Determining Certainty Factors with the Analytic Hierarchy Process

Ernest H. Forman
Professor of Mangement Science
George Washington University
Washington, D.C. 20052

Abstract

Certainty factors are intended measure the certainty of expert system rules. Since certainty factors represent a *change* in the probability of a hypothesis, given additional information about an event, a rule's certainty factor depends on the difference between posterior and prior probabilities. Developers of the MYCIN expert system (originators of the certainty factor concept) abandoned Bayes' Theorem and the P-function because they felt there were large areas of expert knowledge and intuition that, although amenable in theory to the frequency analysis of statistical probability, defied rigorous analysis, in part, because experts resisted expressing their reasoning process in coherent probabilistic terms. The Analytic Hierarchy Process (AHP) facilitates the practical acquisition of experts' knowledge and intuition in a way that produces ratio scale likelihoods with a theoretical basis that conforms to Bayes theorem and the P-function. We show how AHP can be used to develop prior and posterior probabilities and how these probabilities can be used to calculate certainty factors for expert system rules.

Certainty factors (CF's) in expert systems are intended to reflect a measure of expert system rule certainty. Originally developed for the MYCIN expert system[1], certainty factors represent a *change* in the probability of a hypothesis, given additional information about an event. Bayes rule has long been used to calculate revised probabilities, e.g. the probability of a hypothesis h given an event e, or P(h/e), given the prior probability of h, P(h), and conditional probabilities of the possible events given h. In order to reduce the amount of

---

[1]Bruce G. Buchanan and Edward H. Shortliffe, editors, *Rule-Based Expert. Systems*, Reading, MA: Addison-Wesley, 1984.

information required for the calculation of revised probabilities, certainty factors for rules were defined in terms of the change from the prior to a revised probability, i.e.,

$$P(h/e) - P(h),$$

which can of course be positive or negative. If it is positive, then the difference is said to increase the *belief* in the rule. If it is negative, then the difference is said to increase the *disbelief* in the rule.

The originators of CF's expressed the certainty factor as a *percentage change*. If the difference $P(h/e) - P(h)$ is positive, the percentage change (increase in belief) is proportional to the maximum possible increase, i.e, $(1 - P(h))$. If the difference $P(h/e) - P(h)$ is negative, the percentage change is proportional to the maximum amount it could have decreased, that is, $P(h)$.

When the difference $P(h/e) - P(h)$ for a rule is positive, the CF is also called the measure of belief, MB, and the measure of disbelief for the rule is zero. When the difference is negative, the absolute value of the CF is called the measure of disbelief, MD, and the measure of belief for the rule is zero.

In summary:

If $P(h/e) - P(h) > 0$ then $CF = (P(h/e)-P(h))/(1-P(h))$ and is > 0.

$$\text{and } MB = CF \text{ and } MD = 0.$$

If $P(h/e) - P(h) < 0$ then $CF = (P(h/e)-P(h))/(P(h))$ and is < 0.

$$\text{and } MD = -CF \text{ and } MB = 0.$$

284

While the above definitions are theoretically sound, there has been much debate about the theoretical foundation of certainty factors as they are derived and applied in practice. In an internal memo to the MYCIN researchers in 1976, Shortliffe "outlined five alternative CF models and argued for careful consideration of one that would require the use of *a priori* probabilities of hypotheses in addition to conventional CF's on rules. The proposed model was never implemented, however, partly due to time constraints but largely because MYCIN's decision making performance was proving to be excellent despite the theoretical limitations of CF's."[2]

A major attraction of asking experts for certainty factors is that it reduces the task to assigning only one number to each rule, rather than having to supply both prior and posterior probabilities. Buchanan and Shortliffe write "Some of us felt that a one-number calculus was preferable in this domain to a more theoretically sound calculus that requires experts to supply estimates of two or more quantities per rule."[3]

The objective of creating a workable system may have been adequate reason for the MYCIN researchers to abandon a theoretically sound calculus, and the fact that the resulting system performed so well is to their credit. However, when a system lacks a theoretical foundation, one is never sure of where one is headed.[4] The MYCIN researchers had to "fine tune" their system in order to make it work, and there were many questions that

---

[2]*Ibid.*, p 214.

[3]*Ibid.*, p 214.

[4]Edward Demming, in a keynote speech to the international meeting of the Institute of Management Science in Australia in 1986 repeated numerous times, that without a theoretical foundation, "Off you go to the milky way."

arose. Buchannan and Shortliffe write that "Victor was developing the meningitis rule set at the time and was having frequent problems deciding what CF's to assign to individual rules and how to anticipate the ramifications of any decisions made."[5] Many of the doubts about certainty factors arose in electronic exchanges and have been documented by Buchanan and Shortliffe.[6] The following excerpts are some examples:

> From Carli Scott: "Since Victor has defined guidelines on deciding how strong the evidence must be in order for a rule to be assigned a certain CF, and since he has tested these guidelines within the framework of MYCIN's combining functions, he believes that it all works as it should. Furthermore, he believes that he can define these "points of reference" so that future medical people can add rules, using the same guideline that Victor has used, and they should fit into the system and work fine with his rules."[7]

> From Bruce Buchanan "..my reservations with the meningitis system stem from my uneasiness with the CF model, which we all know needs improving. I don't want Victor to become dependent on a particular mechanism for combining CF's because we hope the mechanism will be improved soon. I have no doubt that the rules work well now, and I don't disagree at all with the need for firm reference points for the CF's. ... "I foresee no difficulty in mapping the CF's from existing rules (meningitis as well as bacteremia) into whatever numbers are appropriate for a new CF model when we have one- with firm reference points if at all possible."[8]

The Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP) is an attractive methodology for dealing with uncertainty in expert systems. AHP was developed by Thomas Saaty [5] and enables

---

[5] *Op. cit.,* p 221.

[6] *Ibid.,* pp 221-232.

[7] *Ibid.,* p 222.

[8] *Ibid.,* p 223.

decision makers to structure a complex decision problem in the form of a hierarchy. Both objective as well as subjective factors can be accommodated. Each factor and alternative can be identified and evaluated with respect to other related factors. The ability to structure a complex problem and then focus attention on individual components amplifies one's decision making capabilities. The capacity for making judgments is expanded beyond the limits imposed by what psychologists call the limited channel capacity and short term memory of the human mind. The AHP methodology goes beyond conventional decision analysis techniques by not requiring numerical guesses. Relative judgments are used, rather than absolute judgments. Not only do relative judgments provide a more accurate representation of the decision makers thoughts, but relative judgments can easily accommodate subjectivity when appropriate. Judgments can be made entirely in a verbal mode; no numerical guesses are required.

AHP overcomes the limitations that led the MYCIN researchers away from a theoretical basis for addressing uncertainty. According to Buchanan and Shortliffe,

> "We have had to abandon Bayes' Theorem and the P-function simply because there are large areas of expert knowledge and intuition that, although amenable in theory to the frequency analysis of statistical probability, defy rigorous analysis because of insufficient data and, in a practical sense, because experts resist expressing their reasoning process in coherent probabilistic terms."[9]

AHP facilitates the practical acquisition of experts' knowledge and intuition in a way that produces ratio scale likelihoods or priorities with a theoretical basis that conforms to Bayes theorem and the P-function. AHP does this with a process of pairwise relative verbal

---

[9]*Ibid.*, p 245.

comparisons.[10] The process is easy to use because the pairwise relative verbal comparison process is a natural way of expressing judgments. Accurate results can be derived from "fuzzy" verbal judgments because of the way redundancy is incorporated when eliciting judgments. This has been demonstrated with an experiment (described in an Appendix) consisting of exercises in which an individual or a group of individuals derived accurate ratio scale priorities from "fuzzy" verbal judgments about the relative sizes of geometric shapes.

Derivation of Prior Probability Distribution

Priorities (or likelihoods) derived from an expert's pairwise verbal judgments should be accurate representations of the experts prior probability distribution for the events in question in the same way that the relative sizes of the five shapes were derived from fuzzy verbal judgments in the experiment described in the Appendix. Suppose we identified a set of mutually exclusive, collectively exhaustive events h1, h2, ...h4 (where h4 might consist of all possibilities other than h1 ....h3) and then elicit an expert's judgments about their relative likelihood using pairwise verbal judgments such as the one illustrated in Figure 1. A set of such judgments and the resulting likelihoods for the prior probability distribution are shown in Figure 2.

---

[10]The comparisons can also be made numerically or graphically.

```
                    GOAL: Deriving Certainty Factors
                            With respect to
                               GOAL

H3
    is MODERATELY more LIKELY than
H1
```

```
        EXTREME-----------

        VERY STRONG-------

        STRONG------------

        MODERATE----------  <--

        EQUAL-------------
```
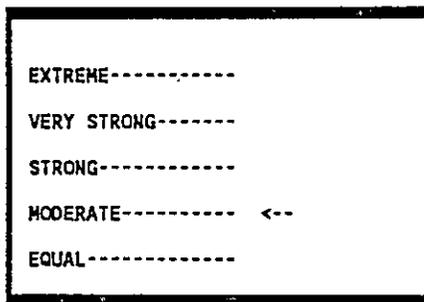
Figure 1 -- Verbal Judgment for relative Likelihood of H1 vs H3

```
           JUDGMENTS AND PRIORITIES WITH RESPECT TO
                            GOAL

        H1        H2        H3        H4
   H1            (2.0)     (3.0)     1.0
   H2                       2.0      3.0
   H3                                2.0
   H4
```

```
Matrix entry indicates that ROW element is ___
  1 EQUALLY   3 MODERATELY   5 STRONGLY   7 VERY STRONGLY   9 EXTREMELY
more LIKELY than COLUMN element
    unless enclosed in parenthesis.
```

```
0.144
H1   ████████████████

0.419
H2   ████████████████████████████████████████████

0.297
H3   ████████████████████████████████

0.139
H4   ██████████████
```
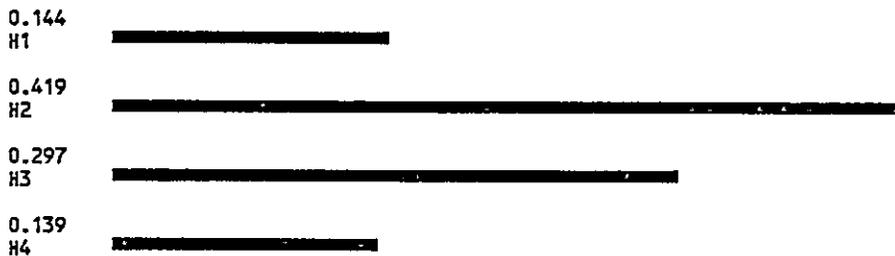
```
             INCONSISTENCY RATIO = 0.039
```

Figure 2 -- Prior Probability Distribution

Next, we can elicit a similar set of judgments to derive the conditional probability distributions as shown in Figures 3 and 4 for the conditional distribution given H1. Then conditional probability distributions given H2 and H3 are derived in a similar fashion.

```
                GOAL: Deriving Certainty Factors
                        ·With respect to
                          GOAL > H1

E2
     is MODERATELY more LIKELY than
E1
```

```
┌──────────────────────────────────────┐
│                                        │
│    EXTREME-----------                  │
│                                        │
│    VERY STRONG-------                  │
│                                        │
│    STRONG-----------                   │
│                                        │
│    MODERATE---------   <--             │
│                                        │
│    EQUAL-------------                  │
│                                        │
└──────────────────────────────────────┘
```
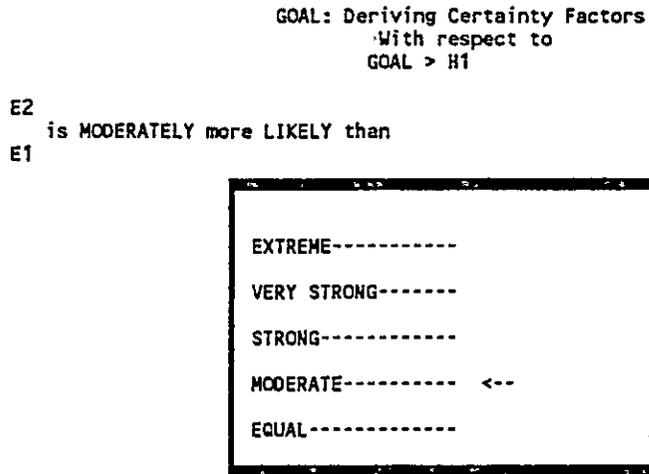
Figure 3 -- Verbal Judgment for Likelihood of E1 vs E2, Given H1

```
                JUDGMENTS AND PRIORITIES WITH RESPECT TO
                              GOAL > H1

              E1        E2        E3
        E1             (3.0)      2.0
        E2                        3.0
        E3

Matrix entry indicates that ROW element is ___
   1 EQUALLY   3 MODERATELY   5 STRONGLY   7 VERY STRONGLY   9 EXTREMELY
more LIKELY than COLUMN element
     unless enclosed in parenthesis.
```

```
0.249
E1      ████████████████████████

0.594
E2      ██████████████████████████████████████████████

0.157
E3      ████████████
```

INCONSISTENCY RATIO = 0.051
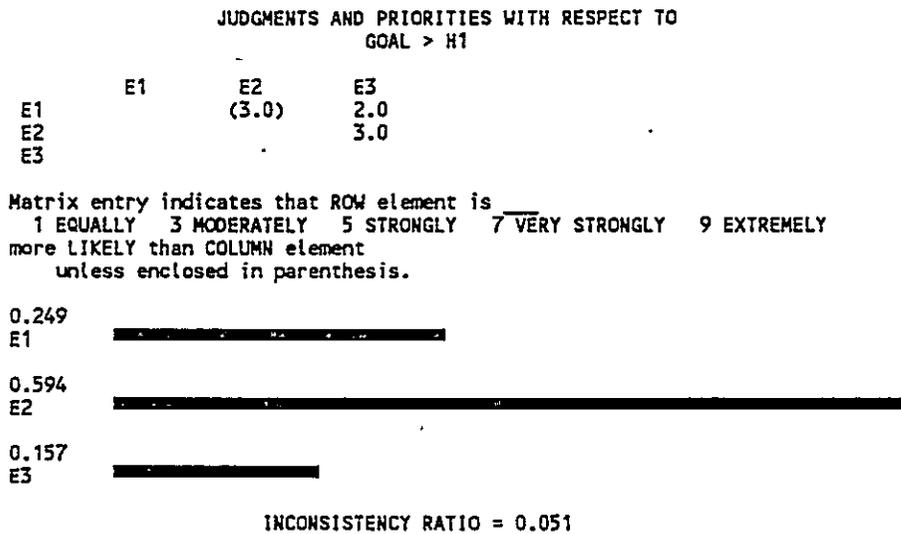
Figure 4 -- Conditional Probability Distribution, Given H1

The prior and conditional probabilities are shown as "local" priorities in Figure 5. The "global" priorities in Figure 5 are the priorities relative to the goal, and are obtained by multiplying a nodes local priority by the global priority of its parent. Therefore, the local priorities for E1, E2, and E3 represent the conditional probabilities, whereas the global priorities represent the joint probability of the Event and its Parent both taking place. For example, the probability of E1 given H1 is .249, while the joint probability of E1 and H1 is .036. The marginal probability of E1, .279, can be found by adding up the global priority of E1 under each of H1, H2, H3, and H4 (similarly for the marginal probability of E2 and E3.) This is achieved by "synthesizing" the AHP model as shown in Figure 6.

Deriving Certainty Factors

```
                          GOAL
                        L 1.000
                        G 1.000
                           |
        ┌──────────────┬───┴──────┬──────────────┐
       H1             H2         H3              H4
     L 0.144        L 0.419    L 0.297         L 0.139
     G 0.144        G 0.419    G 0.297         G 0.139

     -E1            -E1        -E1             -E1
      L 0.249        L 0.192    L 0.493         L 0.117
      G 0.036        G 0.080    G 0.147         G 0.016
     -E2            -E2        -E2             -E2
      L 0.594        L 0.634    L 0.311         L 0.268
      G 0.086        G 0.266    G 0.092         G 0.037
     -E3            -E3        -E3             -E3
      L 0.157        L 0.174    L 0.196         L 0.614
      G 0.023        G 0.073    G 0.058         G 0.086
```

L   --- LOCAL PRIORITY: PRIORITY RELATIVE TO PARENT
G   --- GLOBAL PRIORITY: PRIORITY RELATIVE TO GOAL

Figure 5 -- Local and Global Priorities

```
Details for Synthesis of Leaf Nodes with respect to   GOAL
     LEVEL 1          LEVEL 2          LEVEL 3          LEVEL 4          LEVEL 5
     -------          -------          -------          -------          -------

   H1       =0.144
   .                  E1       =0.036
   .                  E2       =0.086
   .                  E3       =0.023
   H2       =0.419
   .                  E1       =0.080
   .                  E2       =0.266
   .                  E3       =0.073
   H3       =0.297
   .                  E1       =0.147
   .                  E2       =0.092
   .                  E3       =0.058
   H4       =0.139
   .                  E1       =0.016
   .                  E2       =0.037
   .                  E3       =0.086

   E1       0.279
   E2       0.481
   E3       0.240
```

Figure 6 -- Marginal Probabilities of E1, E2, and E3

The above calculations and computer output were produced with the Expert Choice [3] implementation of AHP. By using the Expert Choice spreadsheet link, the above priorities can be routed to a spreadsheet for the calculation of the posterior probabilities and certainty factors. These are shown in Figure 7.

	P(H)		P(E and H)	P(H/E)	(P(H/E)-P(H)/(P(H)) or ------> (P(H/E)-P(H)/(1-P(H))	CF
H1	0.144					
		E1	0.036	0.129		-0.107
		E2	0.086	0.178		0.039
		E3	0.023	0.095		-0.344
H2	0.419					
		E1	0.080	0.288		-0.313
		E2	0.266	0.552		0.229
		E3	0.073	0.305		-0.272
H3	0.297					
		E1	0.147	0.525		0.324
		E2	0.092	0.192		-0.354
		E3	0.058	0.243		-0.183
H4	0.139					
		E1	0.016	0.059		-0.580
		E2	0.037	0.078		-0.442
		E3	0.086	0.358		0.253

P(E) or Marginals
E1   0.279
E2   0.481
E3   0.240

Figure 7 -- Calculation of Posterior Probabilities
and Certainty Factors

Conclusion

The Analytic Hierarchy Process (AHP) facilitates the practical acquisition of experts'

knowledge and intuition in a way that produces ratio likelihoods with a theoretical basis that

conforms to Bayes theorem and the P-function. We have shown how AHP can be used to

develop prior and posterior probabilities and how these can be used to calculate certainty

factors for expert system rules. We believe that pairwise verbal comparisons are a more

meaningful way for experts to express their measures of belief about conditions of

uncertainty · rathern than making absolute statements about prior and conditional

probabilities or about changes in probabilities given the acquisition of additional

information. While we expect that the elicitation of certainty factors using the approach described in this paper will be superior to the conventional way of "assigning" certainty factors, further research is required to verify this expectation.

Appendix

Deriving Ratio Priorities from Fuzzy Verbal Judgments

A wide variety of people participated in exercises to illustrate how accurate ratio level priorities can be derived from fuzzy verbal judgments.. In addition to participants from the business, government and educational communities in the United States, study participants also included Japanese business executives, Soviet scientists, and Chinese city planners. The subjects were given the following instructions:

> Suppose you were allocating funds for environmental quality purposes and wanted to determine the relative need for funds for clean air, clean water, noise reduction, industrial dumps, and acid rain. As an analogy, suppose your insight about the relative needs coincide with the relative sizes (areas) of the five objects shown in Figure 8. Although you *could* look at these objects and estimate their relative sizes numerically, the experiment is designed to show how we can derive accurate ratio scale priorities for qualitative factors from "fuzzy"verbal judgments. The analogy with the geometric shapes is necessary in order to measure the accuracy of the priorities derived from your "fuzzy" verbal judgments.
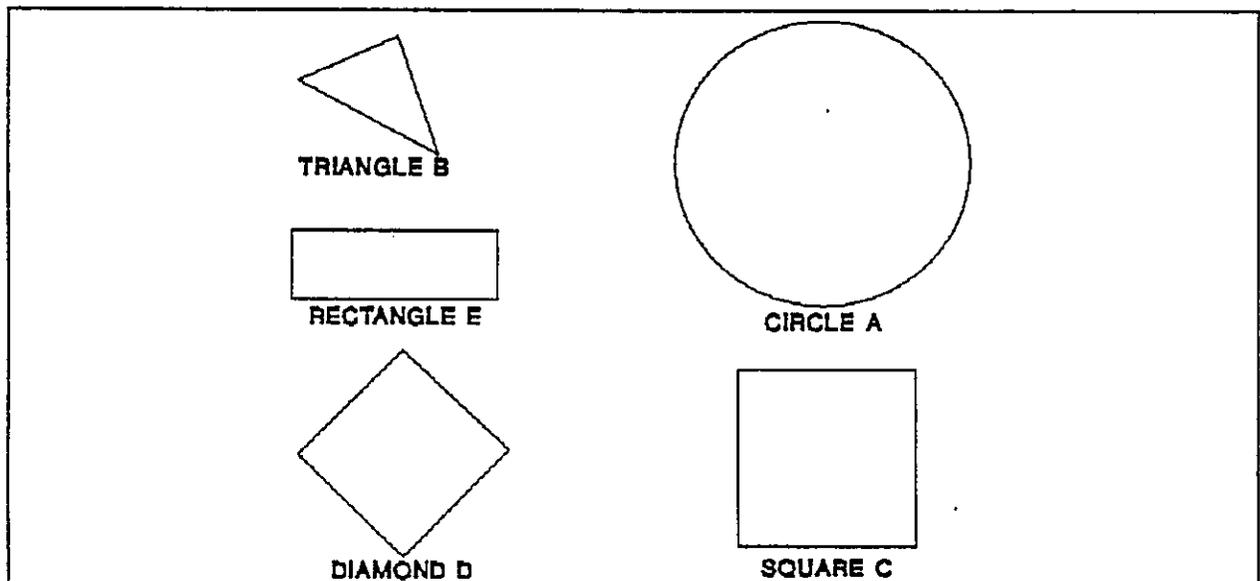

Figure 8 - Estimating Relative Areas Using Words

An example of a pairwise verbal judgment shown in Figure 9.

```
           GOAL: DETERMINE IMPORTANCE OF ENVIRONMENTAL FACTORS
        With respect to                                              GOAL
     CIRCLE A
        is VERY STRONGLY more IMPORTANT than
     TRINGLE B

              EXTREME-----------

              VERY STRONG-------  <--

              STRONG------------

              MODERATE----------

              EQUAL-------------
```
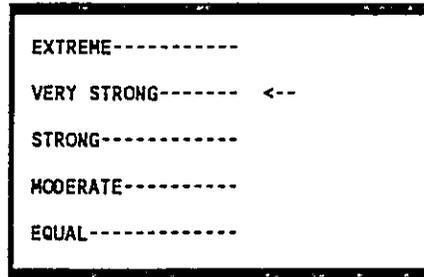
Figure 9 - Pairwise Verbal Judgment

Subjects were not told the numerical value[11] associated with each of the verbal judgments.[12] They were told only that the meaning of "equal" was obvious, and that "extreme" meant an order of magnitude or more, but not necessarily that the largest shape was being compared to the smallest. Because the words were not precisely defined, some subjects or groups of subjects tended to use words higher up on the scale than did other subjects or groups.

A typical set of judgments is shown in Figure 10, where each verbal judgment is shown numerically. The judgments are actually the upper part of a reciprocal matrix -- the diagonal elements are equal to 1 and the elements below the diagonal are reciprocal to the elements above the diagonal. The derivation of priorities from a set of pairwise

---

[11]the numerical representations are from the AHP scale: EQUAL (1); MODERATE (3); STRONG (5); VERY STRONG (7); EXTREME (9). Intermediate values are 2,4,6and 8. See Saaty [5] and Forman, et al. [3]

[12]We expect that if they had been so informed, and had made their judgments with precise numbers instead of "fuzzy"words, their results would have been even more accurate since they were in fact comparing objects that posessed attributes. However, this must be established in a future study.

comparisons was accomplished using the Analytic Hierarchy Process[13] procedure of calculating the largest eigenvector of the comparison matrix. These priorities are also shown in Figure 10.

```
              JUDGMENTS AND PRIORITIES WITH RESPECT TO
                               GOAL

            CIRCLE A  TRINGL B  SQUARE C  DIAMND D  RECT E
   CIRCLE A             7.0       3.0       5.0      6.0
   TRINGL B                      (5.0)     (3.0)    (2.0)
   SQUARE C                                 3.0      3.0
   DIAMND D                                          2.0
   RECT E

Matrix entry indicates that ROW element is ___
   1 EQUALLY   3 MODERATELY   5 STRONGLY   7 VERY STRONGLY   9 EXTREMELY
more IMPORTANT than COLUMN element
      unless enclosed in parenthesis.

0.514
CIRCLE A  ████████████████████████████████████████

0.048
TRINGL B  ████

0.242
SQUARE C  ████████████████████

0.117
DIAMND D  ████████

0.078
RECT E    ██████

              INCONSISTENCY RATIO[14] = 0.029
```

Figure 10 -- Matrix of judgments and resulting priorities

Algebraically, a minimum of four comparisons are required to derive the relative areas. For example, *if* one assumed that the numerical representations of the verbal judgments shown in the first row of the matrix in Figure 10 were accurate the resulting priorities would be A = .543; B = .078; C = .181; D = .109; and E = .090. Notice that for

---

[13]Saaty, T.L., The Analytic Hierarchy Process, (New York: McGraw Hill, 1980). p49-51.

[14]Had the inconsistency ratio been greater than about 10%, the judgments would have been re-examined. See Saaty[5], Forman et al.[3], and Golden et al. [4] for discussions of inconsistency.

this set of priorities, A is seven times B, 3 times C, 5 times D, and 6 times E, corresponding to the values in the first row of the matrix. However, the values in the matrix can *not* be assumed to be accurate. Since the judgments were fuzzy verbal judgments, the numbers are at best crude representations of the words.[15] The AHP derived priorities[16] are based on the entire set of judgments and are almost always more accurate[17] than results based on the minimal number of required judgments. This is indeed the case here as can be seen by comparing the derived priorities from the minimal set of judgments and those derived from the full set of judgments (shown in Figure 10) to the true priorities of the geometric shapes, as shown in Table 1 (A = .475; B = .049; C = .232; D = .151; and E = .093). The priorities derived from the full set of judgments (those shown in Figure 10) are in quite close agreement with the actual priorities. Notice that the derived priorities do *not* necessarily agree with the numerical representation of any one judgment (as shown in the matrix in Figure 10). For example, the ratio of the derived priorities of A and B is about 10.7, which is closer to the true ratio 9.7 than is the 7.0 numerical representation of the verbal judgment VERY STRONG.

---

[15]Even if the judgments were made numerically, we would expect some error in each of the judgments.

[16]Which are the normalized values of the eigenvector corresponding to the largest eigenvalue of the pairwise comparison matrix.

[17]Churchman [2] has said that "Accuracy itself is a measurement ..ofthe degree to which a given measurement may deviate from the truth. No procedure can claim the name of measurement unless it includes methods of estimating accuracy." The truth in this case is known because we are dealing with measured geometric shapes. The degree of deviation from the truth is, however relative. Churchman also observed that "'Deviation from the truth' must be defined in terms of the uses to which the measurement is put. [with the ...] consequence that accuracy is a highly relative term, the meaning of which depends on the individual decision maker."

Value of Redundancy

There is value in making more than the minimal number of judgments necessary to algebraically calculate priorities. We will refer to this as the value of redundancy. Redundancy is often used in statistical estimation. If a scientist wants to measure a physical property accurately, he/she will make numerous measurements (beyond the one required to make an estimate) and use the average of the measurements as the estimate of the quantity being measured. The number of measurements required for a desired level of accuracy of the estimate depends on the variance (error) of the measurement process. The greater the variability, the larger the number of measurements required. If there is no error, only one measurement is required, and there is no value in making more than one measurement. However, if there is error (and there almost always will be) then measuring the same object more than once (i.e., making redundant judgments) will produce a more accurate estimate -- the larger the sample size (more redundancy), the more likely the sample average will be close to the true value.

The value of redundancy in the pairwise comparisons process is similarly related to the accuracy of the judgments. If there were no error in the judgments comparing one factor with each of the others then there would be no value in making redundant judgments. Since the subjects in this experiment used imprecise words to compare the relative areas of the five geometric shapes, there is likely to be a considerable amount of error in the verbal judgments, and consequently considerable value in using redundancy when deriving the priorities. The results of the experiment confirm this.

WITH REDUNDANCY:

	ACTUALS	AREA01	AREA02	AREA03	AREA04	AREA48	AREA49	AREA50	AREA51
A	0.475	0.460	0.445	0.451	0.499	0.511	0.500	0.487	0.518
B	0.049	0.076	0.058	0.060	0.047	0.057	0.052	0.051	0.043
C	0.232	0.227	0.235	0.239	0.264	0.197	0.230	0.208	0.233
D	0.151	0.136	0.153	0.155	0.131	0.129	0.147	0.162	0.141
E	0.093	0.101	0.109	0.095	0.058	0.105	0.071	0.091	0.064

WITHOUT REDUNDANCY

		AREA01	AREA02	AREA03	AREA04	AREA48	AREA49	AREA50	AREA51
A	0.475	0.460	0.418	0.472	0.543	0.519	0.528	0.506	0.550
B	0.049	0.077	0.042	0.094	0.090	0.074	0.075	0.072	0.079
C	0.232	0.230	0.279	0.157	0.181	0.104	0.176	0.169	0.183
D	0.151	0.131	0.167	0.157	0.109	0.130	0.132	0.127	0.110
E	0.093	0.102	0.093	0.118	0.078	0.173	0.088	0.127	0.079

Squared Errors	A		0.00023	0.00092	0.00057	0.00059	0.00130	0.00063	0.00014	0.00189
With Redundancy	B		0.00073	0.00008	0.00013	0.00001	0.00007	0.00001	0.00001	0.00003
	C		0.00002	0.00001	0.00005	0.00105	0.00120	0.00000	0.00056	0.00000
	D		0.00022	0.00001	0.00002	0.00039	0.00048	0.00002	0.00013	0.00010
	E		0.00006	0.00027	0.00000	0.00119	0.00014	0.00049	0.00000	0.00084
Sum Sq. Errors			0.00126	0.00128	0.00076	0.00323	0.00319	0.00114	0.00084	0.00287
Mean Square Error			0.00025	0.00026	0.00015	0.00065	0.00064	0.00023	0.00017	0.00057
Avg. Mean Square Error		0.00049								

Squared Errors	A		0.00023	0.00320	0.00001	0.00457	0.00195	0.00284	0.00096	0.00559
W/O Redundancy	B		0.00076	0.00005	0.00207	0.00172	0.00063	0.00070	0.00054	0.00087
	C		0.00000	0.00220	0.00555	0.00261	0.01643	0.00312	0.00401	0.00238
	D		0.00038	0.00027	0.00004	0.00180	0.00045	0.00036	0.00060	0.00169
	E		0.00008	0.00000	0.00063	0.00024	0.00641	0.00002	0.00112	0.00021
Sum Sq. Errors			0.00147	0.00572	0.00830	0.01095	0.02587	0.00705	0.00724	0.01073
Mean Square Error			0.00029	0.00114	0.00166	0.00219	0.00517	0.00141	0.00145	0.00215
Avg. Mean Square Error		0.00197								

Ratio of
Average Mean Square Errors   404%

Table 1 -- Experiment Results

Significance

The average mean square error without redundancy is more than four times, or 400% greater than the average mean square error with redundancy. This illustrates the value of using redundant judgments. While this result is obviously statistically significant, the practical significance is of much greater importance. Suppose an expert were supplying judgments about the likelihoods of five (mutually exclusive and collectively exhaustive) events. And suppose that the relative areas of the shapes in Figure 8 represent the expert's (imprecise) estimates of the relative likelihoods of the five events. First of all, pairwise comparisons would be easier to make than absolute comparisons. That is, it is easier to

estimate how much larger A is than B, and how much larger B is than C, etc, than it would be to estimate the number of square inches in A, B, etc. This is one aspect of the AHP comparison process. Secondly, since the expert typically has only a fuzzy knowledge of the relative likelihood of events, it would be much easier for an expert to make (and justify) a judgment in the form:

"A is strong or very strongly more likely than B"

than it is to say that

"A is 10 times more likely than B".

References

1. Buchanan, B. G. and Shortliffe, E. H., editors, *Rule-Based Expert Systems*, Reading, MA: Addison-Wesley, 1984.

2. Churchman, C.W., "Why Measure?", *Measurement: Definition and Theories*, Wiley 1967, pp 83-94.

3. Forman, E. H., Saaty, T. L., Selly, M. A., Waldron, R., *Expert Choice*, Decision Support Software, McLean, VA., 1983.

4. Golden, B. L., Wang, Q. 1989. An Alternate Measure of Consistency. *The Analytic Hierarchy Process: Applications and Studies* edited by Golden, B., Wasil, E., and Harker, P. T. Springer-Verlag, New York, p68-81.

5. Saaty, T.L., *The Analytic Hierarchy Process*. McGraw Hill, New York, 1980.