EXPERIMENTS WITH A SEARCH ALGORITHM FOR THE DATA BASE OF A HUMAN BELIEF STRUCTURE

Kenneth Mark Colby-
Lawrence Tesler
Horace Enea

Summary

A large data base was collected from a human informant. The data consisted of beliefs regarding parent-child relations. A variety of factors in searching the data base were manipulated in an attempt to discover which were the more important in contributing to estimates of credibility. Problems of data collection, data representation and a searching algorithm are discussed in detail.

Introduction

It is clear that people hold beliefs. What is not so clear is how these beliefs are processed to judge the credibility of an input proposition. As an aid in understanding credibility processes, we constructed a computer model which was intended to simulate the belief processes of a particular human informant. We shall begin with a clarification of terminology used in describing the model.

Terminology

A model consists of a set of interacting components. The major components of the model to be described are state-descriptions (data base), process-descriptions (procedures) and an interpreter whose logic governs the applications of process-descriptions to state-descriptions in accordance with the aim or task of judging credibility. Thus the model is two levelled with the interpreter at the top supervising interactions between procedures and data base.

The data base is made up of a conceptual graph whose basic structures consists of conceptualizations in turn composed of elementary conceptions. Conceptualizations in the model are held propositions which symbolically represent states of affairs or situations. Conceptualizations can be represented in both natural and computer languages by an ordered set of name-tokens. Thus a given conceptualization might be described in English as 'Bill likes Mary' and in a programming language as the list ((agent Bill) (action like) (object Mary)). The particular conceptualizations we focussed on were those which semantically involved certain relations between humans. Since people conceptualize their experience with persons in terms of human action, the elementary conceptions of the data base involve agents, actions, objects and (optionally) settings, modalities and rationales.

The term 'belief' in human belief structures refers to (a) an affective attitude of acceptance, rejection or neutral judgment towards, (b) a held conceptualization. Each conceptualization held or prehended by a belief structure is either accepted to some degree as true, rejected to some degree as false or held in suspension as a neutral candidate for belief. 'True' here means that the situation conceptualized is accepted as being certainly, probably, or possibly being the case while 'false' stands for the opposite of these three modal quantifiers. It is important to note that the attitude of rejection or incredibility is towards a conceptualization prehended within a structure. A disbelieved conceptualization is not expelled from the structure but is prehended with an attitude of rejection. A conception is thus judged to be credible, incredible or somewhere in between.

We postulate credibility to be a function of foundation and consistency. The foundation of a given belief is a measure of those beliefs which imply it as opposed to its negation. Consistency refers to a degree of consonance and dissonance found in those beliefs a given belief implies. The term 'imply' does not refer to logical implication but to psychological implication which involves rules of expectancy. We also assume a weight which determines the relative importance foundation and consistency have for one another in a particular domain of interest.

Conceptualizations with their associated credibilities make up one component of the data base. We term these conceptualizations 'facts' since they stand for that which is or is not the case in the structure. A second component of the data base consists of rules. By the term 'rule' we mean a connectivity relation holding between two or more conceptualizations. The components of this relation contains variables as well as name-tokens. Hence, a rule might read 'if x likes y then x helps y', where both x and y are variables to which the name-tokens of persons can be bound. As mentioned, this if-then relation represents a type of psychological implication. Our interpretation of psychological implication is that given conceptualization A to be the case, conceptualization B is expected to be the case by the structure. For example, humans commonly expect that if a person likes another person, the first person will help the second person. Such general expectancy rules allow a variety of inference processes to be carried out. A further discussion of psychological implication can be found in Abelson[2].

The remaining components of the data base consist of definitions and classifications. For example, the name-token 'love[1] is defined as similar-to and stronger-then 'like'. 'Dislike[1] is definable as negation-of 'like'. These definitions are used in finding similarities and contrasts between conceptualizations. Classifications consist of set memberships and set inclusions.

In summary, the data base represents information in the form of various kinds of state descriptions. When the model runs, this information is subjected to procedures
governed by a top-level interpreter. The procedures are called into operation by the interpreter in accordance with the task involved. The main task we were interested in involves estimating the credibility of a given proposition describing some actual or hypothetical situations. Given such a proposition, how might a person judge its credibility using the information-processing capacities he has available?

We approached this problem by selecting an informant, collecting certain beliefs, and representing them in a data base. We planned first to conduct certain information-processing experiments on the data base and second to attempt a validation of the simulation. This report will be concerned with the first phase of experimentation.

Data Collection

To find an informant for this research, we advertised in a college newspaper for persons who might be interested. Out of 65 applicants, we interviewed 26 and then selected a 30 year old married woman on grounds that she was intelligent, articulate, interested in the research and serious of purpose.*

Several times a week she would write down in natural language beliefs which occurred to her about events in her life. Each week we would try to reduce these natural language statements to a simpler form suitable for the model's data base and processing. At intervals we would show the data base to the informant for her corroboration or correction of our paraphrasings. Initially we planned to obtain her beliefs regarding all the important people in her life space. Preliminary experience showed that while collection of such data is possible, the labor required to organize and represent this amount of data in a computer model makes the task extremely difficult with currently available methods.

We found two main disadvantages to this method of data-collection. First, it is cumbersome and time-consuming, requiring hundreds of man-hours to obtain a data base of 700 facts, rules, definitions and classifications. A better method should be developed whereby an informant could type information directly into a data-base by means of a man-machine dialogue. This input might be in an artificial and simplified language which an informant could learn. However an artificial language is distracting and constrains expressiveness. It would be better to allow the rich freedom of natural language but there are great problems involved in the machine handling of this sort of input. We have had some experience along these lines3 and we are continuing an attempt to make a conceptual parsing of natural language sentences in order to translate them into conceptualizations suitable for the data base of a belief structure •

A second drawback to our initial data collection method involved the problem of extensiveness versus intensiveness. An extensive data base is one in which there are a great variety of conceptualizations but not a great number around any one theme. While containing a large amount of information, this type of data is too sparse to permit the model to come to very many conclusions regarding credibility. Few beliefs of relevance can be found for a given proposition unless it is of a very general and hence trivial nature.

We then attempted to concentrate on a particular theme in order to make the data base dense around selected conceptions. Since our informant was the mother of a three year old child and interested in the problem of child raising, we concentrated on her beliefs in this area. For each of her beliefs in this domain, we obtained a weighting of a degree of credibility. We used crude categories of strong, medium and weak for these weightings. To obtain data rules we would ask the informant for reasons for each belief. For example, if a belief were 'a child ought not hit another child' and the reason given by the informant was 'because if a child hits another child then the second child gets hurt', a general if-then expectation rule can be constructed about the relation between hitting and hurting.

One difficulty to be anticipated in simulating a human belief structure involves keeping the model updated along with the informant. If a person's beliefs are continually changing, one cannot keep a model in close enough correspondence to test out comparisons between the person's and the model's performances in estimating credibility. However, we found that with our informant, these particular beliefs about parent-child behavior changed very little over a period of several months. In the case of only two beliefs did she change a credibility weight from strong to medium. Hence the structure appeared quite stable over this period of time. It should also be remarked that there occurred no major environmental event in this domain of interest which might be expected to have great impact on a belief structure.

We attempted to model the credibility processes of a single individual. This approach is in the research tradition of an intensive design in contrast to an extensive design. An extensive design might make one observation on 1000 persons while an intensive design would make 1000 observations on one person. Both designs attempt to account for variation in the phenomena observed. In an extensive design, the unit of variability is an individual and variation between individuals is studied whereas in

*For obvious reasons we cannot give her name, but we would like to use this footnote to express our gratitude for her help.

an intensive design we are studying the variability within an individual. In modelling a single case we are trying to understand the mechanisms involved in intra-individual processes. An intensive design attempts to show what can and does happen. The frequency of this sort of happening in a population and which population is another matter. After learning how to model one person we can model another and so build up a series of cases. The inductive problem of generalizing then becomes one of sampling and of statistical measures to discover how general the informant's beliefs might be in a population. This was not our problem at this state of the inquiry. Our problem was how to construct a good model of the informant's belief processes. The criteria for 'good' can be varied. And are we getting at what the informant 'really' believes? What 'really' means here is obscure but it is common knowledge that people have limited accessibility to their beliefs at a given moment. Even worse, they have the capacity to deceive themselves to rationalize, and to distort their own beliefs. Over time we hoped to increase accessibility, realizing there are always limits. In worrying about what is 'really' believed we found it useful to keep in mind that we were constructing a model of a model. A belief structure is a representation and in giving information about himself, an informant tells us what he believes he believes. He simulates himself and it is his accessible model of himself that becomes the data base of a computer model. Humans' ability to simulate themselves and to make models of other models is of course a most interesting property for a symbolic system to have.

Data Representation

In building a data base for the model we thought of the collected facts, rules, definitions and classifications as a graph. Physically in the model they were lists in the programming language MLISP[5,6]. MLISP is a high level list processing language which translate ALGOL-like meta-statements (M-expressions) into the symbolic statements (S-expressions) of LISP 1.5. The program runs on the PDP-6/10 time-sharing system of the Stanford Artificial Intelligence Project.

Each conceptualization was represented as a list of elements consisting of English-like name tokens or atoms, of lists of atoms, and of lists of lists which contained semantic and numerical informationt As mention in the section 'Terminology', the conceptualizations reflect a human action model of situations in the interpersonal world. From this perspective, agents carry out actions towards objects which in turn can be agents or other situations. In the data base each agent, action, object, etc., is identified by an atom. On the property list of each atom is a list of pointers to all conceptions in the data base in which that particular atom occurs. A hash coding scheme is used for rapid look-ups and retrieval of relevant conceptions.

The representation of a fact such as 'Barb likes children[1] appears on the list:

(F(agent Barb) (action like) (object children) (credibility 0.9)) with the symbol F indicating this is a fact. More complicated is the representation of a rule because of the problem of binding variables unambiguously. For example, the natural language statement 'parents spank children' has a number of possible semantic interpretations. It was necessary to check carefully with the informant in order to be clear about which interpretation she intended. In this case she did not mean that the set 'parents' spank the set 'children' nor did she mean each member of the set 'parents' spank each member of the set 'children'. By the expression 'parents spank children' she meant that a given parent spanks his children. More formally, if x is a parent and y is a child and x is a parent of y, then x spanks y.

The relation 'is a parent of must first be defined in terms of certain constrained variables. For instance, the variable P .is defined as a parent who is a parent of C_{x4} and C_{x4} is defined as a child who is a child of P_{x4}. When assignments are made to such variables only those name tokens which qualify can be substituted. Thus the rule 'if a parent likes his child, then his child is happy' would be represented as the list:

(R($P_{x4}$$C_{x4}$)(((agent P_{x4}) (action like) (object C_{x4}))

((agent C_{x4}) (action has) (object happiness))))

with the symbol R indicating that it is an implications! rule. When facts are search to match the components of a rule, the fact 'John likes Mary' would fit this rule only if John is held to be a parent of Mary so that 'John' can be substituted for P_{x4} and 'Mary' for C_{x4}. These constrained variables are global in the program. They permit the binding of variables to be unambiguous and allow rules to be arbitrarily complex since the qualifications required for the variable may involve multiple conditions.

Representation of definitions is in the form of a simple list. To indicate conceptual relations between 'love' and 'like', the list appears as

(love similar like S)

where the symbol S indicates 'love' is stronger in intensity than 'like'. The following are some relations represented in definitions, S meaning stronger, W weaker, and E equal.

similar	(love similar like S)
different	(men different women)
negative	(notlike negative like E)
opposite	(love opposite hate E)
kindof	(spank kindof aggression)

Classifications take the simple form,

(F(agent matches) (action is a) (object things))

Initial experience with a data base of 700 facts, rules, definitions and classifications not only taught us about the density requirements of data but also brought to light an implementation problem. When output from a running model is not satisfactory, it may be due to errors in the data as well as to inadequacies of the procedures.■ A small data error, (such as the term 'notlike'. in a conceptualization instead of 'like'), originating from human mistakes in inputting data into the data base, can give rise to an incorrect credibility estimate. When a data base is large, it becomes extremely difficult to trace entirely by hand what happened in a given run of the model. We tried frequently to check the data for errors and the informant repeatedly studied a listing of the data-base searching for mistakes. In spite of this labor of scrutiny, disconcerting data errors would still crop up. To make sure the procedures were operating as postulated, we first selected a very restricted subset of the data base and then gradually added to it as the program became debugged.

Procedures

The modelling program scales variables such as credibility, foundation, and consistency into the range 0 to 100. The interpretation of these numbers is as follows:

90-100	Strongly positive
6O-89	Weakly to moderately positive
41-59	Undecided
11-40	Moderately to weakly negative
0-10	Strongly negative.

Credibility is a function of two components: foundation and consistency.

The foundation of a proposition is a measure of the model's evidence for and against the proposition. If the positive evidence outweighs the negative, then the foundation is high; if the negative evidence is stronger, the foundation is low.

Credibility is a function mainly of foundation. When foundation is moderate, consistency has more influence on credibility than when foundation is extreme. Thus, if the evidence concerning a proposition is not dominantly pro or con, then the model gives extra weight to its consistency in determining its credibility. A formula for credibility which incorporates this factor is given below.

$$\text{credibility} = \text{foundation} + (\text{consistency}-50) \times (50 - |\text{foundation}-50|) \times \text{weight}$$

The "weight" is a number between 0 and 0.02 which indicates the relative importance of consistency in this computation. If consistency is irrelevant, the weight is zero. If it is dominant, the weight is 0.02.

EXAMPLES

CREDIBILITY as a function of Foundation and Consistency

	Consistency				
	100	80	50	20	0
Foundation					
80	90	86	80	74	70
50	75	65	50	35	25
20	30	26	20	14	10

The consistency of a proposition P is computed by finding a few highly relevant beliefs and measuring the consonance of P with these beliefs. Relevance is defined objectively. If P is a proposition of the predicate form f(p) and if there is a rule in the model that says:

if f(x) then g(x)

and if g(p) is the predicate form of a belief Q_i, then Q_i is highly relevant to P. If the model already disbelieves Q_i. — or believes -1 Q_i. — then P is dissonant with it. If the model neither believes or disbelieves Q_i. then the consistency of P is not affected.

The computation of consistancy consists of determining the percentage of Q_i.'s with which P is consonant. More weight is given to consonance with more credible Q_i.'s than to consonance with less credible Q_i.'s. A set of formulas that incorporate this weighting is given below.

sc = Z credibility (Q_i.), where P Q_i. and

credibility (Q_i) > 50

cc= count of Q_i's contributing to sc

cn = count of Q_i 's where P o -1 Q_i. and

credibility (q_i) > 50

$$\text{consistency} = \frac{sc + 1}{cc + en + .02}$$

The foundation of a proposition P is computed by finding relevant beliefs and seeing whether they imply that P is or is not the case. In the search for relevant beliefs, graph paths through beliefs consonant with P are searched harder than paths through beliefs dissonant with P if P seems to be highly consistent. The reverse strategy is used if P seems to be inconsistent. This is done so that model can attempt to limit its search for evidence in such a way as to maintain the consistency of its entire belief structure.

Formulas for foundation in terms of evidence for and against P are:

$sc = \Sigma$ credibility (Q_i), where $Q_i \supset P$ and credibility $(Q_i) > 50$

cc = count of Q_i's contributing to sc

cn = count of Q_i's where $Q_i \supset \neg P$ and credibility $(Q_i) > 50$

$$\text{foundation} = \frac{sc + 1}{cc + cn + .02}$$

The search for relevant beliefs is controlled by a "work" factor. A consistency search will do, say, 200 units of work while a foundation search will do, say, 1000 units of work. This work allotment is apportioned among the possible graph paths that lead from the proposition in question to relevant areas of the graph.

The algorithm for searching is as follows. The directly relevant beliefs in the graph are found. A directly relevant belief is one which can be derived from P in one step by any one of these methods:

1) Replace the verb by a similar (or opposite) verb.

2) Replace the subject or object by an analogous (or complementary) noun.

3) Replace the predicate adjective by a similar (or opposite) adjective.

4) Generate a belief which implies P (or P implies) according to any one rule.

These beliefs are the heads of paths to be searched. A certain amount of work is used up just in finding them; say, 2 units for each relevant rule used, 2 for each step of an analogy that is drawn, and 3 for each similar verb that is found, plus 6 units overhead even if nothing relevant is found.

If there is any work that remains unused after finding the heads of these paths, it is divided up among the paths for further searching.

In the consistency search, all paths receive equal treatment. However, in the foundation search, the division among paths is affected by consistency. To compute the consistency of P with these paths, a recursive short-depth search is performed along each path; these searches are alotted, say, 1/3 of the remaining work. From the resulting consistencies, proportions are computed according to the following formulas:

eg = consistency of P with whole system

cp = consistency of P with this path

$$\text{relative proportion} = \frac{110 - |cg - cp|}{110 - |cg - 50|}$$

The paths with highest proportions are searched first and receive a proportionately greater work allotment.

If not all the work along a path is exhausted, the remainder is divided among the remaining paths. If, after searching any path, enough relevant beliefs have been found to compute a credibility exceeding 60 or below 40, then the search of the rest of the paths is cancelled.

Experiments

The program performs two major experiments. The first experiment assumes that the belief structure is unchanging. One proposition at a time is presented to the structure and its credibility is judged. In the second experiment, the belief structure does change. After each proposition's credibility has been evaluated, it becomes incorporated into the structure as a belief.

The first experiment is run by presenting each belief in the structure to all the other beliefs and judging its credibility. The result can be compared with the prestipulated credibility of the belief. Then, a list of new propositions is presented to the structure for evaluation. In both experiments, many factors of the evaluation are varied.

One factor to vary is the means of finding relevant beliefs. There are four variations:

1) Use only rules - no definitions.

2) Use (1) plus rules to find supersets.

3) Use (1) and (2) plus "similar" and "opposite" rules.

4) Use (1), (2), and (3) plus rules to find instances.

Another factor to vary is the use of consistency. There are two variations:

1) Use foundation and not consistency.

2) Use also consistency.

Other factors varied are:

1) Weight of consistency relative to foundation in computing credibility.

2) Amount of work expended in search.

3) The initial credibilities assigned each belief.

All values of these factors are combined with every meaningful combination of other factors.

Experimental Results

For the data base used in the experiments so far, a few interesting results were obtained.

The search for relevant beliefs was effective when both rules and supersets were utilized in the search. Without supersets, many relevant beliefs were missed. The addition of similar and opposite rules expanded the search enough to discover nearly all beliefs considered relevant by the experimenters. In only a handful of cases did the application of instance rules improve the relevance search.

The use or disuse of consistency made no noticeable difference in the credibility computation. It is planned to see whether consistency will make a difference with different data or with work allotments that have not yet been tried.

The amount of work alloted made a difference in the success of finding relevant beliefs. It is intended to measure this difference quantitatively, but techniques for this have not yet been developed.

Scaling the credibilities of all the beliefs in the system by a factor x seemed to affect the credibility computed for an input proposition by that same multiple, x. This showed that the complex search combined with the quotient formulas for credibility still preserved linearity.

Further values of the variable f .ctors are in the process of being tested, as well as improved searching algorithms.

Discussion

The only other program we know of which judges credibility is that of Abelson and Carroll[1]. There are a number of similarities and differences between the two programs. Perhaps the most important difference lies in the way the search algorithm is controlled. In the Abelson and Carroll program searches through a large data base are cut off probabilistically, depending on a random number exceeding some fixed value. In our model, three factors govern the search, consistency, work and firmness. Search along consistent paths is preferred to search along inconsistent paths. Also search along a path is cut off if (a) alloted work runs out, (b) a firm credibility of >60 or *<k0* is reached or (c) the path is exhausted.

Another interesting difference lies in the way the two programs treat instances and supersets. The Abelson and Carroll program looks "down" at instances and "up" at supersets to an equal degree. We found that with this data base searching for instances contributed to credibility less reliably than supersets. Therefore we alloted more work to searching supersets than to searching instances.

Our experiments with this data base collected from an informant constituted an attempt to discover what search factors made a significant difference in estimating the credibility of input propositions. We were not attempting to validate a particular search algorithm. Instead we explored a variety of procedures in an effort to learn more about their respective merits in processing the same data base. We intend to discuss the validation problem in a future report.

References

(I) Abelson, R.P. and Carroll, J., Computer simulation of individual belief systems. American Behavioral Scientist, 8, 24-30 (1965).

(2) Abelson, R.P., Psychological implication. In Theories of Cognitive Consistency (Abelson,R., Aronson, E., McGuire, W., Newcomb, T., Rosenberg, M., Tannenbaum, P. Edc. Rand-McNally, New York, 1969)

(3) Colby, K.M. and Enea, H., Heuristic methods for computer understanding of natural language in context restricted on-line dialogues. Mathematical Biosciences, 1, 1-25 (1967)-

00 Schank, R.C and Tesler, L.G. A conceptual parser for natural language. Stanford Artificia Intelligence Project Memo No. AI-76, 1969 (See also this volume).

(5) Enea, H., MLISP. Stanford Computer Science Department Technical Report, No. 92, March 14, 1965:

(6) Smith, D.C. MLISP Users Manual. Stanford Artifi cial Intelligence Memo, No. 1969. (In preparation.

Acknowledgements

This research is supported by Grant PHS MH 06645-07 from the National Institute of Mental Health, by (in part) Research Scientist Award (No. *K-l4,433)* from the National Institute of Mental Health to the senior author and (in part) by the Advanced Research Project Agency of the Office of the Secretary of Defense (SD-183).