	2 Blue Collar	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125
	3 Families	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240
	4 Teenagers	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036
5 Merchandise	1 Low Cost	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043
	2 Quality	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034
	3 Variety	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028
6 Characteristic	1 Lighting	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2 Organization	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3 Cleanliness	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4 Employees	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5 Parking	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 21. The Limit Supermatrix - Part II

		4 Customer Groups				5 Merchandise			6 Characteristics of Store				
		1 White Collar	2 Blue Collar	3 Families	4 Teens	1 Low Cost	2 Quality	3 Variety	1 Light'ng	2 Organization	3 Clean	4 Employees	5 Park
1Alternatives	1 Walmart	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057
	2 Kmart	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024
	3 Target	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015
2 Advertising	1 TV	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079
	2 Print Med.	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053
	3 Radio	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009
	4 Direct Mail	0.039	0.039	0.039	0.039	0.039	0.039	0.039	0.039	0.039	0.039	0.039	0.039
3 Locations	1 Urban	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022
	2 Suburban	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062
	3 Rural	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069	0.069
4Customers	1 White Collar	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068
	2 Blue Collar	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125
	3 Families	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240
	4 Teenagers	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036
5Merchandise	1 Low Cost	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043
	2 Quality	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034
	3 Variety	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028
6 Characteristics	1 Lighting	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2 Organization	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3 Cleanliness	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4 Employee	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5 Parking	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 22. The Synthesized Results for the Alternatives

Alternatives	Values FromLimit Supermatrix	Normalized Values
Walmart	0.057	0.599
Kmart	0.024	0.248
Target	0.015	0.154

was compared with the sales values by constructing a pairwise matrix from the results vector in column 1 below and a pairwise matrix from results vector in column 3 and computing the compatibility index using the Hadamard multiplication method. The index is equal to 1.016. As that is about 1.01 the ANP results may be said to be close to the actual relative market share.

Table 23. Comparison of Results to Actual Data

Competitor	ANP Results	Dollar Sales	Relative Market Share (normalize the Dollar Sales)
WALMART	59.8	$58.0 billion	54.8
KMART	24.8	$27.5 billion	25.9
TARGET	15.4	$20.3 billion	19.2
		COMPATIBILITY INDEX	**1.016**

14 Outline of Steps of the ANP

1. Describe the decision problem in detail including its objectives, criteria and subcriteria, actors and their objectives and the possible outcomes of that decision. Give details of influences that determine how that decision may come out.

2. Determine the control criteria and subcriteria in the four control hierarchies one each for the benefits, opportunities, costs and risks of that decision and obtain their priorities from paired comparisons matrices. If a control criterion or subcriterion has a global priority of 3% or less, you may consider carefully eliminating it from further consideration. The software automatically deals only with those criteria or subcriteria that have subnets under them. For benefits and opportunities, ask what gives the most benefits or presents the greatest opportunity to influence fulfillment of that control criterion. For costs and risks, ask what incurs the most cost or faces the greatest risk. Sometimes (very rarely), the comparisons are made simply in terms of benefits, opportunities, costs, and risks in the aggregate without using control criteria and subcriteria.

3. Determine the most general network of clusters (or components) and their elements that apply to all the control criteria. To better organize the development of the model as well as you can, number and arrange the clusters and their elements in a convenient way (perhaps in a column). Use the identical label to represent the same cluster and the same elements for all the control criteria.

4. For each control criterion or subcriterion, determine the clusters of the general feedback system with their elements and connect them according to their outer and inner dependence influences. An arrow is drawn from a cluster to any cluster whose elements influence it.

5. Determine the approach you want to follow in the analysis of each cluster or element, influencing (the preferred approach) other clusters and elements with respect to a criterion, or being influenced by other clusters and elements. The sense (being influenced or influencing) must apply to all the criteria for the four control hierarchies for the entire decision.

6. For each control criterion, construct the supermatrix by laying out the clusters in the order they are numbered and all the elements in each cluster both vertically on the left and horizontally at the top. Enter in the appropriate position the priorities derived from the paired comparisons as subcolumns of the corresponding column of the supermatrix.

7. Perform paired comparisons on the elements within the clusters themselves according to their influence on each element in another cluster they are connected to (outer dependence) or on elements in their own cluster (inner dependence). In making comparisons, you must always have a criterion in mind. Comparisons of elements according to which element influences a given element more and how strongly more than another element it is compared with are made with a control criterion or subcriterion of the control hierarchy in mind.

8. Perform paired comparisons on the clusters as they influence each cluster to which they are connected with respect to the given control criterion. The derived weights are used to weight the elements of the corresponding column blocks of the supermatrix. Assign a zero when there is no influence. Thus obtain the weighted column stochastic supermatrix.

9. Compute the limit priorities of the stochastic supermatrix according to whether it is irreducible (primitive or imprimitive [cyclic]) or it is reducible with one being a simple or a multiple root and whether the system is cyclic or not. Two kinds of outcomes are possible. In the first all the columns of the matrix are identical and each gives the relative priorities of the elements from which the priorities of the elements in each cluster are normalized to one. In the second the limit cycles in blocks and the different limits are summed and averaged and again normalized to one for each cluster. Although the priority vectors are entered in the supermatrix in normalized form, the limit priorities are put in idealized form because the control criteria do not depend on the alternatives.

10. Synthesize the limiting priorities by weighting each idealized limit vector by the weight of its control criterion and adding the resulting vectors for each of the four merits: Benefits (B), Opportunities (O), Costs (C) and Risks (R). There are now four vectors, one for each of the four merits. An answer involving marginal values of the merits is obtained by forming the ratio BO/CR for each alternative from the four vectors. The alternative with the largest ratio is chosen for some decisions. Companies and individuals with limited resources often prefer this type of synthesis.

11. Governments prefer this type of outcome. Determine strategic criteria and their priorities to rate the four merits one at a time. Normalize the four ratings thus obtained and use them to calculate the overall synthesis of the four vectors. For each alternative, subtract the costs and risks from the sum of the benefits and opportunities. At other times one may subtract the costs from one and risks from one and then weight and add them to the weighted benefits and opportunities. This is useful for predicting numerical outcomes like how many people voted for an alternative and how many voted against it. In all, we have three different formulas for synthesis.

12. Perform sensitivity analysis on the final outcome and interpret the results of sensitivity observing how large or small these ratios are. Can another outcome that is close also serve as a best outcome? Why? By noting how stable this outcome is. Compare it with the other outcomes by taking ratios. Can another outcome that is close also serve as a best outcome? Why?

15 A Complete BOCR Example

Disney Decision: A New Theme Park in Greater China (By Amber Ling-Hui, Lin SzuLun Peng)

15.1 Introduction / Background

In order to enhance operations in foreign market, Disney is constantly searching for areas where it can expand into new markets. According to the projected number of foreign visitors, Walt Disney World expects to increase the current level from 20 percent foreign visitors in domestic parks to 50 percent as well as to expand its theme park business outside the U.S. To achieve these projected numbers Disney needs to make an aggressive attempt to expand its presence in foreign markets, especially Greater China. However, considering the diverse social and economic backgrounds within this area, Disney needs to carefully evaluate the possible benefits as well as the costs and potential risks. In this model, we narrow down the alternatives to Hong Kong, Shanghai, Taiwan and no investment in Greater China. In fact, an awakening and growing middle class in these three areas is exactly the prime target audience for a Disney theme park.

15.2 Ultimate goal for Disney

Disney's intention is to make a minimal equity investment in any operating entity and generate most of its returns through royalty, licensing, and fee income streams.

15.3 Main Model

15.3.1 BOCR Networks and Cluster Definitions

Under the benefits, opportunities, costs, and risks (BOCR) models, different clusters define interactions with respect to the control hierarchy established. The benefits networks indicate the alternatives that yield the most benefit and the opportunities networks indicate the alternative that offers the most opportunities, whereas the costs and risks networks indicate the alternatives that are the most costly or pose the most risk on each alternative.

The flow of the decision process is to first build the networks and sub-networks for each of the BOCR models, make the judgments and evaluate which is the best alternative in each case for this particular decision. The importance of the BOCR must then be determined by rating them with respect to the strategic criteria of the organization or decision maker.

15.3.2 Control Criteria and Subnets of the BOCR

Each of the BOCR in Figure 13 has control criteria whose priorities are established through pairwise comparison. The control criteria in turn have associated network sub-models that contain the alternatives of the decision and clusters of elements. Thus priorities for the alternatives are determined in each of the subnets. These are weighted by their control criterion, and these results are multiplied by the BOCR weights from the rating model and combined to give the final results. The alternatives appear in a cluster in every decision subnet, so we define them only once here. There are three locations being considered for the first Disney theme park in Greater China plus the alternative of not building at all.

15.3.3 Alternatives (in every subnet)

- **Don't invest in Greater China**
- **Hong Kong**
- **Shanghai**
- **Taiwan**

Moving on to the first subnet, under the Social control criterion for Benefits we show the clusters in that network below:

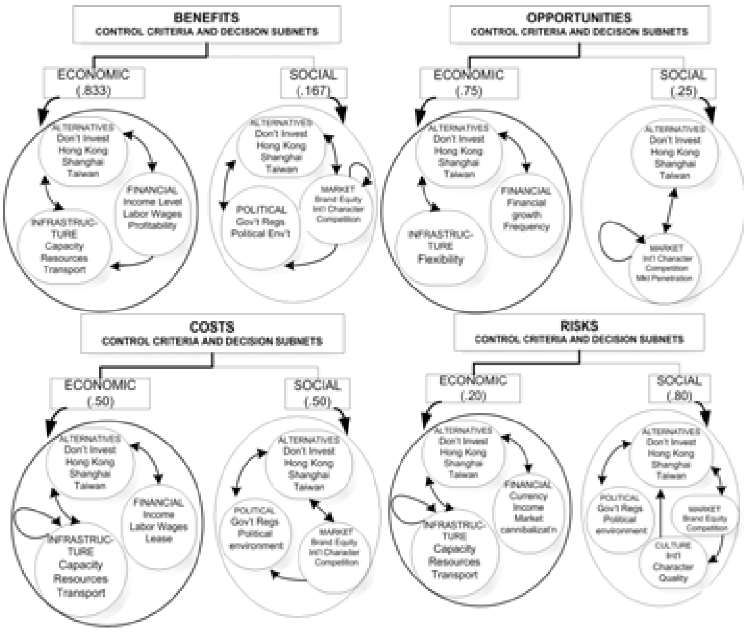

Figure 13. Decision Sub-networks with Clusters and Nodes for each of the BOCR

15.3.4 Clusters in Benefits/Social Subnet

- **Alternatives**
- **Market**

Brand Equity: For the brand equity, we consider it as an intangible asset to Walt Disney. Brand equity represents Disney's reputation and image in the market. Within this subnet, we will examine how much benefit each alternative can bring to Disney in terms of increasing their brand equity.

International Character: International character refers to having a diversified visitor base. The higher the diversification of the visitor base, the more it benefits Disney.

Market Competition: Market competition refers to the number of competitors with *comparable scale* in one market. Within the benefit cluster, we will discuss the level that Disney can benefit from the competition in the market under each alternative.

- **Political Factors**

Government Regulation: We believe a favorable local government regulation on the theme park business will definitely benefit Disney's operation in that area and vise versa.

Political Environment: We believe a stable political environment will create a promising investment environment. Thus, the benefits will be measured base on the current political stability and potential political instability of each alternative.

15.3.5 Interactions between Clusters in the Benefits/Social subnet

In this subnet, we can see the interactions among clusters as well as interactions within clusters.

Market Factors: First of all, since the government regulations and political environment will affect the international character and the market competition in a market, we can see an interaction between market cluster and political factors cluster. Besides, different choices that Disney makes will affect the company itself in terms of brand equity, international character and competition in the market. Finally, the competitive ability of the company and the international character of the market may also affect Disney's brand equity at the end. Thus, we can see another interaction within the market cluster itself.

Political Factors: Besides the interaction with the market cluster, the political factors cluster also interacts with the alternative cluster because the political factors are also affected by different alternatives.

Alternatives: While each alternative affects factors in the market and political clusters, those factors also have effect on Disney's decision among alternatives in return. Thus, there are also backward interactions between the alternatives cluster and the other two clusters.

15.3.6 Nodes in the Benefits/Economic Subnet Clusters

- **Alternatives**
- **Financial Factors**

Gross and disposable income level: Under this factor, only the current gross and disposable income level of the area's citizens will be considered. We assume that a higher income level in the local area will bring more business to the Disney facility and further increase Disney's revenue.

Labor Wage: Labor refers to the current level of local labor wage. A lower labor wage will benefit Disney from reducing operating overheads.

Profitability: Profitability refers to the forecasted profits based on the current market situation.

- **Infrastructure**

Accommodation Capacity: This refers to the current hotel accommodation capacity of that area.

Resources: The resources factor refers to the current construction quality and efficiency of the area.

Transportation: Transportation here means the current development of local railroads, airports, tunnels, etc. If the area is already well developed, Disney can benefit from an instant resource of transportation system for customers.

Table 24. Alternative Rankings from the Benefits/Economic Subnet

Graphic	Alternatives	Total	Normal	Ideal	Ranking
	Don't invest in Greater China	0.0273	0.0579	0.1242	4
	Hong Kong	0.2201	0.4662	1.0000	1
	Shanghai	0.1379	0.2922	0.6267	2
	Taiwan	0.0867	0.1837	0.3940	3

Combining the outcomes from the social and economic decision subnets for the benefits model produces the results shown below. The normalized values (in bold) show that Hong Kong offers the most benefits, and by a significant amount, at 46.4%.

In the opportunities, costs and risks models, the decision subnets are built based on the same logic as that of the benefits subnets. The details of their clusters and nodes are similar to that of benefits and will not be shown here. A general idea of what they are can be obtained from the figure above showing the decision sub-networks. The results for each of the control criteria for opportunities, costs and risks are given below.

We show only the final synthesized results for opportunities, costs, and risks

Table 25. Alternative Rankings from the Benefits/Social Subnet

Graphic	Alternatives	Total	Normal	Ideal	Ranking
	Don't invest in Greater China	0.0045	0.00994	0.0219	4
	Hong Kong	0.2059	0.4521	1.0000	1
	Shanghai	0.1556	0.3417	0.7558	2
	Taiwan	0.0894	0.1963	0.4342	3

Table 26. Synthesized Result for the Benefits Model

Graphic	Alternatives	Total	Normal	Ideal	Ranking
	Don't invest in Greater China	0.107	**0.050**	0.1074	4
	Hong Kong	1.0000	**0.464**	1.000	1
	Shanghai	0.648	**0.301**	0.648	2
	Taiwan	0.401	**0.186**	0.401	3

Table 27. Synthesized Results for the Opportunities Model

Graphic	Alternatives	Total	Normal	Ideal	Ranking
	Don't invest in Greater China	0.019	**0.010**	0.019	4
	Hong Kong	0.428	**0.224**	0.428	3
	Shanghai	1.000	**0.524**	1.000	1
	Taiwan	0.462	**0.242**	0.462	2

Table 28. Synthesized Results for the Costs Model

Graphic	Alternatives	Total	Normal	Ideal	Ranking
	Don't invest in Greater China	0.104	**0.040**	0.105	4
	Hong Kong	0.610	**0.233**	0.617	3
	Shanghai	0.989	**0.378**	1.000	1
	Taiwan	0.912	**0.349**	0.922	2

Table 29. Synthesized Results for the Risks Model

Graphic	Alternatives	Total	Normal	Ideal	Ranking
	Don't invest in Greater China	0.116	**0.051**	0.118	4
	Hong Kong	0.425	**0.188**	0.434	3
	Shanghai	0.981	**0.434**	1.000	1
	Taiwan	0.736	**0.326**	0.751	2

15.3.7 Decision Model for Rating Strategic Criteria

The final step in the decision is to determine the strategic criteria that are more or less the same for the organization or individual in making any decision and use them to rate the BOCR with respect to competition, income level, infrastructure, international character and political support as shown in the table below. We thought the five strategic criteria below pretty well captured Disney's main corporate concerns about their theme parks.

To prepare to rate the strategic criteria one first pairwise compares them for importance in a hierarchy resulting in the priorities shown underneath their names in Table 30. Then one establishes intensities to indicate the degree of fulfillment (in the case of benefits and opportunities) or impact (in the case of costs and risks). The intensities and their priorities (in the ideal form) are Very Strong (1.000), Strong (0.627), Medium (0.382), Moderate (0.232) and Weak (0.148). Priorities are determined for them by pairwise comparing. In this case the same intensities and priorities are used for each strategic criterion, although they could be different.

Table 30. BOCR Ratings and Priorities

Very Strong(1.000), Strong(0.627), Medium(0.382), Moderate(0.232) and Weak(0.148)						
	Competition (0.127)	**Income Level** (0.190)	**Infrastructure** (0.147)	**International Character** (**0.323**)	**Political Support** (**0.214**)	**Priorities**
Benefits	strong	very strong	Strong	very strong	very strong	**0.319**
Opportunities	very strong	strong	Strong	very strong	medium	**0.264**
Costs	very strong	medium	Strong	strong	strong	**0.223**
Risks	very strong	strong	Strong	medium	medium	**0.193**

15.3.8 Strategic Criteria Definitions

The strategic criteria are defined below and pairwise compared for importance with respect to Disney's goal. Ratings are then established for each of these criteria and pairwise compared to establish their priorities in turn. These ratings are then used to determine the priority or importance of Benefits, Opportunities, Costs and Risks and these values are used to weight the results in the submodels attached to them.

Competition. Successful theme parks in the area of the Disney Facility may be viewed both positively and negatively. Other theme parks already in the areas represent competition for Disney; however, competitors may also bring more people to the area to visit both facilities at the same time.

Income Level. Gross and disposable income levels of the area's citizens may also affect the success of the park. Consider Tokyo Disney Land for example. Approximately 95% of its visitors are local Japanese; thus, the high average income level of Japanese does appear to contribute to the tremendous success of Disney in Japan.

Infrastructure. Infrastructure in the area of the park and the regional support are also important. Visitors should be able to access the park easily. The transportation system should be well established or enhanced while the park is being constructed. A good area should have the infrastructure to support a park efficiently. Besides, the region should also contribute to extending the time visitors are able to spend at the Disney facilities. For example, a stock of hotel rooms to support park visitors is important and rooms at a variety of price levels, from economy all the way to luxury, should be available when the park opens.

International Character. Disney is looking for "international character" for any theme park it builds in Greater China. A diversified visitor base will reduce the risks of problems in one country having an adverse effect on the flow of international visitors.

Political Support. In all Disney's international operations, support from local government is critical to the Disney Company. This support ranges from providing a good location to build the theme park to insuring sufficient capital flow.

15.3.9 Rate Benefits, Opportunities, Costs and Risks

To select the ratings in Table 30 for Benefits, for example, one must keep in mind the alternative in the synthesized results for the benefits model given in Table 26 that has the highest priority, Hong Kong. For example, Hong Kong's benefits to fulfill the Competition strategic criterion or objective is thought to be strong. For fulfilling benefits for Income Level, Hong Kong would be very strong as people there have high disposable income, and so on for all the Strategic Criteria.

When making ratings for Costs and Risks, keep in mind that the highest priority alternative is the most costly or most risky. To select the ratings for Risks keep in mind Shanghai. Shanghai has very strong risks so far as Competition is concerned, and strong risks for Income Level as people have less disposable income there, and medium risks for Political Support which means the risk is not too great for Disney in Shanghai as they believe they would have the support of the Chinese Government.

The overall priorities for the BOCR are computed by multiplying and adding across each row and normalizing the final result shown in the last column of Table 30. The priorities show that the most important merit is Benefits at 31.9% followed by Opportunities at 26.4%. This means that the priorities of the alternatives under benefits are weighted more heavily. Benefits at 31.9% drive the decision more than the Risks at 19.3%.

The final results shown in Table 31 are obtained by using the formula $bB + oO - cC - rR$ where b, o, c and r are the priorities for Benefits, Opportunities, Costs and Risks obtained from rating the strategic criteria in Table 30. This formula is applied for the alternatives using the priority vectors from the synthesized results (the B, O, C, and R of the formula) in Tables 25–28. Since this formula involves negatives, the overall synthesized results in Table 30 may be negative, meaning that the alternative is undesirable. Sometimes all results are negative, and one is forced to take the least undesirable one. In Table 31 positive results are labeled blue and negative red. Here Hong Kong is best with the highest positive value and Taiwan is worst with the highest negative value.

Table 31. BOCR Model: Overall Synthesized Results

Graphic	Alternatives	Total	Normal	Ideal	Ranking
(red)	Don't invest in Greater China	−0.006	−0.017	−0.030	3
(blue)	Hong Kong	0.214	0.567	1.000	1
(blue)	Shanghai	0.061	0.161	0.284	2
(red)	Taiwan	−0.096	−0.255	−0.449	4

As we can see, from the overall synthesized results in Table 31, Disney's best option is to build their new theme park in *Hong Kong*.

15.3.10 Sensitivity Analysis Graphs

In Figures 14, the vertical dotted lines represent the priorities of the BOCR. Starting on the left, the curves from top to bottom are for 1) benefits: Do Nothing, Hong Kong Shanghai, Taiwan, 2) opportunities: Hong Kong, Do Nothing, Shanghai, Taiwan, 3) and for both costs and risks: Hong Kong, Shanghai, Taiwan, Do nothing. Sensitivity analysis in Figure 15 shows that when the importance of benefits is greater than 0.05, investing in Hong Kong is the best choice. The dotted vertical line indicates the priority of Benefits, for example. To see what happens as the importance of benefits increases, the vertical line can be moved to the right. Above a benefits priority of about 0.39, the least preferred alternative changes from Taiwan to not investing in Greater China. The curve immediately under Hong Kong is Shanghai. One might interpret this as indicating that investing in China somewhere is imperative in terms of benefits.

At an importance of 22.3% for costs, Hong Kong (the top curve) is most costly and Taiwan (the bottom curve) is least costly, perhaps because of the political uncertainty and lack of supporting infrastructure in Hong Kong. As costs increase, not investing in China is the top curve (after about 39%). At an importance of 19.3% for risks, the top curve is Hong Kong, so it is most risky and the bottom curve is Taiwan, meaning least risky. When risk is less than about 0.50, the preferred alternative is to invest in Hong Kong. As the

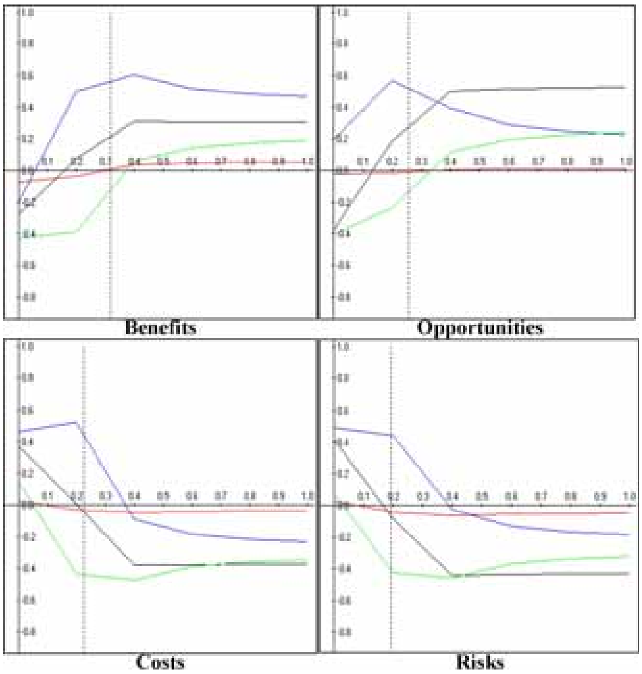

Figure 14. Sensitivity Graphs for the BOCR Networks

importance of risk increases the preferred alternative is to not to invest at all in Greater China, but since the priority is negative and lies below the x-axis, this is not a particularly good alternative, though it is the least negative. Thus, it is extremely risky not to invest in China at all. In sum, the greatest benefits and opportunities lie in mainland China, but also the greatest costs and risks. Netting out both sides, Hong Kong is the preferred alternative for Disney's investment.

By way of validation, this study was done before implementation took place. Now, built on lush Lantau Island just 24 minutes from downtown Hong Kong, Disneyland Hong Kong is the fifth Disneyland style park in the world offering people a world of adventure and entertainment since its opening in September 2005.

16 The Question of Rank Preservation and Reversal in the Discrete Case

In relative measurement the final priority value associated with each object depends on the values of all the other objects with which it is compared so that its measurement does not remain constant but depends on what other objects are being measured. Because of this, if objects were to be rank ordered through comparisons, their priorities and overall rank order may change when other objects are added or old ones deleted. In other words rank reversal can occur. It is known that such reversal can never happen when the objects are compared with respect to one criterion and the judgments are consistent, but only with respect to several criteria. In addition, if it is known that rank must not be allowed to reverse it can be preserved in the Analytic Hierarchy Process (AHP) using both its "Rating" one at a time approach and its comparisons approach. In fact it has been shown experimentally that rank reversal occurs on the average only 8% of the time, but there is good reason for that [16].

One consequence of the belief in scales of measurement as the incontestable and unique way of measurement is the idea that the measurement of objects can and should be done independently of one another and thus the measurement of one object cannot affect the rank order of other objects no matter how important or unimportant the given object may be. It is a natural step to believe that must also be true in multicriteria decisions where we have measurements of the same objects on not one property but several, using different scales, and wish to combine these measurements using criteria weights or other kinds of numbers to develop an overall measurement rank for the alternatives of a decision. This kind of dogma results in what has come to be known as rank preservation. At first glance it seems reasonable and compelling when an alternative is added or deleted, the criteria weights do not depend on the alternatives and particularly when no criteria are added or deleted and judgments are changed, a requirement that naturally suggests itself to suit the process of assigning numbers to objects one at a time.

All decision making methods that rank alternatives one at a time suffer from the following observation. Adding copies or near copies of an alternative until the universe is full of them can depreciate the value and also the rank of that alternative and, as a counter example, negates the possibility of proving a theorem that the rank of independent alternatives must always be preserved when the judgments remain the same and no criteria are added or deleted or their weights changed. A criterion such as "manyness" that represents the number of alternatives cannot be used in the ranking because it forces the dependence of the ranking of each alternative on the existence of every other alternative thus contradicting the assumption of independence. It is illogical (or we might say also wrong) for all multicriteria methods that all use the rating of alternatives one at a time not to take this into account.

There have been numerous objections by various authors to the assumption of rank preservation even when things are measured one at a time in the manner described above. In other words, rank is preserved by some multicriteria methods in cases when, people have argued, it is inappropriate for it to do so and should in fact reverse. In other words there are people who have published in the literature who do not believe in the unconditional rank preservation principle for this kind of measurement.

Three reasons why always preserving rank is not a reliable principle are:

1. There is no theorem or law of nature that confirms the erroneous observation that rank preservation must always hold in multicriteria measurement when one uses one or several scales of measurement each of which measures things one at a time. If one does not use such scales, any fact known about the behavior of things that are measured on such scales need not apply. Any convincing example that violates rank preservation proves that ranking alternatives one at a time is an inappropriate method for multicriteria decision making with independence assumptions all the time.

2. Many papers with examples of rank reversal have been published by distinguished scholars that show legitimate rank reversals and some authors have gone on to make important statements that forcing rank to be always preserved would produce unreasonable results.

3. Paired comparisons can lead to rank reversal among alternatives when new ones are added or deleted. In fact they show that rank preservation is desirable only when the criteria used to evaluate alternatives are extrinsic outside or top down conditions (not intrinsic attributes) imposed on the alternatives. In ranking the intrinsic attributes one must think of the alternatives, and that is a case of the dependence of the criteria on the alternatives and when criteria depend on alternatives, anything can happen to the ranks of the alternatives, that is their ranks can change. But when extrinsic conditions are applied to the alternatives, their ranks must be preserved. This is also true of intrinsic conditions whose importance acquires autonomy over a long time and is not tied to the alternatives one has when a decision is being made.

17 Group Decision Making

Here we consider two issues in group decision making [14]. The first is how to aggregate individual judgments into a representative group judgment, and the second is how to construct a group choice from individual choices. In reality group decisions should not go by consensus because not all people feel the same about things. A minority can have very strong commitments to a cause and can give rise to disruptions that the majority feels lukewarm about. There is no hiding from this issue in the real world. The reciprocal property plays an important role in combining the judgments of several individuals to obtain a judgment for a group. Judgments must be combined so that the reciprocal of the synthesized judgments must be equal to the syntheses of the reciprocals of these judgments. It has been proved that the geometric mean is the unique way to do that. If the individuals are experts, they may not wish to combine their judgments but only their final outcome from a hierarchy. In that case one takes the geometric mean of the final outcomes. If the individuals themselves have different priorities of importance their judgments (final outcomes) are raised to the power of their priorities and then the geometric mean is formed. The following is a summary of results of research I did with Janos Aczel many years ago [2].

17.1 How to Aggregate Individual Judgments

Let the function $f(x_1, \ldots, x_n)$ for synthesizing the judgments given by n judges, satisfy the

(i) *Separability condition* (S): $f(x_1, \ldots, x_n) = g(x_1) \cdots g(x_n)$, for all $x_1, \ldots, x_n$ 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 judgments can be separated as above.]

(ii) *Unanimity condition* (U): $f(x, \ldots, x) = x$ for all x in P. [(U) means that if all individuals give the same judgment x, that judgment should also be the synthesized judgment.]

(iii) *Homogeneity condition* (H): $f(ux_1, \ldots, ux_n) = uf(x_1, \ldots, x_n)$ where $u > 0$ and x_k, ux_k ($k = 1, 2, \ldots, n$) are all in P. [For ratio judgments (H) means that if all individuals judge a ratio u times as large as another ratio, then the synthesized judgment should also be u times as large.]

(iv) *Power conditions* ($\mathrm{P_p}$): $f(x_1^p, \ldots, x_n^p) = f^p(x_1, \ldots, x_n)$. [(P2) for example means that if the kth individual judges the length of a side of a square to be x_k, the synthesized judgment on the area of that square will be given by the square of the synthesized judgment on the length of its side.] Special case ($\mathrm{R} = \mathrm{P}_{-1}$):

$$f\left(\frac{1}{x_1}, \ldots, \frac{1}{x_n}\right) = \frac{1}{f(x_1, \ldots, x_n)}.$$

[(R) is of particular importance in ratio judgments. It means that the synthesized value of the reciprocal of the individual judgments should be the reciprocal of the synthesized value of the original judgments.] Aczel and Saaty [13] proved the following theorem:

Theorem 2. *The general separable (S) synthesizing functions satisfying the unanimity (U) and homogeneity (H) conditions are the geometric mean and the root-mean-power. If moreover the reciprocal property (R) is assumed even for a single n-tuple of the judgments 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: $f(x_1, \ldots, x_n) = g_1(x_1) \cdots g_n(x_n)$ [(WS) implies that not all judging individuals have the same weight when the judgments are synthesized and the different influences are reflected in the different functions $(g_1, \ldots, g_n)$.]

In this situation, Aczel and Alsina [1] proved the following theorem:

Theorem 3. *The general weighted-separable (WS) synthesizing functions with the unanimity (U) and homogeneity (H) properties are the weighted geometric mean* $f(x_1, x_2, \ldots, x_n) = x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}$ *and the weighted root-mean- powers* $f(x_1, x_2, \ldots, x_n) = \sqrt[\gamma]{q_1 x_1^\gamma + q_2 x_2^\gamma + \cdots + q_n x_n^\gamma}$, *where* $q_1 + q_2 + \cdots + q_n = 1$, $q_k > 0$, $k = 1, 2, \ldots, n$, $\gamma > 0$, *but otherwise* $q_1, q_2, \ldots, q_n, \gamma$ *are arbitrary constants.*

If f also has the reciprocal property (R) and for a single set of entries $(x_1, \ldots, x_n)$ of judgments 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:

Theorem 4. *If* $x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)}$, $i = 1, \ldots, m$ *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* $\sum_{i=1}^m a_i = 1$, *then* $(\prod_{i=1}^m x_1^{a_i}), \ldots, (\prod_{i=1}^m x_n^{a_i})$ *are the combined ranks of the alternatives for the* m *judges.*

Note that essentially the geometric mean raises each judgment to the default power of $1/n$ as if each individual's judgment is as important as any other individual. The power or priority of judge i is simply a replication of the judgment 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 following useful theorem was proved in a joint, yet to be published work, with my friend and colleague Luis G. Vargas. It shows how important it is to debate issues in order to mitigate extreme positions as much as we can to make it easier for groups to agree and act cooperatively on their decisions.

Theorem 5 (Bounded Inconsistency)**.** *For a group of inconsistent individuals, the inconsistency of group judgments aggregated by the geometric mean from individual judgments is at most equal to the largest of the individual inconsistencies.*

PROOF. First we show that the inconsistency of a group along cycles of length three satisfies this theorem and is related to the inconsistency index given by

$$\mu_k \equiv \frac{\lambda_{\max}^{(k)} - n}{n-1} = -1 + \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} \left(\varepsilon_{ijk} + \frac{1}{\varepsilon_{ijk}} \right)$$

where $\lambda_{\max}^{(k)}$ is the principal eigenvalue of the matrix $(a_{ijk} = \frac{w_{ik}}{w_{jk}} \varepsilon_{ijk})$. Thus, we have $\lambda_{\max}^{(k)} w_{ik} = \sum_j a_{ijk} w_{jk} = \sum_j \frac{w_{ik}}{w_{jk}} \varepsilon_{ijk} w_{jk} = w_{ik} \sum_j \varepsilon_{ijk}$ and $\lambda_{\max}^{(k)} = \sum_j \varepsilon_{ijk}$ for all i.

Let $(a_{ijk} = \frac{w_{ik}}{w_{jk}} \varepsilon_{ijk})$, $k = 1, 2, \ldots, m$ be the n-by-n matrices of pairwise comparisons of m individuals. Let $\delta_{ijk} = \varepsilon_{ijk} - 1$. By definition inconsistency in the judgments of an individual means that $a_{ijk} a_{jlk} \neq a_{ilk}$ for at least some i, j and l. The inconsistency of a 3-cycle for the kth individual can be measured by $a_{ijk} a_{jlk} a_{lik}$. Clearly, if these judgments were consistent $a_{ijk} a_{jlk} a_{lik} = 1$ for all i, j and l. Thus, inconsistency can be measured with an index that must exclude loops that only connect judgments with themselves preventing the formation of 3-cycles. We have

$$I_k(n,3) = \frac{1}{n(n-1)(n-2)} \sum_{i \neq j \neq l} a_{ijk} a_{jlk} a_{lik} = \frac{1}{n(n-1)(n-2)} \sum_{i \neq j \neq l} \varepsilon_{ijk} \varepsilon_{jlk} \varepsilon_{lik}$$

that is equal to one with consistency. The index corresponding to the matrix of geometric means for the group is given by

$$\begin{aligned}
I(n,3) &= \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l} \prod_k (a_{ijk}a_{jlk}a_{lik})^{1/m} \equiv \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l} a_{ij}a_{jl}a_{li} \\
&= \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l} \varepsilon_{ij}\varepsilon_{jl}\varepsilon_{li} \equiv \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l} \prod_k (\varepsilon_{ijk}\varepsilon_{jlk}\varepsilon_{lik})^{1/m} \\
&\leq \prod_k \left(\frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l} \varepsilon_{ijk}\varepsilon_{jlk}\varepsilon_{lik} \right)^{1/m} = \prod_k [I_k(n,3)]^{1/m} \\
&\leq \max_k \{I_k(n,3)\}
\end{aligned}$$

again having used Holder's inequality. *Thus, the* 3*-cycle inconsistency of a group does not exceed the* 3*-cycle inconsistency of the most inconsistent individual*. Now we show that $I_k(n,3)$ is analytically related to μ_k. It serves as the link to connect μ for the group to $\mu_1, \mu_2, \ldots, \mu_m$, the inconsistency measures of the individuals, so we can say how the inconsistency of individuals judgments affects the inconsistency of the geometric mean of these judgments, i.e., that $\mu \leq \max_k\{\mu_k\}$.

Substituting $1+\delta_{ijk}$ for ε_{ijk} we obtain

$$\begin{aligned}
I_k(n,3) &= \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l} (1+\delta_{ijk})(1+\delta_{jlk})(1+\delta_{lik}) \\
&= \frac{1}{n(n-1)(n-2)} \Big[n(n-1)(n-2) + n\sum_{i\neq j}\delta_{ijk} + n\sum_{j\neq l}\delta_{jlk} + n\sum_{l\neq i}\delta_{lik} \\
&\quad + \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk} + \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk} + \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk} + \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk}\delta_{lik} \Big] \\
&= 1 + \frac{3}{(n-1)} \sum_{i\neq j}\delta_{ijk} + \frac{3}{n(n-1)(n-2)} \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk} \\
&\quad + \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk}\delta_{ilk}
\end{aligned}$$

In addition, $\lambda_{\max}^{(k)} = \sum_j (1+\delta_{ijk}) = n + \sum_j \delta_{ijk}$ and $\mu_k \equiv \frac{\lambda_{\max}^{(k)} - n}{n-1} = \frac{1}{n-1}\sum_j \delta_{ijk}$. Note that for any i, $\delta_{iik} = 0$ for all k. In addition, by summing over i we obtain $\mu_k = \frac{1}{n(n-1)}\sum_i\sum_j \delta_{ijk}$. Now we can write the index $I_k(n,3)$ of the kth individual as a function of μ_k as follows:

$$\begin{aligned}
I_k(n,3) &= 1 + \frac{3}{(n-1)(n-2)} \sum_{i\neq j}\delta_{ijk} + \frac{3}{n(n-1)(n-2)} \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk} \\
&\quad + \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk}\delta_{lik} \\
&= 1 + \frac{3n}{(n-2)}\mu_k + \frac{3}{n(n-1)(n-2)} \sum_{i\neq j}\delta_{ijk}\sum_l \delta_{jlk} + \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk}\delta_{lik} \\
&= 1 + \frac{3n}{(n-2)}\mu_k + \frac{3\mu_k}{n(n-2)} \sum_{i\neq j}\delta_{ijk} + \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk}\delta_{lik} \\
&= 1 + \frac{3n}{(n-2)}\mu_k + \frac{3(n-1)}{(n-2)}\mu_k^2 + \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq l}\delta_{ijk}\delta_{jlk}\delta_{lik}
\end{aligned}$$

In a neighbohood of $\varepsilon_{ij} = 1$ corresponding to near consistency we have

$$I_k(n,3) = 1 + \frac{3n}{(n-2)}\mu_k + o(\delta_{ijk})$$

and for the group index $I(n,3)$ by simply suppressing the subscript k above:

$$\begin{aligned} I(n,3) &= 1 + \frac{3}{(n-1)(n-2)}\sum_{i\neq j}\delta_{ij} + \frac{3}{n(n-1)(n-2)}\sum_{i\neq j\neq l}\delta_{ij}\delta_{jl} + \frac{1}{n(n-1)(n-2)}\sum_{i\neq j\neq l}\delta_{ij}\delta_{jl}\delta_{li} \\ &= 1 + \frac{3n}{(n-2)}\mu + o(\delta_{ij}) \end{aligned}$$

Thus, finaly $I(n,3) = 1 + \dfrac{3n}{n-2}\mu + o(\delta_{ij}) \le \max_k \left\{1 + \dfrac{3n}{n-2}\mu_k + o(\delta_{ijk})\right\}$ and $\mu \le \max_k\{\mu_k\}$. ■

17.2 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, Kenneth Arrow [3] who won the Nobel Prize for his work on this theorem 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:

- *Decisiveness:* the aggregation procedure must generally produce a group order.
- *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.
- *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.
- *No dictator:* no single individual preferences determine the group order.

Using Fundamental Scale of absolute numbers of the AHP, it can be shown that because now the individual preferences are cardinal rather than ordinal, it is *possible* to derive a rational group choice satisfying the above four conditions. We proved this with L. G. Vargas [18]. It is possible because: a) Individual priority scales can always be derived from a set of pairwise cardinal preference judgments as long as they form at least a minimal spanning tree in the completely connected graph of the elements being compared; and b) The cardinal preference judgments associated with group choice belong to an absolute scale that represents the relative intensity of the group preferences.

18 Generalization to the Continuous Case

The foregoing discrete formulation with

$$\sum_{j=1}^{n} a_{ij} w_j = \lambda_{\max} w_i \tag{1}$$

$$\sum_{i=1}^{n} w_i = 1 \tag{2}$$

And with $a_{ji} = 1/a_{ij}$ or $a_{ij}a_{ji} = 1$ (the reciprocal property), $a_{ij} > 0$ generalizes to the continuous case through Fredholm's integral equation of the second kind (which we can also derive directly from first principles to describe the response of a neuron to stimuli) [10]:

$$\int_a^b K(s,t)w(t)\,\mathrm{d}t = \lambda_{\max} w(s) \tag{3}$$

where instead of the matrix A we have a positive kernel, $K(s,t) > 0$. A solution $w(s)$ of this equation is a right eigenfunction.

The standard way in which (3) is written is to move the eigenvalue to the left hand side which gives it the reciprocal form

$$\lambda \int_a^b K(s,t)w(t)\,\mathrm{d}t = w(s) \tag{4}$$

with the normalization condition:

$$\int_a^b w(s)\,\mathrm{d}s = 1 \tag{5}$$

An eigenfunction of this equation is determined to within a multiplicative constant. If $w(t)$ is an eigenfunction corresponding to the charateristic value λ and if C is an arbitrary constant, we see by substituting in the equation that $Cw(t)$ is also an eigenfunction corresponding to the same λ. 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. Here also, we have the reciprocal property

$$K(s,t)K(t,s) = 1 \tag{6}$$

so that $K(s,t)$ is not only positive, but also reciprocal. An example of this type of kernel is $K(s,t) = \mathrm{e}^{s-t} = \mathrm{e}^s/\mathrm{e}^t$. As in the finite case, the kernel $K(s,t)$ is consistent if it satisfies the relation

$$K(s,t)K(t,u) = K(s,u), \qquad \text{for all } s,\ t, \text{ and } u \tag{7}$$

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.

The most important part of what follows is the derivation of the fundamental equation, a functional equation whose solution is an eigenfunction of our basic Fredholm equation.

Theorem 6. *$K(s,t)$ is consistent if and only if it is separable of the form:*

$$K(s,t) = \frac{k(s)}{k(t)} \tag{8}$$

Theorem 7. *If $K(s,t)$ is consistent, the solution of* (4) *is given by*

$$w(s) = \frac{k(s)}{\int_S k(s)\,\mathrm{d}s} = \alpha k(s) \tag{9}$$

We note that this formulation is general and applies to all situations where a continuous scale is needed. It applies equally to the derivation or justification of ratio scales in the study of scientific phenomena. We now determine the form of $k(s)$ and also of $w(s)$.

In the discrete case, the normalized eigenvector was independent of whether all the elements of the pairwise comparison matrix A are multiplied by the same constant a or not, and thus we can replace A by aA and obtain the same eigenvector. Generalizing this result we have:

$$K(as,at) = aK(s,t) = \frac{k(as)}{k(at)} = a\frac{k(s)}{k(t)} \tag{10}$$

which means that K is a homogeneous function of order one.

Theorem 8. *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*

$$w(as) = bw(s) \tag{11}$$

where $b = \alpha a$.

It is clear that whatever aspect of the real world we consider, sight, sound, touch, taste, smell, heat and cold, at each instant, their corresponding stimuli impact our senses numerous times. A stimulus S of magnitude s, is received as a similarity transformation as, $a > 0$ referred to as a dilation of s. It is a *stretching* if $a > 1$, and a *contraction* if $a < 1$. When relating response to a dilated stimulus of magnitude as to response to an unaltered stimulus whose magnitude is s, we have the proportionality relation between responding to an altered stimulus and to the original stimulus given in (11) as

$$\frac{w(as)}{w(s)} = b$$

We refer to equation (11) as: The Functional Equation of Ratio Scales. Because of its wider implications in science, we may call it: *The Fundamental Equation of Proportionality and Order* [10]. We could have written down this equation directly from what we know about responding to stimuli with the idea of exercising control for which there also has to be proportionality. This would be true for a composite stimulus consisting of many stimuli thus defining s function instead of a single variable.

If we substitute $s = a^u$ in (11) we have $w(a^{u+1}) - bw(a^u) = 0$ (I am grateful to my friend Janos Aczel the world leader in functional equations for his help in deriving (12) and (13) below):

Again if we write,

$$w(a^u) = b^u p(u)$$

we get:

$$p(u+1) - p(u) = 0$$

which is a periodic function of period one in the variable u (such as $\cos u/2\pi$). Note that if a and s are real, then so is u which may be negative even if a and s are both assumed to be positive.

If in the last equation $p(0)$ is not equal to 0, we can introduce $C = p(0)$ and $P(u) = p(u)/C$, we have for the general response function $w(s)$:

$$w(s) = Ce^{\log b \frac{\log s}{\log a}} P\left(\frac{\log s}{\log a}\right) \tag{12}$$

where P is also periodic of period 1 and $P(0) = 1$. Note that $C > 0$ only if $p(0)$ is positive. Otherwise, if $p(0) < 0$, $C < 0$.

Near zero, the exponential factor, which is equal to $s^{\log b/\log a}$, "slims" $w(s)$ if $\log b/\log a > 0$ and "spreads" $w(s)$ if $\log b/\log a < 0$. Because s is the magnitude of a stimulus and cannot be negative, we do not have a problem with complex variables here so long as both a and b are real and both positive. Our solution in the complex domain has the form:

$$w(z) = z^{\ln b/\ln a} P\left(\frac{\ln z}{\ln a}\right) \tag{13}$$

Here $P(u)$ with $u = \ln z/\ln a$, is an arbitrary multivalued periodic function in u of period 1. Even without the multivaluedness of P, the function $w(z)$ could be multivalued because $\ln b/\ln a$ is generally a complex number. If P is single-valued and $\ln b/\ln a$ turns out to be an integer or a rational number, then $w(z)$ is a single-valued or finitely multivalued function, respectively. This generally multivalued solution is obtained in a way analogous to the real case.

This Solution Leads to the Weber-Fechner Law

Note in (12) that the periodic function $P(u)$ is bounded and the negative exponential leads to an alternating series. Thus, to a first order approximation in series expansion, one obtains the Weber-Fechner law $a \log s + b$ for response to a stimulus s. We assume that $b = 0$, and hence the response belongs to a ratio scale. Recall that in making comparisons, we form the ratio of two ratio scale readings which is a dimensionless number and belongs to the Fundamental Scale of absolute numbers.

Stimuli received by the brain from nature are transformed to chemical and electrical neural activities that result in summation and synthesis. This is transformed to awareness of nature by converting the electrical synthesis (vibrations caused by a pattern) to a space-time representation. The way the brain goes back and forth from a pattern of stimuli to its electro-chemical synthesis and then to a representation of its response to that spacio-temporal pattern is by applying the Fourier transform to the stimulus and the inverse Fourier transform to form its response. What we have been doing so far is concerned with the inverse Fourier transform. We now need to take its inverse to develop expressions for the response.

We now show that the space-time Fourier transform of (13) is a combination of Dirac distributions. Our solution of Fredholm's equation here is given as the Fourier transform,

$$f(\omega) = \int_{\infty}^{+\infty} F(x) \mathrm{e}^{-2\pi i \omega x} \, \mathrm{d}x = C \mathrm{e}^{\beta\omega} P(w) \tag{14}$$

whose inverse transform is given by:

$$\frac{1}{2\pi} \log a \sum_{-\infty}^{\infty} a'_n \left[\frac{(2\pi n + \theta(b) - x)}{(\log a|b| + (2\pi n + \theta(b) - x))} i \right] \delta(2\pi n + \theta(b) - x) \tag{15}$$

where $\delta(2\pi n + \theta(b) - x)$ is the Dirac delta function which means that the way we respond to stimuli has the form of the delta firings as indeed it is known that our neurons behave. Neural responses are impulsive and hence the brain is a discrete firing system. This is supporting evidence in favor of the generalized comparisons model.

19 The Formation of Images and Sounds with Dirac Distributions

Complex valued functions with their imaginary part cannot be drawn as one does ordinary functions of three real variables. Nevertheless, one can make a plot of the modulus or absolute value of such a function. The basic assumption we made to represent the response to a sequence of individual stimuli is that all the layers in a network of neurons are identical, and each stimulus value is represented by the firing of a neuron in each layer. A shortcoming of this representation is that it is not invariant with respect to the order in which the stimuli are fed into the network. It is known in the case of vision that the eyes do not scan pictures symmetrically if they are not symmetric, and hence our representation must satisfy some order invariant principle. Taking into account this principle would allow us to represent images independently of the form in which stimuli are input into the network. For example, we recognize an image even if it is subjected to a rotation, or to some sort of deformation. Thus, the invariance principle must include affine and similarity transformations. This invariance would allow the network to recognize images even when they are not identical to the ones from which it recorded a given concept, e.g., a bird. The next step would be to use the network representation given here with additional conditions to uniquely represent patterns from images, sounds and perhaps other sources of stimuli such as smell. Our representation focuses on the real part of the magnitude rather than the phase of the Fourier transform. Tests have been made to see the effect of phase and of magnitude on the outcome of a representation of a complex valued function. There is much more blurring due to change in magnitude than there is to change in phase. Thus we focus on representing responses in terms of Dirac functions, sums of such functions, and on approximations to them without regard to the coefficients in (15).

Because finite linear combinations of the functions $\{t^{\alpha}\mathrm{e}^{-\beta t},\ \alpha, \beta \geq 0\}$ are dense in the space of bounded continuous functions $\mathcal{C}[0, b]$ we used them to represent images and sounds [17].

We created a 2-dimensional network of neurons consisting of layers. For illustrative purposes, we assumed that there is one layer of neurons corresponding to each of the stimulus values. Thus, if the list of stimuli consists of n numerical values, we created n layers with a specific number of neurons in each layer. Under the assumption that each numerical stimulus is represented by the firing of one and only one neuron, each layer of the network must also consist of n neurons with thresholds varying between the largest and the smallest values of the list of stimuli. We also assumed that the firing threshold of each neuron had the same width. Thus, if the perceptual range of a stimulus varies between two values θ_1 and θ_2, and each layer of the network has n neurons, then a neuron in the ith position of the layer will fire if the stimulus value falls between

$$\theta_1 + (i-1)\frac{\theta_2 - \theta_1}{n-1} \qquad \text{and} \qquad \theta_1 + i\frac{\theta_2 - \theta_1}{n-1}$$

19.1 Picture Experiment

In the graphics experiment the bird and rose pictures required 124 and 248 data points, respectively, whereas the sound experiment required 1000 times more data points. Once the (x, y) coordinates of the points were obtained, the x-coordinate was used to represent time and the y-coordinate to represent response to a stimulus. The numerical values associated with the drawings in Figures 15 and 16 were tabulated and the numbers provided the input to the neurons in the networks built to represent the bird and the rose.

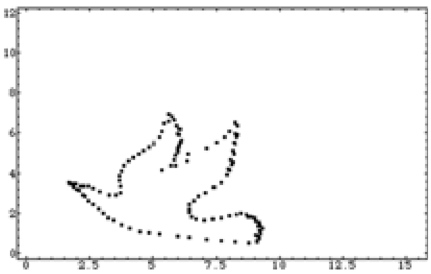

Figure 15. Bird

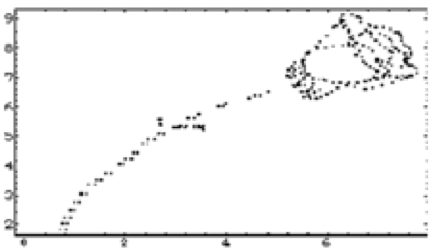

Figure 16. Rose

19.2 Sound Experiment

In the sound experiment we first recorded with the aid of Mathematica the first few seconds of Haydn's symphony no. 102 in B-flat major and Mozart's symphony no. 40 in G minor. The result is a set of numerical amplitudes between -1 and 1. Each of these amplitudes was used to make neurons fire when the amplitude falls within a prescribed threshold range. Under the assumption that each neuron fires in response to one stimulus, we would need the same number of neurons as the sample size, i.e., $117,247$ in Haydn's symphony and $144,532$ in Mozart's symphony. Our objective was to approximate the amplitude using one neuron for each amplitude value, and then use the resulting values in Mathematica to play back the music. A small sample of the numerical data for Mozart's symphony is displayed in Figure 17.

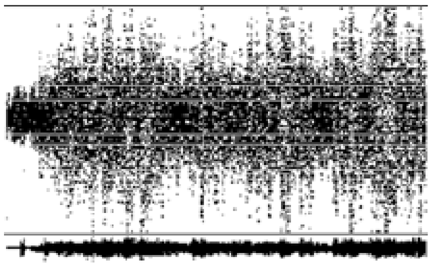

Figure 17. Mozart's Symphony No. 40

This task is computationally demanding even for such simple geometric figures as the bird and the flower shown in Figures 15 and 16. For example, for the bird picture, the stimuli list consists of 124 values, and we would need $124^2 = 15376$ neurons, arranged in 124 layers of 124 neurons each. The network and the data sampled to form the picture given in Figure 15, were used to create a 124 by 124 network of neurons consisting of 124 layers with 124 neurons in each layer. Each dot in the figures is generated by the firing of a neuron in response to a stimulus falling within the neuron's lower and upper thresholds.

19.3 Inverse Squares Laws

Earlier we took the Fourier transform of the complex valued solution to our fundamental equation. Let us examine what happens by doing the same to the real valued solution [10]. Our solution is a product of two factors, the inverse transform can be obtained as the convolution of two functions, the inverse Fourier transform of each of which corresponds to that of just one of the factors.

Now the inverse Fourier transform of $\mathrm{e}^{-\beta u}$ is given by

$$\frac{\sqrt{2/\pi}\,\beta}{\beta^2 + \xi^2}$$

Also from the theory of Fourier series, we know that we can write

$$P(u) = \sum_{k=-\infty}^{\infty} \alpha_k \mathrm{e}^{2\pi i k u}$$

whose inverse Fourier transform involves the familiar Dirac delta function and is given by:

$$\sum_{-\infty}^{\infty} \alpha_k \delta(\xi - 2\pi k)$$

The product of the transforms of the two functions is equal to the Fourier transform of the convolution of the two functions themselves that we just obtained by taking their individual inverse transforms. We have, to within a multiplicative constant:

$$\int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \alpha_k \delta(\xi - 2\pi k) \frac{\beta}{\beta^2 + (x-\xi)^2} \,\mathrm{d}\xi = \sum_{k=-\infty}^{\infty} \alpha_k \frac{\beta}{\beta^2 + (x - 2\pi k)^2}$$

The inverse Fourier transform of $w(u) = C\mathrm{e}^{-\beta u} \cos \frac{u}{2\pi}$, $\beta > 0$ is given by:

$$C\left[\frac{1}{\beta^2 + (x+2\pi)^2} + \frac{1}{\beta^2 + (x-2\pi)^2}\right]$$

When the constants in the denominator are small relative to x we have C_1/x^2 which we believe is why in optics, gravitation (Newton) and electricity (Coulomb) our response to the forces is according to inverse square laws. This is the same law of nature in which an object responding to a force field must decide to follow that law by comparing infinitesimal successive states through which it passes. If the stimulus is constant, the exponential factor in the general response solution is constant, and the solution would be periodic of period one. When the distance x is very small, the result varies inversely with the parameter $\beta > 0$. It appears that the inverse square law is not strictly an inverse square and its use may sometimes lead to difficulty as found with the orbit of Mercury.

20 Generalization to Multiple Stimuli

With multiple stimuli we have the functional equation

$$w(a_1 z_1, \ldots, a_n z_n) = b w(z_1, \ldots, z_n)$$

whose solution for the real variable case with $b > 0$, $a_k > 0$, and $s_k > 0$, is

$$w(s_1, \ldots, s_n) = b^{\sum_{k=1}^{n} (\log s_k / \log a_k)/k} P\left(\frac{\log s_1}{\log a_1}, \ldots, \frac{\log s_n}{\log a_n}\right), \qquad k = 1, \ldots, n,$$

where P is an arbitrary periodic function of period one of n variables, that is,

$$P(u_1 + 1, \ldots, u_n + 1) = P(u_1, \ldots, u_n)$$

Solution of the functional equation in the complex domain was communicated to me by my friend the mathematician Karol Baron of Katowice, Poland.

$$w(z_1, \ldots, z_N) = z_1^{k_1} \cdots z_N^{k_N} G\left(|z_1|, \ldots, |z_M|, \frac{\log|z_{M+1}|}{\log|a_{M+1}|}, \ldots, \frac{\log|z_N|}{\log|a_N|}, \arg z_1, \ldots, \arg z_N\right)$$

where G is a doubly periodic function with periods

$$(0, \underset{M}{\ldots}, 0, 1, \underset{M-N}{\ldots}, 1, \arg a_1, \ldots, \arg a_N) \quad \text{and} \quad (0, \ldots, 0, \underset{(n+l)}{2\pi}, 0, \ldots, 0)$$

for $l \in \{1, \ldots, N\}$. It is known that a single-valued analytic function cannot possess more than two linearly independent periods.

For a continuum number of stimuli let $K(\mathbf{X}, \mathbf{Y})$ be a compact reciprocal kernel i.e. $K(x, y)K(y, x) = 1$, for all $x \in \mathbf{X}$ and $y \in \mathbf{Y}$, where $\mathbf{X}$ and $\mathbf{Y}$ are compact subsets of the reals. We have the equation $w(\mathbf{X} = \lambda \int_\Omega K(\mathbf{X}; \mathbf{Y}) w(\mathbf{Y}) \,\mathrm{d}Y$ for which by analogy with the necessary condition for the solvability of Fredholm's equation (4) and ensuing discussion, the following operator functional equation must be satisfied in order for the above equation to have a solution: $w[af(x)] = b[f(x)]$. This operator linear functional equation has been formally and completely solved at Nantes France by Nicole Brillouet-Belluot [5]. Because of space limitations here, its solution and interpretation will be described in some detail in my forthcoming book concerned with the generalization of pairwise comparisons to the continuous case.

21 Criticisms of the AHP

Essentially, there have been five types of criticisms of the AHP all of which have been addressed in the literature [19]. One, already mentioned in the paper, is the concern with illegitimate changes in the ranks of the alternatives, called rank reversal, upon changing the structure of the decision. It was believed that rank reversal is legitimate only when criteria or priorities of criteria or changes in judgments are made. Rank reversals were shown by critics to occur when using comparisons and relative measurement that is *essential* in prioritizing criteria and also alternatives on intangible criteria in two ways: First, when new alternatives are added or old ones deleted; and second, when new criteria are added or old ones deleted with the caveat that the priorities of the alternatives would be tied under these criteria and hence argued that the criteria should be irrelevant when ranking the alternatives. Rank reversals that followed such structural changes were attributed to the use of relative measurement and normalization. Rating alternatives one at a time with respect to the criteria using the ideal mode, always preserves rank. Also, the ideal mode is used with paired comparisons to preserve rank. But rank can and should reverse under more general conditions than had previously been recognized as in introducing copies or near copies of alternatives and criteria turn out not to always be so strictly independent among themselves and from the alternatives. The second concern is about inconsistent judgments and their effect on aggregating such judgments or on deriving priorities from them. A modicum of intransitivity and numerical inconsistency, usually not considered or thought to be permissible in other theories, is permissible in the AHP so that decisions can be treated realistically rather than axiomatically truncated. A condition that may not hold with inconsistent judgments is Pareto optimality. Pareto optimality is an ordinal relation which demands of a method used to aggregate judgments of individuals in a group to a representative collective judgment for that group that when all individuals in the group prefer A to B then the group judgment must prefer A to B. Because in the AHP judgments are not ordinal, it is possible to aggregate the individual judgments into a representative group judgment with or without Pareto optimality. Another condition also inherited from expected utility theory has to do with a relation called Condition of Order Preservation (COP): For all alternatives x_1, x_2, x_3, x_4, such that x_1 dominates x_2 and x_3 dominates x_4, if the evaluator's judgments indicate the extent to which x_1 dominates x_2 is greater than the extent to which x_3 dominates x_4, then the vector of priorities w should be such that, not only $w(x_1) > w(x_2)$ and $w(x_3) > w(x_4)$ (preservation of order of preference) but also that $\frac{w(x_1)}{w(x_2)} > \frac{w(x_3)}{w(x_4)}$ (preservation of order of intensity of preference). This condition holds when judgments are consistent but may or may not hold when they are inconsistent. It is axiomatically imposed, sacrificing the original intent of the AHP process to derive priorities that match the reality represented by the judgments without forcing consistency. The third criticism has to do with attempts to preserve rank from irrelevant alternatives by combining the comparison judgments of a single individual using the geometric mean (logarithmic least squares) to derive priorities and also combining the derived priorities on different criteria by using multiplicative weighting synthesis. The fourth criticism has to do with people trying to change the fundamental scale despite the fact that it is theoretically derived and tested by comparing it with numerous other scales on a multiplicity of examples for which the answer was known. The fifth and final criticism has to do with whether or not the pairwise comparisons axioms are behavioral and spontaneous in nature to provide judgments.

Interestingly, the AHP/ANP provides a way to make complex decisions in the most general structures encountered in real life. It makes it possible to derive priorities for all the factors in such structures and synthesize them for an overall outcome, as no other method can because one can build scales for tangibles and intangibles, yet we know little about criticisms of framing and validating problems within such a wide perspective that includes structures, not only for dependence and feedback, but also for benefits, opportunities, costs and risks analyzed separately and then synthesized for the final outcome or in conflict resolution with or without a moderating negotiator

22 Conclusions

The AHP/ANP is a useful way to deal with complex decisions that involve dependence and feedback analyzed in the context of benefits, opportunities, costs and risks. It has been applied literally to hundreds of examples both real and hypothetical. What is important in decision making is to produce answers that are valid in practice. The AHP/ANP has also been validated in many examples. People often argue that judgment is subjective and that one should not expect the outcome to correspond to objective data. But that puts one in the framework of garbage in garbage out without the assurance of the long term validity of the outcome. In addition, most other approaches to decision making are normative. They say, “If you are rational you do as I say.” But what people imagine is best to do and what conditions their decisions face after they are made can be very far apart from what can happen in the real world. That is why the framework of the AHP/ANP is descriptive as in science rather than normative and prescriptive giving license to unrealistic assumptions like insisting on transitivity of preferences when we know that people are intransitive and inconsistent. The AHP/ANP produce outcomes that are best not simply according to the decision maker’s values, but also to the external risks and hazards faced by the decision.

It is unfortunate that there are people who use fuzzy sets without proof to alter the AHP when it is known that fuzzy applications to decision making have been ranked as the worst among all methods. Because of the Weber-Fechner observation, the fundamental scale used in the AHP/ANP to represent judgments is already fuzzy. To fuzzify it further does not improve the outcome as we have shown through numerous examples [15]. The intention of fuzzy seems to be to perturb the judgments in the AHP. It is already known in mathematics that perturbing the entries of a matrix perturbs the eigenvector by a small amount but not necessarily in a more valid direction.

The SuperDecisions software used to analyze complex decisions is named after the supermatrix. It can be downloaded free from creativedecisions.net and is available on the internet along with a manual and with numerous applications to enable the reader to apply it to his/her decision. Alternatively, go to www.superdecisions.com/ saaty and download the SuperDecisions software. The installation file is the `.exe` file in the software folder. The serial number is located in the `.doc` file that is in the same folder. The important thing may be not the software but the models which are in a separate folder called models. The military are constantly involved in making complex decisions and appear to like using the ANP and investing in its development. Why do we do all this with so much effort? Because we believe strongly in the creativity of the human race and hope that our world will become increasingly more rational in making its decisions and in resolving its conflicts.

Since 1976, a very large number of case studies using the Analytic Hierarchy Process in conflict resolution, on which I co-authored a book, have been developed. The hope in this work was that, by teaching people to analyze their conflicts and to lay bare the structure, reason would prevail and antagonist would be motivated to come to a reasonable and just solution.

But, in a conflict, particularly one of long duration, reason rarely prevails. Instead, positions become entrenched and people seek not only to satisfy their own needs but also to punish their opponents for having opposed them -or, at least, to pay a price for their opposition.

New light may often be shed on a conflict. Experience has shown that a conflict may be alleviated by the use of a mediator who introduces “bargaining chips” into the negotiations. If A and B are the participants, A will look at this particular item and evaluate not only the incremental benefits to be received but also the cost to the opponent of providing the concession: The greater each value, the greater the gain.

Thus A's gain for a given item may be described as the product of A's benefits and B's cost (as A perceives them). We have the following ratios for the two parties A and B:

(as perceived by A)

$$A\text{'s ratio} \qquad \frac{\text{gain to } A}{A\text{'s perception of gain to } B} = \frac{\sum A\text{'s benefits} \times B\text{'s costs}}{\sum B\text{'s benefits} \times A\text{'s costs}} = \frac{\text{gain to } A}{\text{loss to } A}$$

where $\sum$ is the sum taken over all concessions by B in the numerator and by A in the denominator. A's *perceived ratio* for B is the reciprocal of the above:

(as perceived by B)

$$B\text{'s ratio} \qquad \frac{\text{gain to } B}{B\text{'s perception of gain to } A} = \frac{\sum B\text{'s benefits} \times A\text{'s costs}}{\sum A\text{'s benefits} \times B\text{'s costs}} = \frac{\text{gain to } B}{\text{loss to } B}$$

If both A and B perceive costs and benefits in the same way, these ratios would be reciprocal of each other. This almost never happens, however.

Obviously, each party would like its ratio to be as high as possible, If A's ratio for some package is less than 1, then A will perceive B's ratio as being greater than 1 and will feel that it has not been treated fairly. The aim must be to find single concessions and groups of concessions where each party perceive its own ratio to be greater than 1. This requires skilled mediation.

It is wise to involve the parties and to gain their trust so that the necessary estimates may be made. If this is not possible, then the analyst must use his or her best judgements for these.

The ideas were applied to the conflict in South Africa in 1986 in a study paid for at the time by the government of South Africa and whose conclusions appear to have been followed fairly accurately. It is summarized as Chapter 11 of my book on Conflict Resolution. In addition to applying the process to analyze the conflict in Northern Ireland by my colleague Joyce Alexander since the 1970's, I have been involved in work and publication of material that deals with the conflict between Israel and the Palestinians in the Middle East with the participation of distinguished individuals from that part of the world. What one learns is that the AHP/ANP provides an opportunity to map out and quantify all the factors in a conflict as desired by the parties and subject them to sensitivity analysis in the hope of finding an acceptable solution without resort to uncivilized behaviour and to violence. One derives a salutary degree of satisfaction from doing this kind of mathematics that may help save the lives of many people and move us into a new era of useful negotiation before commitment to take violent action.

Acknowledgement. My thanks go to Luis Vargas, Rozann Whitaker, Orrin Cooper and José Manuel Magallanes for their editorial help in preparing the manuscript.

References

[1] ACZEL, J. AND ALSINA, C., (1986). On synthesis of judgments, *Socio-Economic Planning Sciences*, **20**, 333–339.

[2] ACZEL, J. AND SAATY, T. L., (1983). Procedures for synthesising ratio judgments, *Journal of Mathematical Psychology*, **27**, 93–102.

[3] ARROW, K. J., (1963). *Social Choice and Individual Values*, 2nd ed., Wiley, New York.

[4] BAUER, R. A., COLLAR, E. AND TANG, V., (1992). *The Silverlake Project*, Oxford University Press, New York.

[5] BRILLOUET-BELLUOT, N., (1999). On a Simple Linear Functional Equation on Normed Linear Spaces, Ecole Centrale de Nantes, F-44 072 Nantes - cedex 03, France, October 1999.

[6] DAVIS, P. J. AND HERSH, R., (1986). *Descartes Dream*, Harcourt Brace and Jovanovich, New York.

[7] FECHNER, G., (1966). *Elements of Psychophysics*, Adler, H. E. (Trans.), Vol. **2**, Holt, Rinehart and Winston, New York. See also Batschelet, S., *Introduction to Mathematics for Life Scientists*, Springer, 1971.

[8] LEBESGUE, H., (1928). *Leçons sur l'integration*, 2nd ed., Gauthier-Villars, Paris.

[9] MACKAY, A. F., (1980). *Arrow's Theorem : The Paradox of Social Choice - A Case Study in the Philosophy of Economics*, Yale University Press, New Haven.

[10] SAATY, T. L., (2000). *The Brain, Unraveling the Mystery of How it Works : The Neural Network Process*, RWS Publications, 4922 Ellsworth Avenue, Pittsburgh, PA 15213.

[11] SAATY, T. L., (2005). *Theory and Applications of the Analytic Network Process*, Pittsburgh, PA : RWS Publications, 4922 Ellsworth Avenue, Pittsburgh, PA 15213.

[12] SAATY, T. L., (2006). *Fundamentals of Decision Making ; the Analytic Hierarchy Process*, Pittsburgh, PA : RWS 3. Publications, 4922 Ellsworth Avenue, Pittsburgh, PA 15213.

[13] SAATY, T. L., AND OZDEMIR, M., (2005). *The Encyclicon*, RWS Publications, 4922 Ellsworth Avenue, Pittsburgh, PA 15213

[14] SAATY, T. L. AND PENIWATI, K., (2008). *Group Decision Making*, RWS Publications, 4922, Ellsworth Avenue, Pittsburgh, PA 15213.

[15] SAATY, T. L. AND TRAN, L. T., (2007). On the invalidity of fuzzifying numerical judgments in the Analytic Hierarchy Process, *Mathematical and Computer Modelling*, **46**, 7-8, 962–975.

[16] SAATY, T. L. AND VARGAS, L. G., (1993). Experiments on Rank Preservation and Reversal in Relative Measurement, *Math. Comput. Modelling*, **17**, 4/5, 13–18.

[17] SAATY, T. L. AND VARGAS, L. G., (1997). Implementing Neural Firing : Towards a New Technology, *Mathl. Comput. Modelling*, **26**, 4, 113–124.

[18] SAATY, T. L. AND VARGAS, L. G., (2008). The Possibility of Group Choice : Pairwise Comparisons and Merging Functions, (in publication).

[19] SAATY, T. L., VARGAS, L. G. AND WHITAKER, R., (2008). Addressing Criticisms of the Analytic Hierarchy Process, (in publication).

[20] WILKINSON, J. H., (1965). *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford.

Thomas L Saaty
saaty@katz.pitt.edu