NASA Technical Memorandum 88224

Automatic Probabilistic Knowledge Acquisition from Data

William B. Gevarter

April 1986

National Aeronautics and
Space Administration

NASA Technical Memorandum 88224

Automatic Probabilistic Knowledge Acquisition from Data

William B. Gevarter, Ames Research Center, Moffett Field, California

April 1986

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

AUTOMATIC PROBABILISTIC KNOWLEDGE ACQUISITION FROM DATA

William B. Gevarter

ABSTRACT

This memorandum documents an outline for a computer program for extracting significant correlations of attributes from masses of data. This information can then be used to develop a knowledge base for a probabilistic "expert system." The method determines the "best" estimate of joint probabilities of attributes from data put into contingency table form. A major output from the program is a general formula for calculating any probability relation associated with the data. These probability relations can be utilized to form IF-THEN rules with associated probability, useful for expert systems.

INTRODUCTION

Knowledge acquisition is a major bottleneck in developing "expert systems." Thus a recent focus of the artificial intelligence (AI) community has been on "machine learning." Though this has been a theme in AI for several decades, it has only been in the last few years, spurred on by the popularity of expert systems, that machine learning has received major attention.

Because of the emergence of sophisticated expert system building tools such as KEE (Intellicorp, 1985) and ART (Williams, 1985), and a host of follow-on simpler systems, the main difficulty in building conventional expert systems has now shifted to knowledge acquisition and choosing the most appropriate knowledge structures and representations. Thus knowledge acquisition is an important and timely research area for NASA to investigate.

Approaches to knowledge acquisition have included psychological techniques for interviewing experts (Boose, 1984 and Kahn et al., 1985) and automatic production of classification-oriented expert systems from examples as exemplified by the TIMM (General Research Corp., 1985) expert system building tool. Thus far, such approaches have had only a limited range of successful applications. Therefore, methods of knowledge acquisition and knowledge extraction from data are important current AI research topics. Knowledge acquisition can be construed as learning.

AI learning systems can be classified according to the following strategies:

1. Rote learning
2. Learning from instruction
3. Learning by analogy

4. Learning by examples

5. Learning from observation and discovery

The last learning strategy is perhaps the most recent and the most exciting. This strategy is exemplified by Lenat's (1982) AM and EURISKO systems and Langley et al. (1983) BACON system. These discovery systems are heavily knowledge-based. The usual focus of discovery systems is on discovering concepts. Cheeseman (1984) has explored another facet, that of developing specific correlations from data. His approach is probabilistic in nature and is primarily procedurally (syntactically) oriented. This approach is particularly appropriate when the source of the knowledge is in the form of large masses of undigested data, such as those obtained from wind tunnel tests; spacecraft observations; computer simulations; or psychological, medical, and social surveys.

Commercial AI learning systems such as Expert-Ease (Derfler, 1985) and TIMM are aimed at developing decision aids from examples. In general, learning from examples is predicated on positive examples, which promote generalization ... and negative examples, which reduce generalization. However, commercial learning tools are not designed to extract significant information from data for which no conclusions have yet been reached.

Many researchers believe that truly powerful intelligent systems will be difficult to achieve without a learning component because of the huge amount of knowledge many future expert systems will require, as well as the need to improve performance by learning new search heuristics as the system is used.

This memorandum outlines an approach to probabilistically determining significant relationships in masses of data. This can be particularly important because NASA has masses of unevaluated data from its space explorations. Automatic means to find significant correlations in these data can begin to reduce this mammoth NASA reserve data bank. The approach outlined in this paper draws on previous work by Cheeseman (1983, 1984) in this area. Using this approach, the resultant information, probabilistically extracted from the data, allows calculations of the conditional probability of any proposition associated with the data, given any combination of evidence. This information can be used as clues for discovering more causal explanations. The probabilistically extracted information can also be transformed into IF-THEN ("condition–conclusion") rules (with associated prob bility) found useful in expert systems. For example, the probability of A given B and C is p, written as

$$P(A \mid B,C) = p$$

can also be written as

$$IF\ B\ AND\ C,\ THEN\ A\ (with\ probability\ p)$$

The system described in this memorandum does not generate rules explicitly. It generates and stores significant joint probabilities instead. Particular conditional probabilities can be calculated from this information as required by noting that a conditional probability can be written as the ratio of corresponding joint probabilities. Thus, for example

$$P(A \mid B,C) = \frac{p(A\ B\ C)}{p(B\ C)}$$

PROBLEM DEFINITION

The problem explored in this paper is that of extracting information from data which can then be used to form the knowledge base of a probabilistic expert system. The resultant formulation summarizes all the probabilistic information found from the data and can be used to calculate any probability associated with the data. The information found is the significant joint probabilities of attributes from data (which has resulted from a collection of observations). An illustrative example followed in this paper determines the probabilistic relationships of cancer to smoking given a set of observations on people over the age of 60 whose hypothetical case histories are obtained from the completion of the following questionaire

A. SMOKING HISTORY

1. Smoker
2. Non smoker not married to a smoker
3. Non smoker married to a smoker

B. CANCER

1. Yes
2. No

C. FAMILY HISTORY OF CANCER

1. Yes
2. No

A set of data thus obtained from a survey of 3428 individuals might appear as shown in Figures 1a and 1b (called "contingency tables" [1]).

The numbers shown in each box or cell refer to the total number of individuals who have that combination of attributes. Thus, the number of smokers who do not have cancer despite a family history of cancer is given as 410. This can be written in a shorthand notation as

$$N_{121}^{ABC} = 410$$

This states that there were 410 individuals that had the attribute values

```
A (SMOKING)              = 1   (Smoker)
B (CANCER)               = 2   (No)
C (FAMILY HISTORY        = 1   (Yes)
   OF CANCER)
```

[1] Appendix A indicates how original data can be put into contingency table form.

		B CANCER	
		1 Yes	2 No
A SMOKING	1. Smoker	130	410
	2. Non smoker	62	580
	3. Non smoker married to a smoker	78	520

C FAMILY HISTORY OF CANCER
1. Yes

(a)

		B CANCER	
		1 Yes	2 No
A SMOKING	1.	110	640
	2.	31	460
	3.	22	385

C FAMILY HISTORY OF CANCER
2. No

(b)

Figure 1: DATA ON SMOKING AND CANCER IN U.S. POPULATION OF AGE GREATER THAN 60.

In general, we can refer to the number of individuals with the ith value of attribute A, the jth value of the attribute B and the kth value of the attribute C, as

$$N_{ijk}^{ABC} \text{ or } N_{ijk}$$

where i, j, k are the numbers associated with the values of the attributes.

We will assume that the range of values for each attribute is complete (made so by adding the value "other," if necessary) so that the the number of people obtained by summing the numbers for each of the values of an attribute (e.g., B (CANCER)) will always add up to N (the total number of individuals surveyed).

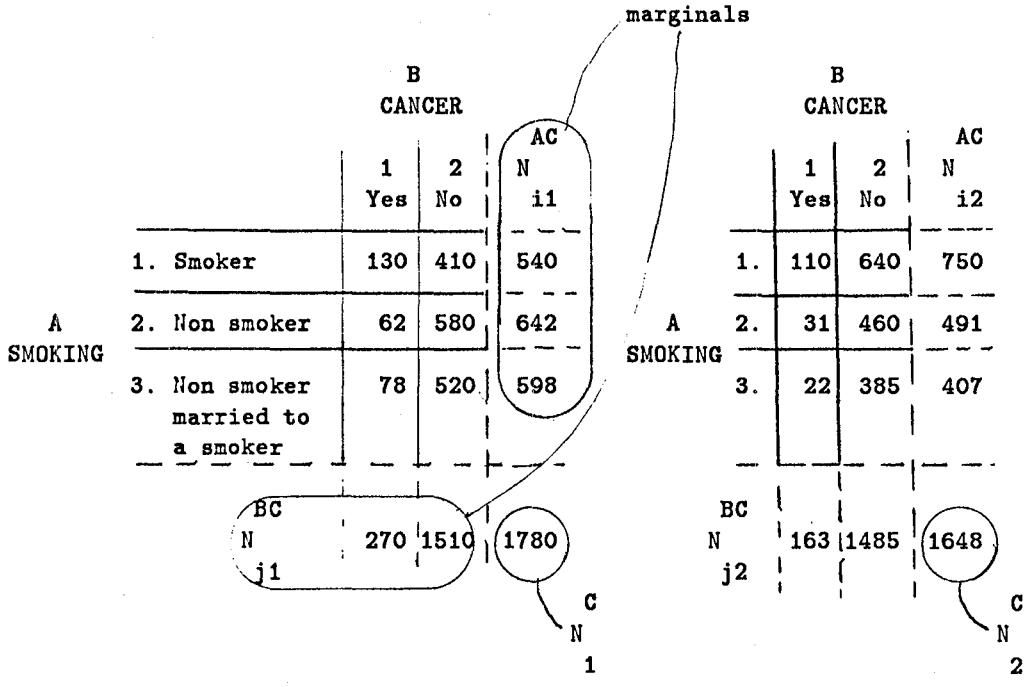
If we add up the numbers in each row or column of Figures 1a and 1b, we obtain the marginal values (placed in the margins) as shown in Figures 2a and 2b. If we sum across C (FAMILY HISTORY OF CANCER), we obtain Figure 2c, which relates SMOKING to CANCER without breaking up the results with respect to family history.

In equation form, these summations can be simply written as

$$N_{ij} = \sum_k N_{ijk} \tag{1}$$

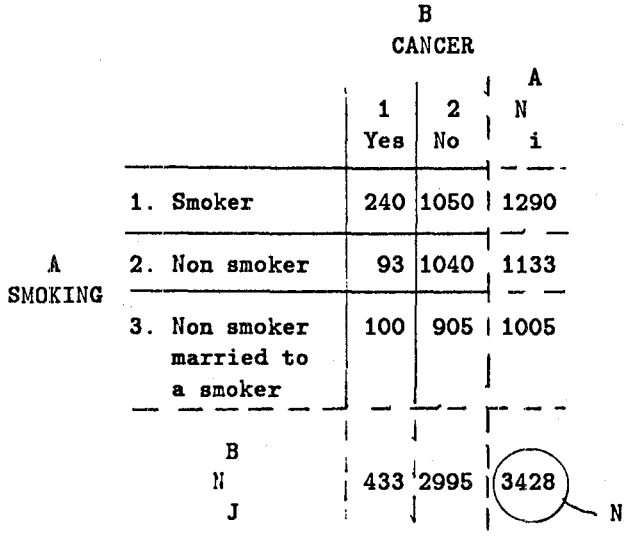
$$N_{jk} = \sum_i N_{ijk} \tag{2}$$

$$N_{ik} = \sum_j N_{ijk} \tag{3}$$



C FAMILY HISTORY OF CANCER
1. Yes
(a)

C FAMILY HISTORY OF CANCER
2. No
(b)



RELATION OF SMOKING TO CANCER
(c)

Figure 2: CANCER DATA ON U.S. POPULATION OF AGE GREATER THAN 60.

and

$$N_i = \sum_j N_{ij} = \sum_j \sum_k N_{ijk} \tag{4}$$

or

$$N_i = \sum_k N_{ik} = \sum_k \sum_j N_{ijk} \tag{5}$$

.
.
.

Similarly, the total number of individuals, N, is simply the summation across all the indices

$$N = \sum_i \sum_j \sum_k N_{ijk} \tag{6}$$

APPROACH

The approach taken for finding joint probabilities of attributes is to maximize the entropy of the discrete probability distribution while satisfying the constraints imposed by all the known probabilities. This can also be thought of as achieving the maximum uncertainty in the values of these higher-order probabilities; as any lesser uncertainty would imply further constraints. "Maximum entropy" probability values distribute the uncertainty (H) as evenly as possible over the underlying probability space in a way consistent with the constraints.

The "known probabilities" (constraints) are determined by applying a significance test to the data. A "minimum message length" criterion is used as the test for significance between the observed values of occurence of higher-order combinations of attributes in the data and the values calculated by the maximum entropy approach using the currently known constraints. If the minimum message length required to encode an observed value of occurence assuming a chance distribution is less than that given by the maximum entropy approach (using the constraints thus far), then the observed value is deemed significant and the joint probability associated with it is added to the list of constraints. Once all the significant joint probabilities are determined, any other probability relationships associated with the data can be readily calculated from the resulting succinct equation.

MAXIMIZING THE ENTROPY

The entropy (uncertainty) is given in terms of the joint probabilities as (Jaynes, 1979)

$$H = -\sum_{ijk\ldots} p_{ijk\ldots} \log p_{ijk\ldots} \tag{7}$$

where i,j,k,... are the indices of the values of the attributes and p_{ijk} is the joint probability (probability of the simultaneous occurrence of the ith value of the attribute A, the jth value of the attribute B, and the kth value of attribute C).

For simplicity, in the rest of this paper, we will only consider three attributes — A, B, C. The extension to a larger number of attributes is straight forward.

To maximize the entropy we first add to H the constraints (associated with the prior known probabilities) using LaGrange multipliers (w's) to form H'. We then take the derviative of H' with respect to each of the unknown variables—the probabilities and the LaGrange multipliers—and set them equal to zero to find values for the variables that maximize H'. Thus

$$\begin{aligned} H' = \; & -\textstyle\sum_{ijk} p_{ijk} \log p_{ijk} \\ & +w(1 - \textstyle\sum_{ijk} p_{ijk}) \\ & +w_i(p_i - \textstyle\sum_{jk} p_{ijk}) + w_j(P_j - \textstyle\sum_{ik} p_{ijk}) + w_k(p_k - \textstyle\sum_{ij} p_{ijk}) \\ & +w_{ij}(p_{ij} - \textstyle\sum_{k} p_{ijk}) + w_{ik}(p_ik - \textstyle\sum_{j} p_{ijk}) + w_{jk}(p_{jk} - \textstyle\sum_{i} p_{ijk}) \\ & +\cdots \end{aligned} \tag{8}$$

Taking derivatives with respect to the probabilities, we obtain

$$\partial H'/\partial p_{ijk} = -\log p_{ijk} - 1 - w - w_i - w_j - ... - w_{ij} - ... = 0 \tag{9}$$

Therefore

$$p_{ijk} = e^{-(w_0 + w_i + ... + w_{ij} + ...)} \tag{10}$$

where we have defined

$$w_0 = w + 1 \tag{11}$$

From Equation 10

$$p_{ijk} = a_0 a_i a_j ... a_{ij} ... \tag{12}$$

where we have defined

$$a_i = e^{-w_i} \tag{13}$$

.
.
.

Taking the partial derivatives of H' with respect to the w multipliers and setting them equal to zero, simply returns our given contraints

$$\partial H'/\partial w = 0 \rightarrow \sum_{ijk} p_{ijk} = 1 \tag{14}$$

$$\partial H'/\partial w_i = 0 \rightarrow \sum_{jk} p_{ijk} = p_i \tag{15}$$

$$\partial H'/\partial w_j = 0 \rightarrow \sum_{ik} p_{ijk} = p_j \tag{16}$$

$$\partial H'/\partial w_k = 0 \rightarrow \sum_{ij} p_{ijk} = p_k \tag{17}$$

$$\partial H'/\partial w_{ij} = 0 \rightarrow \sum_{k} p_{ijk} = p_{ij} \tag{18}$$

$$\partial H'/\partial w_{ik} = 0 \rightarrow \sum_{j} p_{ijk} = p_{ik} \tag{19}$$

$$\partial H'/\partial w_{jk} = 0 \rightarrow \sum_{i} p_{ijk} = p_{jk} \tag{20}$$

(and so on for any higher-order known prior probabilities).

Substituting for p_{ijk} from Equation 12 into Equations 14 – 20, yields

$$a_0 \sum_{ijk} a_i a_j ... a_{ij} a_{ik} ... = 1 \tag{21}$$

$$a_0 a_i \sum_{jk} a_j a_k a_{ij} a_{ik} ... = p_i \tag{22}$$

$$a_0 a_j \sum_{ik} a_i a_k a_{ij} a_{ik} ... = p_j \tag{23}$$

$$a_0 a_k \sum_{ij} a_i a_j a_{ij} a_{ik} \ldots = p_k \tag{24}$$

.
.
.

These summations are simplified if we group the summations by the indices[2]. Thus, for example

$$a_0 \sum_i a_i \sum_j a_j a_{ij} \sum_k a_k a_{ik} a_{jk} = 1 \tag{25}$$

$$a_0 a_i \sum_j a_j a_{ij} \sum_k a_k a_{ik} a_{jk} = p_i \tag{26}$$

$$a_0 a_j \sum_i a_i a_{ij} \sum_k a_k a_{ik} a_{jk} = p_j \tag{27}$$

$$a_0 a_k \sum_i a_i a_{ik} \sum_j a_j a_{ij} a_{jk} = p_k \tag{28}$$

$$a_0 a_{ij} a_i a_j \sum_k a_k a_{ik} a_{jk} = p_{ij} \tag{29}$$

$$a_0 a_{ik} a_i a_k \sum_j a_j a_{ij} a_{jk} = p_{ik} \tag{30}$$

$$a_0 a_{jk} a_j a_k \sum_i a_i a_{ij} a_{ik} = p_{jk} \tag{31}$$

As will be illustrated later using our example, this set of simultaneous equations is iteratively solved for a values. Initially, the a values are calculated from the first-order probabilities derived from the data. Then the a values are recalculated using any known prior second-order probabilities. Based on the resulting a values, predicted second-order joint probabilities of the attributes are then calculated (using Equation 12 or equivalently Equations 25 - 31) and the observed data is evaluated to see whether it differs significantly from the values predicted.

If the predicted probabilities of the observed values of the combinations of attributes are less than the probabilities of their occurence owing to pure chance, then the values observed from the data are statistically significantly different from those calculated from the constraints used thus far. In this case, these significantly different observed values are used to form new constraints and the a values are recalculated. This process is repeated at this level and each successive level until all the observed statistically significant correlations are accounted for.

PREDICTING THE VALUE OF N_{ijk}

The probability of finding the number N_{ijk} of occurences having the ith value of attribute A, the jth value of B, and the kth value of C is given by the well-established "binomial distribution"

$$p(N_{ijk} \mid p_{ijk}, N) = \binom{N}{N_{ijk}} (p_{ijk})^{N_{ijk}} (1 - p_{ijk})^{N - N_{ijk}} \tag{32}$$

[2]A method for calculating such "sum of products" equations is given by Appendix B.

where

N is the total number sampled

p_{ijk} is the prior probability (calculated from the a values) of the ith value of A, the jth value of B, and the kth value of C occurring together in the population being sampled.

The predicted mean of N_{ijk} is given by

$$(N_{ijk})_m = Np_{ijk} \tag{33}$$

and the associated standard deviation of N_{ijk} is given by

$$(N_{ijk})_{sd} = \sqrt{Np_{ijk}(1 - p_{ijk})} \tag{34}$$

The mean and standard deviation are useful for estimating the significance of the difference between the observed value of N_{ijk} and the predicted value.

SIGNIFICANCE TESTING OF THE OBSERVED VALUES OF THE N_{ijk}'s

In this section we determine the significance of the observed values of the N's by comparing the probability of their chance occurence with the probability predicted by the probability formula, Equation 12, derived from the constraints found thus far. If, for example, the probability of occurence by chance of an observed value of N_{ijk} is greater than that predicted by the formula, we regard that N_{ijk} as significant and use it as a constraint to revise our probability formula so that it will predict the observed value. The revised formula is then used to predict the probability of occurence of the remaining observed values of the N's and the procedure recursively repeated until no further significant N's are found.

The procedure, outlined above, starts by comparing the probabilities of the chance occurence of the observed values of the second-order N's with the probabilities predicted by Equation 12 and uses the resultant most significant N to update Equation 12. The remaining second-order N's are then evaluated using the new predictions and the procedure is recursively repeated until all the significant second-order N's have been determined. This procedure is then repeated for the third-order N's and so on.

In determining the probability of an observed value of N_{ijk} ocurring by chance, we note that the values in the cells must add up to their constraining marginal values. Thus no cell can have a value exceeding its significant marginals minus the values of any other significant cells associated with those marginals. This sets a maximum value for a cell. If for a cell all the other cells associated with one of its significant marginals have been found to be significant, then the cell must have the value observed for it. If the cell is not so constrained, then for the chance case it is equally likely that it will have any integer value from zero to its maximum value (discussed above).

We now derive the equation needed for comparing the probabilities of occurence of the observed value of N_{ijk} as calculated by Equation 12 and as calculated for chance.

The well-known "bayesian formula" for calculating the posterior value of a hypothesis, $h1$, given data, D, is

$$p(h1 \mid D) = \frac{p(h1)\, p(D \mid h1)}{p(D)} \tag{35}$$

where $p(h1)$ is the prior probability of the hypothesis. For our purposes, a more convenient relative form of Baye's rule gives the likelihood ratio of the posterior probability of two different hypotheses (given the same data)

$$\frac{p(h1 \mid D)}{p(h2 \mid D)} = \frac{p(h1)\, p(D \mid h1)}{p(h2)\, p(D \mid h2)} \tag{36}$$

Taking the log of the likelihood ratio, we obtain

$$\ln \frac{p(h1 \mid D)}{p(h2 \mid D)} = [-\ln p(h2) - \ln p(D \mid h2)] \;\; - \;\; [-\ln p(h1) - \ln p(D \mid h1)] \tag{37}$$

In information theory, the minimum message length required to encode (communicate) a particular choice (e.g., $h1$) from a set of mutually exclusive and exhaustive hypotheses is proportional to $-\ln p(h1)$ (Jaynes, 1979). Thus, Equation 37 can be interpreted as proportional to the difference in the minimum message lengths required to represent the two hypotheses given the data.

There are two basic hypotheses for the N_{ijk} obtained from the data for a particular cell ijk.

H1 Given that we have found M nth-order significant constraints (joint probabilities), there are no more nth-order significant constraints. (i.e., Equation 12 adequately predicts the probability of occurence of the observed values of the remaining nth-order N's.)

H2 Given that we have found M nth-order significant constraints (joint probabilities), there is at least one more nth-order significant constraint and this cell is the next nth-order significant constraint.

The hypothesis $H2$ can be broken up into two hypotheses

$H2'$ There is at least one more nth-order significant constraint.

$H2''$ This cell is the next nth-order significant constraint.

Thus

$$p(H2) = p(H2'\ H2'') = p(H2')\, p(H2'' \mid H2') \tag{38}$$

where without any other information

$$\begin{aligned} p(H2'' \mid H2') &= \frac{1}{\text{remaining available cells at the current order}} \\ &= \frac{1}{\text{no. of cells at this order } - M} \end{aligned} \tag{39}$$

For the highest order this is simply

$$p(H2'' \mid H2') = \frac{1}{(IJK... - M)} \tag{40}$$

where I, J, K are the total number of values of the A, B, and C attributes, respectively.

For hypothesis $H2$, lacking prior knowledge, the value of N_{ijk} is equally likely to be any integer in the range of values available to it. Thus, for third-order combinations

IF

$$\min \begin{pmatrix} [JK - No_i(N_{iyz}^{ABC})] \\ [IK - No_j(N_{xjz}^{ABC})] \\ [IJ - No_k(N_{xyk}^{ABC})] \\ [J - No_i(N_{iy}^{AB})] \\ [K - No_i(N_{iz}^{AC})] \\ [I - No_j(N_{xj}^{AB})] \\ [K - No_j(N_{jz}^{BC})] \\ [I - No_k(N_{xk}^{AC})] \\ [J - No_k(N_{yk}^{BC})] \end{pmatrix} > 1$$

THEN

$$\begin{aligned} p(D \mid H2) &= p(N_{ijk} \mid H2) \\ &= p(N_{ijk} \mid N_i, N_j, N_k, I, J, K, significant(N_{xyz}^{ABC}s), H2) \\ &= 1 \,/\, \{\, 1 + \min \begin{pmatrix} [N_i^A - \sum_{yz\ yz \neq jk} significant(N_{iyz}^{ABC}s)] \\ [N_j^B - \sum_{xz\ xz \neq ik} significant(N_{xjz}^{ABC}s)] \\ [N_k^C - \sum_{xy\ xy \neq ij} significant(N_{xyk}^{ABC}s)] \\ [significant\ N_{ij}^{AB} - \sum_{z\ z \neq k} significant(N_{ijz}^{ABC}s)] \\ [significant\ N_{ik}^{AC} - \sum_{y\ y \neq j} significant(N_{iyk}^{ABC}s)] \\ [significant\ N_{jk}^{BC} - \sum_{x\ x \neq i} significant(N_{xjk}^{ABC}s)] \end{pmatrix} \,\} \end{aligned} \tag{41}$$

ELSE

$$p(D \mid H2) = 1$$

(as the value of N_{ijk} is then completely determined from the marginal values and the significant values previously found)

where we have defined

$$No_i(N_{iyz}s) = \text{number of } N_{iyz}\text{'s found significant that have the ith value of the A attribute} \tag{42}$$

etc.

Note that in Equation 41 our constraints are the first-order marginals — N_1^A, N_j^B, N_k^C — and any higher-order marginals found significant in our analysis or originally given as significant.

Using Equation 38 in Equation 37, the hypothesis $H1$ (that there are no more significant nth-order constraints so that the prior probability calculated from the a's is adequate) is more likely than $H2$ if

$$[-\ln p(H2') - \ln p(H2'' \mid H2') - \ln p(D \mid H2)] \quad - \quad [-\ln p(H1) - \ln p(D \mid H1)] \; > \; 0 \tag{43}$$

or in abreviated form as

$$m2 - m1 > 0 \tag{44}$$

where using Equations 39 and 41 in Equation 43,

IF

$$\min \left(\begin{array}{c} [JK - No_i(N_{iyz}^{ABC})] \\ [IK - No_j(N_{xjz}^{ABC})] \\ [IJ - No_k(N_{xyk}^{ABC})] \\ [J - No_i(N_{iy}^{AB})] \\ [K - No_i(N_{iz}^{AC})] \\ [I - No_j(N_{xj}^{AB})] \\ [K - No_j(N_{jz}^{BC})] \\ [I - No_k(N_{xk}^{AC})] \\ [J - No_k(N_{yk}^{BC})] \end{array} \right) > 1$$

THEN

$$\begin{aligned} m2 = \quad & -\ln p(H2') + \ln(\text{no. of cells at this order} - \text{M}) \\ & + \ln \{ \min \left(\begin{array}{c} [N_i^A - \sum_{yz\ yz \neq jk} significant(N_{iyz}^{ABC}s)] \\ [N_j^B - \sum_{xz\ xz \neq ik} significant(N_{xjz}^{ABC}s)] \\ [N_k^C - \sum_{xy\ xy \neq ij} significant(N_{xyk}^{ABC}s)] \\ [significant\ N_{ij}^{AB} - \sum_{z\ z \neq k} significant(N_{ijz}^{ABC}s)] \\ [significant\ N_{ik}^{AC} - \sum_{y\ y \neq j} significant(N_{iyk}^{ABC}s)] \\ [significant\ N_{jk}^{BC} - \sum_{x\ x \neq i} significant(N_{xjk}^{ABC}s)] \end{array} \right) + 1 \} \end{aligned} \tag{45}$$

ELSE

$$m2 = -\ln p(H2') + \ln(\text{no. of cells at this order} - \text{M})$$

Similarly, using Equation 32 in Equation 43

$$m1 = -\ln p(H1) \; - \; \left[\ln \binom{N}{N_{ijk}} + N_{ijk} \ln P_{ijk} + (N - N_{ijk}) \ln(1 - P_{ijk})\right] \tag{46}$$

For the observed value of N_{ijk} to be statistically significant requires (from Equation 44) that

$$m2 - m1 < 0 \tag{47}$$

If the observed N_{ijk} is statistically significant, then it forms a new constraint and a new set of a values is calculated that will predict it when using these new values in Equation 12.

CALCULATING THE INITIAL a VALUES BASED ON THE FIRST ORDER PROBABILITIES

The first order probabilities are readily calculated from the data as

$$\begin{aligned} p_i^A &= N_i^A/N \\ p_j^B &= N_j^B/N \\ p_k^C &= N_k^C/N \end{aligned} \tag{48}$$

If we start with these as our only initial constraints, then from Equations 25 - 28, we obtain for the data from our example

$$a_0\, a^A a^B a^C = 1 \tag{49}$$

$$a_0\, a_1^A a^B a^C = p_1^A = .38 \tag{50}$$

$$a_0\, a_2^A a^B a^C = p_2^A = .33 \tag{51}$$

$$a_0\, a_3^A a^B a^C = p_3^A = .29 \tag{52}$$

$$a_0\, a_1^B a^A a^C = p_1^B = .13 \tag{53}$$

$$a_0\, a_2^B a^A a^C = p_2^B = .87 \tag{54}$$

$$a_0\, a_1^C a^A a^B = p_1^C = .52 \tag{55}$$

$$a_0\, a_2^C a^A a^B = p_2^C = .48 \tag{56}$$

where we have defined

$$a^A = a_1^A + a_2^A + a_3^A \tag{57}$$

$$a^B = a_1^B + a_2^B \tag{58}$$

$$a^C = a_1^C + a_2^C \tag{59}$$

It can readily be verified that the solution to Equations 49 – 56 is

$$\begin{aligned} &a_0 = 1 \\ &a^A = 1 \quad a^B = 1 \quad a^C = 1 \\ &a_1^A = .38 \quad a_2^A = .33 \quad a_3^A = .29 \\ &a_1^B = .13 \quad a_2^B = .87 \\ &a_1^C = .52 \quad a_2^C = .48 \end{aligned} \tag{60}$$

which simply means that for this simple case where there are no constraining higher-order probabilities, the a values are just the values of the associated first-order probabilities.

Substituting these values into Equations 29 – 31, we find that the higher-order probabilities are just equal to the product of the corresponding first-order probabilities. This indicates that the maximum entropy approach has distributed the higher-order probabilities based on the attributes being statistically independent — as we would expect, having no other information.

Thus from Equation 12

$$p_{ijk}^{ABC} = p_i^A p_j^B p_k^C \tag{61}$$

and

$$p_{ij}^{AB} = p_i^A p_j^B \sum_k p_k^C = p_i^A p_j^B \tag{62}$$

.
.
.

Lacking other information, we will assume that the probability of there being one more constraint is equal to the probability that there are no more constraints. (If prior information is available about the possibility of remaining constraints, then this is easily incorporated.) Assuming equality,

$$p(H2') = p(H1) \tag{63}$$

resulting in these terms cancelling in Equation 43, simplifying the calculations for $(m2 - m1)$. Note (using Equations 45 and 46 for m_2 and m_1) that

If $p(H2') = .6$ so that $p(H1) = .4$, this makes a difference of $-.40$ in $(m2 - m1)$.

If $p(H2') = .8$ so that $p(H1) = .2$, this makes a difference of -1.39 in $(m2 - m1)$.

For our example, Table 1 gives the values of the predicted second-order probabilities (calculated from conditional independence), and values for the observed N_{ij}, N_{ik}, N_{jk} and their predicted mean and standard deviation. Also given are the values of $(m2 - m1)$ — indicative of the statistical significance of the observed values — and the resulting likelihood ratio of the two hypotheses. For our example, there are 16 second order cells from which to choose the significant cell. (Note that even $p(H2') = .8$ only changes the sign of $(m2 - m1)$ for one of the values in our example.)

CALCULATING a VALUES FOR HIGHER-ORDER CONSTRAINTS

If we select N_{12}^{AC}, from Table 1, as the first statistically significant data value to investigate, then (by including the associated a_{12}^{AC}) we obtain from Equations 25 – 28 and 30 the following equations for finding the new values of the a's

$$c[a_1^A a_1^C + b + (a_2^A + a_3^A)a^C] = 1 \tag{64}$$

$$c[a_1^A a_1^C + b] = p_1^A = .38 \tag{65}$$

$$c a_2^C a^C = p_2^A = .33 \tag{66}$$

$$c a_3^A a^C = p_3^A = .29 \tag{67}$$

$$a_0 a_1^B [a_1^A a_2^C + b + (a_2^A + a_3^A)a^C] = p_1^B = .13 \tag{68}$$

$$a_0 a_2^B [a_1^A a_2^C + b + (a_2^A + a_3^A)a^C] = p_2^B = .87 \tag{69}$$

1	2	3	4	5	6	7
Eq 62	Fig 2	Eq 33	Eq 34	$\frac{2-3}{4}$	Eqs 45,46	$e^{m_2-m_1}$
$p_{ij}^{AB} = p_i^A p_j^B$	N_{ij}^{AB}	$N_{ij}^{AB}{}_m$	$N_{ij}^{AB}{}_{sd}$	No. of sd's	$m_2 - m_1$	$\frac{p(H_1\|D)}{p(H_2\|D)}$
$p_{11} = .376 \times .126 = .048$	240	165	12.5	6.03	-11.57	<.1
$p_{12} = .376 \times .874 = .329$	1050	1128	27.5	-2.83	1.75	5.8
$p_{21} = .331 \times .126 = .042$	93	144	11.7	-4.34	-4.74	<.1
$p_{22} = .331 \times .874 = .289$	1040	990	26.5	1.86	3.83	46.1
$p_{31} = .293 \times .126 = .037$	100	127	11.1	-2.43	2.44	11.5
$p_{32} = .293 \times .874 = .256$	905	888	25.6	1.07	4.97	144.0

$p_{jk}^{BC} = p_j^B p_k^C$	N_{jk}^{BC}	$N_{jk}^{BC}{}_m$	$N_{jk}^{BC}{}_{sd}$	No. of sd's	$m_2 - m_1$	$\frac{p(H_1\|D)}{p(H_2\|D)}$
$p_{11} = .126 \times .519 = .065$	270	223	14.4	3.27	.59	1.8
$p_{12} = .126 \times .481 = .061$	163	209	14.0	-3.29	-.21	.8
$p_{21} = .874 \times .519 = .454$	1510	1556	29.2	-1.59	4.77	118.0
$p_{22} = .874 \times .481 = .420$	1486	1440	28.9	1.56	4.62	101.0

$p_{ik}^{AC} = p_i^A p_k^C$	N_{ik}^{AC}	$N_{ik}^{AC}{}_m$	$N_{ik}^{AC}{}_{sd}$	No. of sd's	$m_2 - m_1$	$\frac{p(H_1\|D)}{p(H_2\|D)}$
$p_{11} = .376 \times .519 = .195$	540	668	23.2	-5.54	-10.54	<.1
$p_{12} = .376 \times .481 = .181$	750	620	22.5	5.75	-9.95	<.1
$p_{21} = .331 \times .519 = .172$	642	590	22.1	2.37	2.87	17.6
$p_{22} = .331 \times .481 = .159$	491	545	21.4	-2.52	2.63	13.9
$p_{31} = .293 \times .519 = .152$	598	593	22.1	.22	-.64	.5
$p_{32} = .293 \times .481 = .141$	407	483	20.4	-3.75	-1.49	.2

Table 1: CALCULATED PARAMETER VALUES USEFUL FOR DETERMINING STATISTICALLY SIGNIFICANT SECOND ORDER ATTRIBUTE DATA

$$ca_1^C a^A = p_1^C = .52 \tag{70}$$

$$c[b + a_2^A a_2^C + a_3^A a_2^C] = p_2^C = .48 \tag{71}$$

$$cb = (p_{12}^{AC})_{data} = \frac{N_{12}^{AC}}{N} = .219 \tag{72}$$

where we have defined

$$c = a_0 a^B \tag{73}$$

$$b = a_1^A a_2^C a_{12}^{AC} \tag{74}$$

Observe that if we add Equations 68 and 69 we obtain Equation 64. Thus Equations 68 and 69 do not contribute to the evaluation of the other a's. This result is to be expected as the B attribute is not part of our latest constraint.

Incorporating Equations 57 - 59, Equations 64 - 74 can be put in the following order for iteratively solving for the a values.

$$b = \frac{.219}{c} \quad \text{from Equation 72} \tag{75}$$

$$c = \frac{.38}{a_1^A a_1^C + b} \quad \text{from Equation 65} \tag{76}$$

$$a_2^A = \frac{.33}{c} a^C \quad \text{from Equation 66} \tag{77}$$

$$a_3^A = \frac{.29}{c} a^C \quad \text{from Equation 67} \tag{78}$$

$$a_2^C = \frac{(\frac{.48}{c} - b)}{a_2^A + a_3^A} \quad \text{from Equation 71} \tag{79}$$

$$a_1^C = \frac{.52}{c} a^A \quad \text{from Equation 70} \tag{80}$$

$$a^C = a_1^C + a_2^C \quad \text{from Equation 59} \tag{81}$$

$$a_1^A = \frac{1/c - b - (a_2^A + a_3^A)a^C}{a_1^C} \quad \text{from Equation 64} \tag{82}$$

$$a^A = a_1^A + a_2^A + a_3^A \quad \text{from Equation 57} \tag{83}$$

$$a_1^B = \frac{.13}{a_0[a_1^A a_2^C + b + (a_2^A + a_3^A)a^C]} \quad \text{from Equation 68} \tag{84}$$

$$a_2^B = \frac{.87}{.13} a_1^B \quad \text{from Equation 69} \tag{85}$$

$$a^B = a_1^B + a_2^B \quad \text{from Equation 58} \tag{86}$$

$$a_0 = \frac{c}{a^B} \quad \text{from Equation 73} \tag{87}$$

Starting with the initial values given by Equation 60 for the a's without the N_{12}^{AC} constraint, Equations 75 - 83 are iteratively solved, in the given order, to obtain a new set of a values that

Iteration	b	c	A a2	A a3	C a2	C a1	C a	A a1	A a	B a1	B a2	B a	a0
0		1	.33	.29	.48	.52	1	.38	1	.13	.87	1	1
1	.219	.91	.36	.32	.45	.55	1	.36	1.04				
2	.24	.85	.39	.34	.44	.59	1.03	.31	1.04				
3	.26	.86	.37	.33	.43	.58	1.01	.34	1.04				
4	.26	.84	.39	.34	.43	.59	1.02	.31	1.04				
5	.26	.85	.38	.33	.43	.59	1.02	.33	1.04	.14	.91	1.05	.81
6										.15	.95	1.10	.77
7										.15	.95	1.10	.77

Table 2: ITERATIVE CALCULATION OF a VALUES TO SATISFY THE N_{12}^{AC} CONSTRAINT

satisfy the N_{12}^{AC} constraint. Then a_1^B, a_2^B, a^B, and a_0 are calculated from Equations 84 - 87 using these values.

Table 2 presents the results of the iterations to find the new a values. After convergence, new values of $(m2-m1)$ are calculated using the probabilities determined from these latest a values. The resultant new highest significant N_{uv} is then selected and its associated probability added as a new constraint. Then, starting with the last previously calculated a values, a new set of a values is iteratively calculated to satisfy this additional constraint. This procedure is repeated until all the significant second-order probabilities are accounted for. Starting with the resultant latest set of a values, the procedure is then repeated for the next higher-order combination of attributes, etc. The overall procedure for finding significant correlations is outlined by the flow diagram shown in Figure 3. The procedure for calculating the a values is outlined by the flow diagram shown in Figure 4.

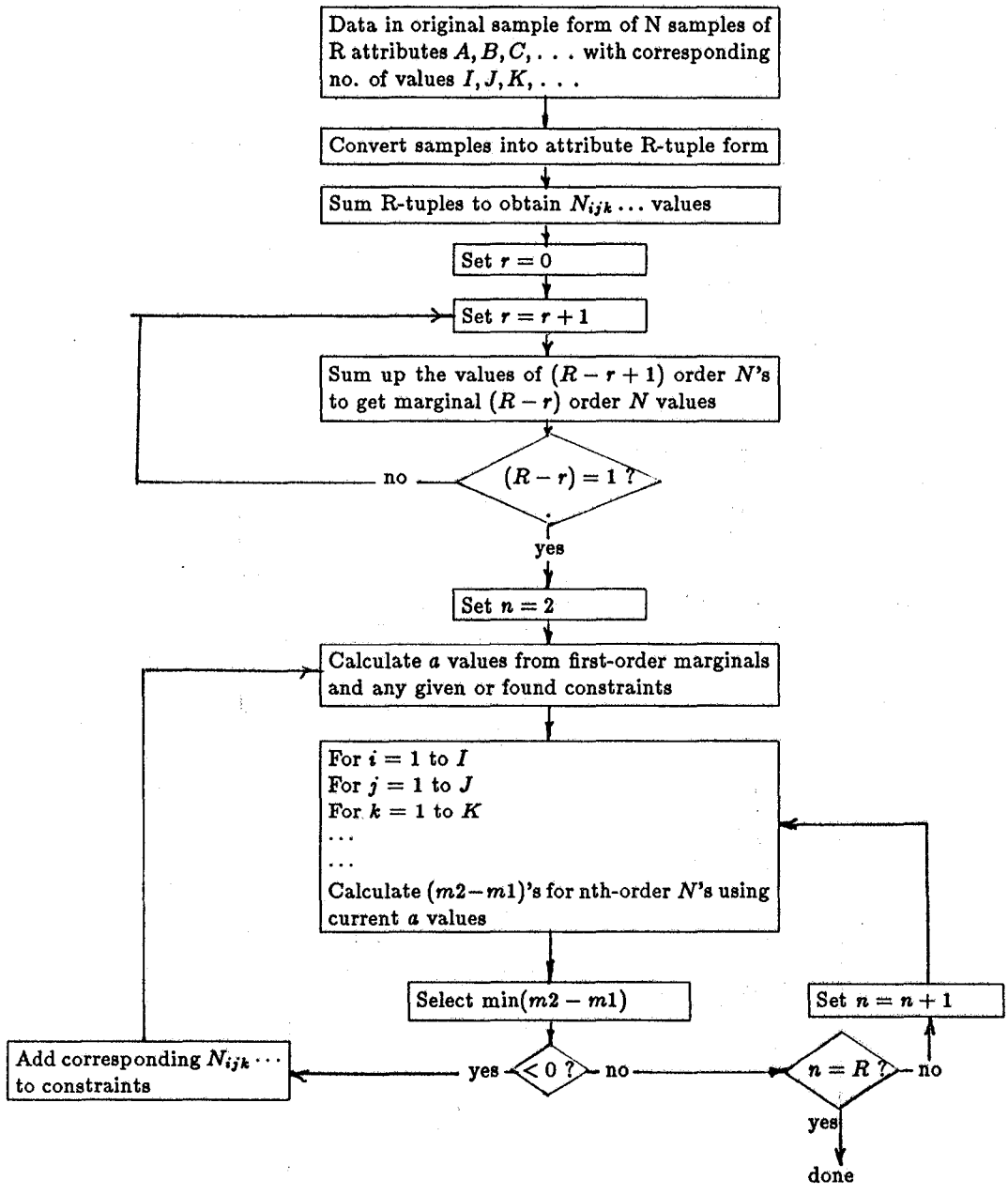


Figure 3: OVERALL PROCEDURE FOR FINDING SIGNIFICANT CORRELATIONS

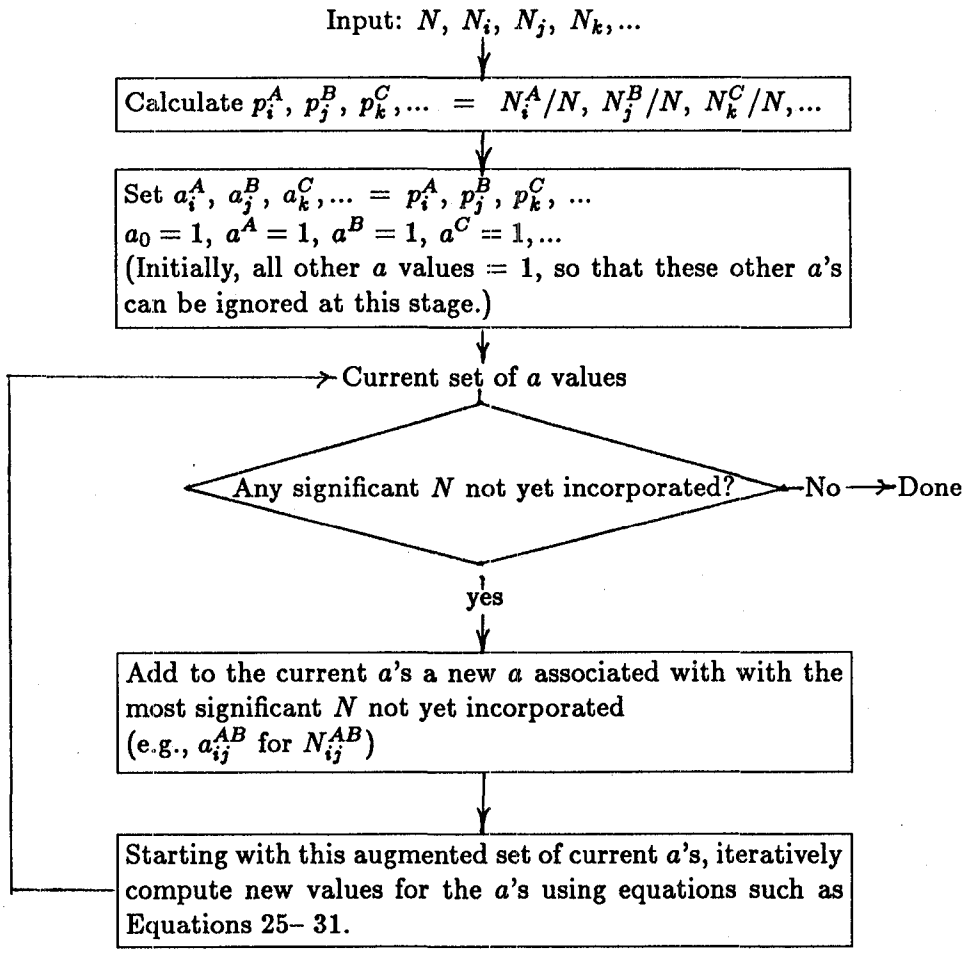


Figure 4: CALCULATING a VALUES

REFERENCES

- Boose, J. H.; Personal Construct Theory and the Transfer of Human Expertise, *Proc. National Conference on Artificial Intelligence*, Austin, TX, Aug 6-10, 1984, pp. 27-33.
- Cheeseman, P. C.; A Method of Computing Generalized Bayesian Probability Values for Expert Systems, *Proc. Eighth International Conference on Artificial Intelligence*, Karlsruhe, W. Germany, Aug 8-12, 1983, pp. 198-202.
- Cheeseman, P. C.; Learning of Expert Systems Data, *Proc. IEEE Workshop on Principles of Knowledge Based Systems*, Denver, Dec 3-4, 1984, pp. 115-122.
- Derfler, F. J.; Expert-Ease Makes Its Own Rules, *PC*, Vol. 4, No. 8, April 16, 1985, pp. 119-124.
- Jaynes, E. T.; Where Do We Stand On Maximum Entropy, *The Maximum Entropy Formalism*, Levine, R. D. and Tribus, M., Eds., MIT Press, Cambridge, 1979. pp. 15-118.
- Kahn, G.; Nowlan, S; and McDermott, J.; MORE: An Intelligent Knowledge Acquisition Tool, *Proc. Ninth International Joint Conference on Artificial Intelligence*, Los Angeles, CA, 1985, pp. 581-584.
- Langley, P.; Bradshaw, G. L; and Simon, H. A.; Rediscovering Chemistry with the BACON System, in *Machine Learning*, 1983.
- Langley, P.; and Carbonell, J. G.; *Approaches to Machine Learning*, Carnegie Mellon University, Pittsburg, CMU-CS-84-108, Feb. 16, 1984.
- Lenat, D. B.; *The Role of Heuristics in Learning: Three Case Studies*, Tioga Publishing Co., Palo Alto, CA, 1982.
- Michalski, R. S.; Carbonell, J. G.; and Mitchell, T. M.; Eds.: *Machine Learning*, Tioga Publishing Co., Palo Alto, CA, 1983.
- Michalski, R. S.; Carbonell, J. G.; and Mitchell, T. M.; Eds.: *Machine Learning: Volume II*, Morgan-Kaufman, Los Altos, CA, 1986.
- Williams, C.; *ART, The Advanced Reasoning Tool - Conceptual Overview*, Infererence Corp., Los Angeles, CA, 1985.
- Intellicorp: *Technical Summary of the Knowledge Engineering Environment (KEE) System*, Mt. View, CA, 1985.
- General Research Corp.: *TIMM - Users Manual*, Santa Barbara, CA, Nov 1985.

	A			B		C	
Sample No	1	2	3	1	2	1	2
1		x		x			x
2	x				x	x	
3		x			x		x
4			x	x		x	

Figure 5: ORIGINAL DATA FORM

Sample No	ABC 111	ABC 121	ABC 112	ABC 122	ABC 211	ABC 221	ABC 212	ABC 222	ABC 311	ABC 321	ABC 312	ABC 322
1							x					
2		x										
3								x				
4									x			
Etc.												
	N_{111}^{ABC}	N_{121}^{ABC}	N_{112}^{ABC}	N_{122}^{ABC}	N_{211}^{ABC}	N_{221}^{ABC}	N_{212}^{ABC}	N_{222}^{ABC}	N_{311}^{ABC}	N_{321}^{ABC}	N_{312}^{ABC}	N_{322}^{ABC}
sum =	130	410	110	640	62	580	31	460	78	520	22	385

Figure 6: SAMPLE DATA IN TRIPLES FORM

APPENDIX A

CONVERTING ORIGINAL DATA TO CONTINGENCY TABLE FORM

For our sample problem, the original data in response to the questionaire might be in the form of Figure 5 (or can readily be placed in that form).

Put into attribute triples form, this data might appear in terms of ijk values of the attributes as shown in Figure 6. Note that the summations of the triples are the values of the cells in Figure 1.

APPENDIX B

CALCULATING SUM OF PRODUCTS EQUATIONS INVOLVING a's

From Equation 12, we have the basic equation for the nth order probability as:

$$p_{ijk}\ldots = a_0 a_i a_j \ldots a_{ij} \ldots \tag{88}$$

based upon which the basic equations for the a's and the p's are given by Equations 21 - 24. These equations all involve the summations of products of a's. If we order these summations, we obtain equations such as Equation 25

$$\frac{1}{a_0} = \sum_i a_i \sum_j a_j a_{ij} \sum_k a_k a_{ik} a_{jk} \tag{89}$$

A convenient way to handle such summation products is to introduce matrices.

Let us define the operator X indicating comparable term-by-term matrix multiplication. Thus for example

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} X \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & 3b \\ 2c & 4d \end{bmatrix} \tag{90}$$

The summation operator, $\sum$, can be considered to be summing terms in such matrices in the following manner

$$M_j = \sum_i M_{ij} = \sum_i \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} = \begin{bmatrix} m_{11}+m_{21} & m_{12}+m_{22} \end{bmatrix} \tag{91}$$

$$M_i = \sum_j M_{ij} = \sum_j \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} = \begin{bmatrix} m_{11}+m_{12} \\ m_{21}+m_{22} \end{bmatrix} \tag{92}$$

In this notation, Equation 89 can be written as

$$\begin{aligned} \tfrac{1}{a_0} &= \textstyle\sum_i a_i \sum_j a_j a_{ij} \sum_k a_k a_{ik} a_{jk} \\ &= \textstyle\sum_i (Q_i \;\; X \;\; \sum_j (Q_{ij} \; X \; \sum_k Q_{ijk})\,) \end{aligned} \tag{93}$$

In Equation 93, for

$$I = 3,\ J = 2,\ \text{and}\ K = 2 \tag{94}$$

(corresponding to the number of values of the attributes A, B, and C in our example)

$$\sum_k Q_{ijk} = S_{ij} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \\ s_{31} & s_{32} \end{bmatrix} \tag{95}$$

where

$$s_{ij} = a_1^C a_{i1}^{AC} a_{j1}^{BC} + a_2^C a_{i2}^{AC} a_{j2}^{BC} \tag{96}$$

and

$$Q_{ij} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \\ q_{31} & q_{32} \end{bmatrix} \tag{97}$$

where

$$q_{ij} = a_j a_{ij} \tag{98}$$

and

$$Q_i = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} \tag{99}$$

where

$$q_i = a_i \tag{100}$$

Using this notation, Equation 93 can be written as

$$\begin{aligned} \tfrac{1}{a_0} &= \textstyle\sum_i a_i \sum_j a_j a_{ij} \sum_k a_k a_{ik} a_{jk} \\ &= \textstyle\sum_i (Q_i \ X \ \sum_j (Q_{ij} \ X \ \sum_k Q_{ijk})) \\ &= \textstyle\sum_i (Q_i \ X \ \sum_j (Q_{ij} \ X \ S_{ij})) \\ &= \textstyle\sum_i (Q_i \ X \ S_i) \\ &= S \end{aligned} \tag{101}$$

where

$$S_i = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \tag{102}$$

$$s_i = q_{i1} s_{i1} + q_{i2} s_{i2} \tag{103}$$

and

$$S = s_1 + s_2 + s_3 \tag{104}$$

Observe the recursive nature of these equations by noting that in general

$$S_n = \sum_{n+1 \ index} (Q_{n+1} \ X \ S_{n+1}) \tag{105}$$

where n is the order of the matrix and the nth index is the index of the nth attribute.

For R being the highest order of the attributes ($R = 3$ in our example),

$$S_R = [I] \text{ , a unity matrix in our notation so that} \tag{106}$$

$$Q_R \ X \ S_R = Q_R \tag{107}$$

If using Equation 88, we desire to calculate a probability, p_k, then Equation 88 takes the form of Equation 28

$$a_0 \, a_k \sum_i a_i a_{ik} \sum_j a_j a_{ij} a_{jk} = p_k \tag{108}$$

or in our matrix notation

$$a_0 \, Q_k \sum_i (Q_{ik} \ X \ \sum_j Q_{ijk}) = P_k = \left[\begin{array}{cc} p_1^C & p_2^C \end{array} \right] \tag{109}$$

where

$$q_{ijk} = a_j a_{ij} a_{jk} \tag{110}$$

$$q_{ik} = a_i a_{ik} \tag{111}$$

$$q_k = a_k \tag{112}$$

$$p_k = p_k \tag{113}$$

Thus, for our example

$$Q_{ik} = \left[\begin{array}{cc} a_1^A a_{11}^{AC} & a_1^A a_{12}^{AC} \\ a_2^A a_{21}^{AC} & a_2^A a_{22}^{AC} \\ a_3^A a_{31}^{AC} & a_3^A a_{32}^{AC} \end{array} \right] \tag{114}$$

and

$$S_{ik} = \left[\begin{array}{cc} a_1^B a_{11}^{AB} a_{11}^{BC} + a_2^B a_{12}^{AB} a_{21}^{BC} & a_1^B a_{11}^{AB} a_{12}^{BC} + a_2^B a_{12}^{AB} a_{22}^{BC} \\ a_1^B a_{21}^{AB} a_{11}^{BC} + a_2^B a_{22}^{AB} a_{21}^{BC} & a_2^B a_{21}^{AB} a_{12}^{BC} + a_2^B a_{22}^{AB} a_{22}^{BC} \\ a_1^B a_{31}^{AB} a_{11}^{BC} + a_2^B a_{32}^{AB} a_{21}^{BC} & a_2^B a_{31}^{AB} a_{12}^{BC} + a_2^B a_{32}^{AB} a_{22}^{BC} \end{array} \right] \tag{115}$$

If, as an example, none of the w_{ik}'s are significant, then we replace the corresponding a_{ik}'s with 1's, as from Equation 13

$$a_{ik} = e^{-w_{ik}} \tag{116}$$

In this case, Equation 114 reduces to

$$Q_i k = \left[\begin{array}{cc} a_1^A & a_1^A \\ a_2^A & a_2^A \\ a_3^A & a_3^A \end{array} \right] \tag{117}$$

so that

$$S_k = \sum_i (Q_{ik} \;\; X \;\; S_{ik}) \tag{118}$$

becomes

$$S_k = \left[\begin{pmatrix} a_1^A(a_1^B a_{11}^{AB} a_{11}^{BC} + a_2^B a_{12}^{AB} a_{21}^{BC}) \\ +a_2^A(a_1^B a_{21}^{AB} a_{11}^{BC} + a_2^B a_{22}^{AB} a_{21}^{BC}) \\ +a_3^A(a_1^B a_{31}^{AB} a_{11}^{BC} + a_2^B a_{32}^{AB} a_{21}^{BC}) \end{pmatrix} \quad \begin{pmatrix} a_1^A(a_1^B a_{11}^{AB} a_{12}^{BC} + a_2^B a_{12}^{AB} a_{22}^{BC}) \\ +a_2^A(a_2^B a_{21}^{AB} a_{12}^{BC} + a_2^B a_{22}^{AB} a_{22}^{BC}) \\ +a_3^A(a_2^B a_{31}^{AB} a_{12}^{BC} + a_2^B a_{32}^{AB} a_{22}^{BC}) \end{pmatrix} \right] \tag{119}$$

$$= [\, s_1 \quad s_2 \,]$$

and

$$\begin{aligned} P_k &= [\, p_1^C \quad p_2^C \,] \\ &= [a_0\, a_1^C s_1 \quad a_0\, a_2^C s_2] \end{aligned} \tag{120}$$

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-88224		

4. Title and Subtitle	5. Report Date
AUTOMATIC PROBABILISTIC KNOWLEDGE ACQUISITION FROM DATA	April 1986
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
William B. Gevarter	A-86154
	10. Work Unit No.
9. Performing Organization Name and Address	T-6125
Ames Research Center Moffett Field, CA 94035	11. Contract or Grant No.
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Technical Memorandum
National Aeronautics and Space Administration Washington, DC 20546	14. Sponsoring Agency Code 506-45-2

15. Supplementary Notes

Point of Contact: William B. Gevarter, Ames Research Center, MS 244-7, Moffett Field, CA 94035 (415) 694-6525 or FTS 464-6525

16. Abstract

This memorandum documents an outline for a computer program for extracting significant correlations of attributes from masses of data. This information can then be used to develop a knowledge base for a probabilistic "expert system." The method determines the "best" estimate of joint probabilities of attributes from data put into contingency table form. A major output from the program is a general formula for calculating any probability relation associated with the data. These probability relations can be utilized to form IF-THEN rules with associated probability, useful for expert systems.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Artificial intelligence Knowledge acquisition Expert systems Machine learning	Unlimited Subject Category - 59

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	30	A03

*For sale by the National Technical Information Service, Springfield, Virginia 22161

End of Document