Applications of Fuzzy Set Theory in Human Factors, edited by W. Karwowski and A. Mital
Elsevier Science Publishers B.V., Amsterdam, 1986 — Printed in The Netherlands

A Fuzzy Model for the Accumulation of Judgments by Human Experts

Alf C. ZIMMER
University of Regensburg, Universitätsstraße 31, 8400 Regensburg, W-Germany

ABSTRACT

A formal model for the justification and checking of arguments is proposed for the development of expert systems. The two major parts of the model consist in the detection of violations of the rules governing rational discourse and in the evaluation of conclusions in fuzzy syllogistic reasoning.

The application of the model consists in the interaction between the expert ('arguer') and the formal model. In the beginning the train of arguments is checked for rule violations and the expert ('arguer') is asked to mend his/her argument accordingly. Then the acceptable train of arguments, that is, a syntactically valid chain of propositions is evaluated according to its fuzziness. Upper and lower bounds for the acceptability of the train of arguments are calculated and fed back to the expert ('arguer') in either numerical or verbal form. After that the expert decides either to choose a level of acceptability in the given bounds or suggests an expansion of the justification. In the latter case the evaluation in both steps starts anew.

First experimental results in predictions for everyday events indicate that the model is viable, fits into the subjects' cognitions, and helps the subjects to arrive at trains of arguments which are more acceptable for them than their initial justifications.

INTRODUCTION

During the last 20 years the role of the human operator in technical settings has changed from immediate manipulation towards a more and more remote monitoring of the equipment and its behavior. Due to this fact, the task of the operator has become more abstract and the required skills pertain mostly to correct decision making, especially in critical situations for which no standard procedures exist (the TMI incident is a typical example for failures in decision making in such a control situation, for other examples see Reason & Mycielska, 1982).

In order to enable the operator to decide correctly even in situations not encountered before or not trained for, it is necessary that he or she has an internal representation, or more specifically, a model (see Johnson-Laird, 1980) of the equipment and process to supervise. There a two major requirement for such a model: (i) it has to be designed in a way consistent with the information-processing capacity and style of the operator, and (ii) it has to contain all states of the equipment and the relations between them that are essential for the operator to plan the necessary interventions ahead.

The results of cognitive psychology suggest that such a model has to be qualitative rather than quantitative not only in regard to the states but also in regard to the relations, that is, instead of a physically correct description in terms of differential equations descriptions of the following kind are to be embodied in a mental model: "a rapid increase of X in A normally causes a drop of pressure in B unless C is in state Y."

Forbus & Stevens (1981) have demonstrated that even for comparatively simple equipments and processes (e.g. steam engines for USN ships) a model in qualitative terms is not only sufficient for the operator to generate a complete control structure for the process, but also that it leads to less errors than an instruction outline based on the set of differential equations underlying the process. This and other results have fostered the development of 'qualitative physics' (for an overview see Gentner & Stevens, 1983) as a means for the development of tutoring systems tailored to the control structure of a given task. In experiments on probability assessment and prediction Zimmer (1983, 1984) has shown that for most people qualitative expressions for quantifiable events have not only a comparatively stable meaning but are also less prone to interference than numerical expressions. Furthermore, for people (even experts) it is easier to combine qualitative information from different sources together with their relations than to integrate quantitative data about components of a system which are functionally related.

There are three main conclusions to be drawn from the indicated evidence about the human operator in control and supervision tasks: (i) the desription of the system in question and the tutoring of the controller have to concentrate on the states of the system and their relations which are relevant for making decisions, (ii) these descriptions ought to be qualitative because in most cases the crucial decisions are about intervening or not intervening; and even the decisions about how to intervene are normally qualitative, too, (iii) the control panels have to be designed in such a way that the location of the instruments and controls are in accordance with the qualitative model.

PRINCIPLES FOR THE DEVELOPMENT OF INTERNAL MODELS

Unfortunately, there is no simple algorithm to translate quantitative into qualitative descriptions and even if this translation has been made it cannot be guaranteed that it fits the information-processing capacity of the human operator and enables him or her to develop a control structure that is simple and at the same time comprehensive enough to plan an intervention in a novel situation. For these reasons the way to come to a functional qualitative description is to find out what kind of internal models experts have and how they derive control decisions from this kind of representation. On the first glance, simply asking experts what they do, for instance, in landing a plane smoothely and safely, seems to be the appropriate approach. However, there are a couple of problems connected with this method. The main problem is the pre-knowledge (correct or incorrect) of the interviewer about the equipment and process in question; not only that it influences the kind of questions posed, furthermore the sequence and the content of the questions can lead to a biased report by the expert, that is, he or she conforms unwittingly to the implicit presuppositions contained in the questions. Of equal importance is that a human interviewer processes the obtained information only partially, usually in dependence on his or her implicit model, and is therefore not able to assess the consitency of a report adequately or to contrast the reports by different experts.

These shortcomings of the traditional approach even if it is centered on the 'critical incidences' of a process (see Flanagan, 1954) have been the motivation to devise a computer-aided interactive system for the elicitation of expert knowledge. This system is aimed at two major and interdependent goals:
(i) to help the expert to arrive at a description that is consistent and complete (as viewed by the expert)
(ii) to render different reports comparable.
If the second goal is attained, it is possible to use already obtained reports for the derivation of novel questions and to decide if there is one unique representation for the process or if there are multiple and possibly disparate representations.

In order to devise such an interactive system it is necessary to make a-priori assumptions about the format of the experts' knowledge. The gist of the assumptions made is that technical and scientific knowledge can be described as consisting of chains of arguments which are tied together by the rules of proposing, questioning, and defending these arguments. Furthermore it is assumed that such arguments differ according to the degree of belief attached to them and according to the importance they have for the overall goal of argumentation (e.g. are they necessary preconditions for further arguments or do they only together with other arguments support a certain chain of argumentation;). There is sufficient empirical support for these assumptions (see Zimmer, 1985, 1986) which underly Toulmin, Rieke & Janik's (1979) treatise on arguments, to envisage them as a viable starting point for a human-computer interaction in the elicitation of expert knowledge.

Toulmin, Rieke, and Janik (1979) assume that the statements of experts (e.g. claims, diagnoses, predictions) together with the given justifications (usually reasons in a given context) can be viewed as arguments in a rational discourse. The analysis of arguments can therefore be used to elucidate the knowledge and rule bases underlying experts' judgements, to check the consistency of their claims and justifications, and finally to derive analytically evaluations of the claims in terms of their probabilities or credibilities.

If such an approach proves to be viable it provides a novel instrument for the accumulation of judgements by different experts. Novel insofar as the consensus opionion is derived from the matching of the experts' knowledge bases underlying their claims and not, as in traditional approaches (e.g. Kim & Roush (1980)), from the experts' global subjective probabilities. However, the knowledge bases of different experts can only be matched if they have been made explicit, checked for their internal consistency, and fed back to the individual experts. This is the only way to make sure that the experts accept the knowledge base analytically derived from their arguments as representative for their initial or revised belief system. In the following, an approach is sketched in which the arguments of experts are interactively analyzed for their structure by a decision-aiding system. The output of this process is an evaluation of the original or the revised claim together with the supporting evidence and the rebuttal of alternative arguments.

THE STRUCTURE OF ARGUMENTS

According to Toulmin, Rieke, and Janik (1979) an argument consists of the following parts:

(i) claims (in the model suggested here, claims are usually predictions)

(ii) grounds or reasons for believing the claims to be valid (the usual form is that of explicitly or implicitly quantified statements)

(iii) warrants, that is, statements about the relations between grounds and claims (e.g. causality, necessity, sufficiency, contingency)

(iv) backing, that is, commonly shared knowledge (Smith, 1982) which provides the rules for a combination of grounds and warrants in order to justify the claims (e.g. rules of syllogistic reasoning, or statistical inference)

(v) modal qualifications (e.g. possibly, usually, necessarily) which apply to to the propositions and to the inferential process (e.g. rules of fuzzy reasoning). The modal quantification of the ground together with the fuzzy truth values resulting from the inferential process determine the modal quantification of the claim.

(vi) rebuttals that is, alternative claims which can be inferred from the grounds, warrants, and the backing, too, because of the modal quantification (fuzziness) of propositions and inferential rules. Rebuttals can be overcome by either showing that they imply a smaller set of consistent propositions than the claim or by comparing the overall modal qualification of the rebuttals with the evaluation of the claims.

If these parts have been identified in an expert's argument and if the credibilities of subjective probabilities of the parts have been determined, it is possible to derive analytically the credibility of the totel claim. This identification process is done interactively with the decision-aiding system.

THE INTERACTIVE ANALYSIS OF ARGUMENTS

In situations for which the suggested model has been investigated, experts are asked for statements concerning a certain state of affairs (e.g. "What kind of critical incidents do you expect in the recently introduced USN steam engines?" or "How can accidents like the one at TMI be prevented?") However, since the described model had been initially developed in the framework of economic forecasting the introductory demonstration is taken from this domain, that is, the question "What will be the further development in the money market?" *) The answers usually consist in a claim (e.g. the prediction "The US Dollar will become even stronger until the election in the US. Afterwards it will remain on a high level if Reagan is reelected, but it will considerably drop if the Mondale ticket wins.") quite often accompanied by a list of grounds (e.g. "Until election day the prime rate will tend to increase." "In the case of Reagan's reelection the budget will be balanced" etc.). Using the technique described in Zimmer (1983) the intended meaning of

---

*) for details see Zimmer (1984 a). The analysis of expert argumentation about forecasting have been made in summer 1984.

terms like ’become stronger’ or ’balanced’ are analyzed resulting in elastic constraints (e.g. "becomes even stronger" means "increases in value at least as fast as during the last quarter and at most twice as fast"). These elastic constraints are fed back to the subjects either numerically or verbally. After the subjects have accepted the analytically derived fuzzy constraints as fitting their intuitive evaluations, the grounds for the given claim are checked by questions like "How does Reagan’s reelection influence the exchange rate for the dollar?"

These questions are usually answered by fairly general statements (e.g. "He causes optimism in the economy") which in turn are scrutinized by ’how’ questions until the expert is as specific as he or she can be. The final grounds (e.g. "He will reduce taxation according to the Shaffer model") together with the warrants (e.g. "Reduced taxation makes higher investments possible. Higher investments lead to a better productivity. High productivity combined with low taxation influences positively the GNP. If the GNP in the US is higher than in other countries, there will be an influx into the dollar market. That keeps the dollar strong"). In the given train of arguments the words indicating causative relations (e.g. ’influences’) are underlined with broken lines. solid lines mark symmetric relations. Furthermore references like ’that’ in the last line of the argument indicate equating relations which can be paraphrased by ’the essence of the given arguments is equivalent to the prediction "the dollar remains strong"’.

It is possible to discriminate three major types of valid causal argumentation: (i) a simple chain, (ii) a chain with alternative parallel paths, and (ii) a tree-like structure with multiple necessary but not sufficent causes. In normal discourse combinations of all three types occur. The most frequent invalid type of causal argumentation is characterized by loops, that is, antecedents are influenced by their consequences.

In describing and analyzing argumentative chains we use the term `causative influence’ for all relations indicating that a certain event (state or action) contributes to the final or an intermediate event. In accordance with the above mentioned valid types of causal argumentation there are two forms of causative influences: (i) causal influences, that is, the occurrence of event U is necessary for the occurrence of event V and (ii) moderating influences, that is, if a certain event W occurs or a disposition G is given, the relation between X and Y is modified. For instance, only if in a landing plane (X), the position of the wings is low (disposition G) the so-called surface effect (Y) occurs immediately before touch-down. This example stems from an investigation into the correspondence between the model underlying the knowledge of flight instructors and the models students have about the behavior of the plane used for training in the German Air Force (a Piaggio Trainer). One important difference is that the instructors know about the moderating influence of the wing position (G), whereas students regard the surface effect (Y) as a necessary consequence of slowing down the speed and approaching the ground (X). This misrepresantation leads to accidents or near-accidents if the students change to a high-wing plane (e.g. a DO 27 or a Cessna 152).

In the framework of fuzzy reasoning interpretations for the relations have been developed which allow to analyze the strength of the argument. The fuzzy constraints on the relations are evaluated and these local evaluations are integrated into a global judgment (see Figure 1 for some examples).

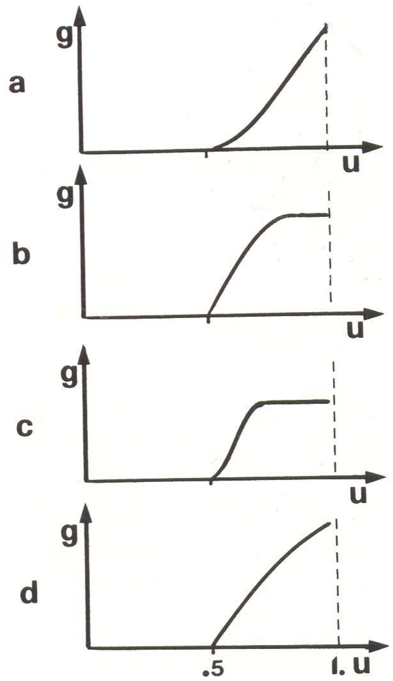

Figure 1 Influence functions; "g" is the weight for a quantified statement combined with "f" by means of a fuzzy multiplication or a similar operation, see Zimmer, 1984a. a) positive influence of X on Y, b) strong positive influence of X on Y, c) causal influence of X on Y, d) is a sufficient condition for Y. "U" is the 'Universe of Discourse'.

A special kind of relations are those with explicit (e.g. "Often X is influenced by Y") or implicit quantification (e.g. "Xs prevail if Y"). If the quantification and the meaning of the quantifier are unambiguous, the quantifier is represented by a fuzzy number in the interval [0, 100 %] giving the expected proportion of positive incidences of X (see Zadeh, 1984; Zimmer, 1984a,b). In the case of ambiguous quantifiers (e.g. usually) the meaning is interactively rendered so precise that the quantifier can be characterized by a unique fuzzy number (see Figure 2).

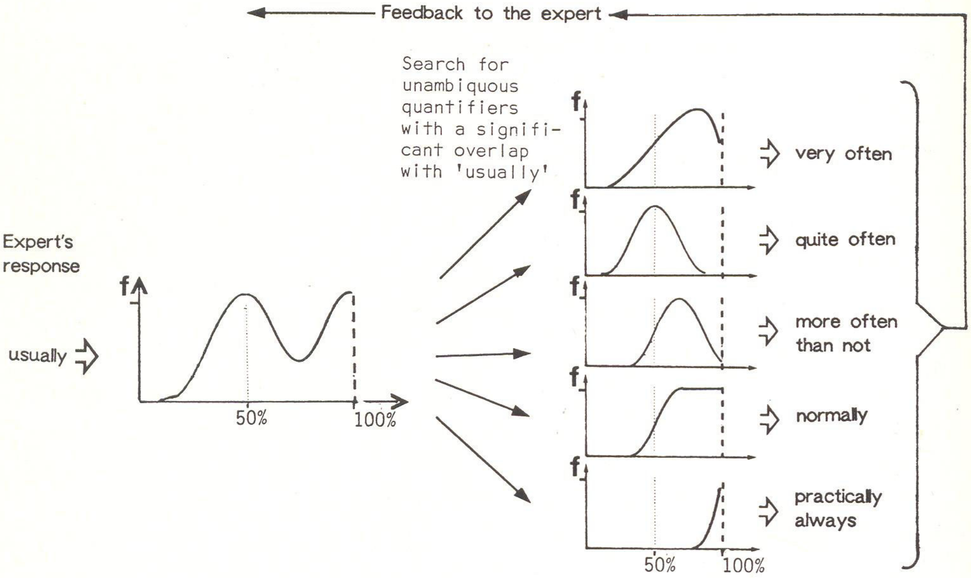

Figure 2 The disambiguation process for the quantifier "usually". The alternatives on the right hand are chosen because their possibility functions are single peaked and they are either entailed in the possibility function of "usually" or their intersections with it are large.

These local evaluations and the propagation of these interpretations to the final evaluation is comparable to the analysis of the backing in Toulmin's schema for arguments. In addition to the structural analysis of grounds and their interconnections it is checked how and if the reasons given are backed by factual knowledge (e.g. "On which observable instances do you base the statement 'X'?") and how reliable and generalizable these supporting facts are. For the final assessment of the argument's credibility the given ratings for the reliability of factual knowledge are integrated as fuzzy truth values modified by the relations between the grounds (see Zadeh, 1984; Zimmer 1984 b).

ELICITATION OF COUNTER ARGUMENTS

In normal discourse the strength of an argument does not only depend on the credibility of the facts underlying the given grounds and on the strength of the connection between the grounds and the claim but also on possibility of a rebuttal of plausible counterarguments.

In many claims counterarguments are explicitly referred to (e.g. "I predict X. One could also argue that Y. However, A cancels out Y as a possible course of development"). In other cases the absence of a certain ground is stated (e.g. "I predict that X, because A, B, C are given and D is not the case"). From a language pragmatic point of view giving negative grounds is not very informative and therefore this way of bringing forward arguments is unlikely to occur. However, if this is interpreted as an indirect refutation of a counterargument (e.g. "D is necessary for Y. D is not the case, therefore prediction Y can be cancelled out"), then stating the absence of facts or grounds makes sense. In the model the implied counterarguments are elicited by conversing the negative statement into the respective positive form and then asking the subject: "If D were given, what prediction would you make?"

The experimental evidence from the applications of the model indicates that in argumentative discourse the normal form of counterarguments is the contrary of the arguments and not its contradiction. From a logical point of view this seems curious. However, there are good reasons for this strategy from a language pragmatic point of view (Grice, 1975):

(i) Contraries preserve the context (universe of discourse). For instance, the counterargument for "A is small" is "A is large". Both arguments refer explicitly to size as the universe of discourse.

(ii) Contraries are usually more informative than contradictions. The negation of "A is small" can have the following meanings:
(a) size is not an applicable predicate for A, (b) A is less than small (minute) or more than small (moderately sized, large, huge, etc.).

(iii) If the linguistic labels in the universe of discourse are assumed to be those of natural language, then their meaning is vague or fuzzy. One property of fuzzy contradiction, however, is that the law of the excluded middle does not apply (Osherson & Smith, 1981). Therefore, the original argument and its contradictory counterargument have an overlap in meaning (see Figure 3), that is, there are instances which are in favor of the argument as well as its contradiction.

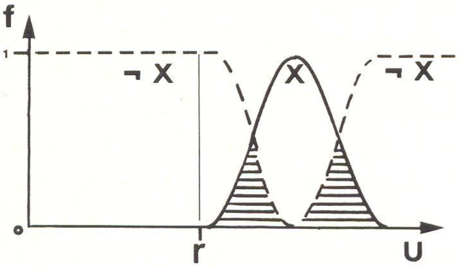

Figure 3 The membership function for X and for not X. The shaded area indicates the overlap in meaning of X and not X. Note that the reference point does not influence the characteristic of the counterargument.

This ambiguity of counterarguments vanishes if the contrary is chosen instead of the contradiction (see Figure 4).

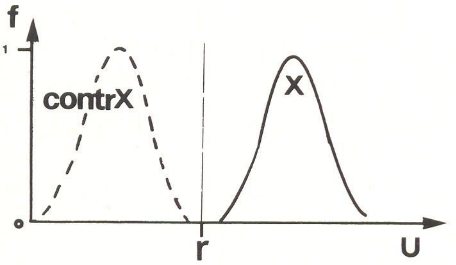

Figure 4 The membershipfunctions for X and the contrary of X. The position of contr(X) is determined by reflection of the membership function of X at the point of reference.

If counterarguments are neither stated explicitly nor implicitly, the experts are confronted with the contraries of circumstantial reasons they have given for their arguments (e.g. "If the Mondale/Ferraro tickets wins (the contrary of 'the Reagan/Bush ticket wins') what kind of development for the dollar would you predict?")

GLOBAL EVALUATION OF ARGUMENTS

After all the components of the argument have been elicited and after the meaning of the used predicates and the credibility of the ground, the warrant, the backing, and of rebuttals have been determined, the credibility of the claim is analytically derived from these values by means of fuzzy syllogistic reasoning (Zadeh, 1984; Zimmer, 1984 b). The credibility values are either transformed into labels (e.g. 'likely', 'unprobable' etc., see Zimmer, 1983) or are given in their original numerical form. If the subjective evaluation of the claim by the expert and the analytical evaluation are about the same, the interactive process ends. If, however, the expert disagrees, he or she is asked to give further grounds or to revise the credibility ratings for the facts given.

CONCLUSION

The interactive model for the elucidation of arguments underlying the claims (e.g. predictions, diagnoses) of experts on the one hand serves as a means for an unbiased probability assessment for claims. Insofar it resembles the procedure proposed by Henrion and Morgan (see Morgan, in press) and the investigation of Shafer & Tversky (1985) into subjective probabilities as formal languages for the expression of evidence and degrees of belief. On the other hand, however, it makes explicit the knowledge base on which the expert grounds his/her claim. The comparison of the knowledge bases underlying the predictions of different experts for the same event shows if these predictions are based on more or less same reasons or not. In the first case an accumulation of the evaluations made by different experts is admissible. In the other case, however, only those judgments can be pooled which are based on comparable knowledge bases.

REFERENCES

1 J.C. Flanagan, Psychological Bulletin, Bd. 51, 1954, pp. 327-355.

2 K. Forbus & A. Stevens, Project Steamer: III. Using qualitative simulation to generate explanations of how to operate complex physical devices, NPRDC TN 81-25, 1981.

3 D. Gentner & A. Stevens (Eds.) Mental models, L. Erlbaum, 1983, Hillsdale.

4 H.P. Grice, in: P. Cole and J.L. Morgan J.L. (Eds.), Syntax and Semantics, Academic Press, New York, 1975, pp.

5 P.N. Johnson-Laird, Cognitive Science,4 (1980) pp. 71-115.

6 K.H. Kim and F.W. Roush, Introduction to Mathematical Consensus Theory, Marcel Dekker Inc., New York, 1980, pp. 179.

7 G. Morgan, in: R. Uppuluri (Ed.) (ed.) Technology Assessment, North Holland, Amsterdam, in press.

8 D.N. Osherson and E.E. Smith, Cognition, 4, (1980) pp. 35-58.

9 J. Reason & K. Mycielska, Absent Minded? Prentice Hall, Englewood Cliffs, pp. 263.

10 G. Shafer & A. Tversky, Cognitive Science, 9, pp. 309-340.

11 S. Toulmin, R. Rieke and A. Janik (Eds.), An introduction to reasoning, Macmillan Publishing Co., Inc., New York, 1979, pp.

12 L.A. Zadeh, Syllogistic Reasoning in Fuzzy Logic and Its Application to Reasoning with Dispositions, UCB/EECS, Berkeley, Memorandum.

13 A. Zimmer, in: R.W. Scholz (Ed.), Decision Making Under Undercertainty, North-Holland, Amsterdam, 1983, pp. 159-182.

14 A. Zimmer, Int. Journal Man-Machine Studies ,20, 1984, pp. 121-134.

15 A. Zimmer, Annual meeting of the Psychometric Society, 1984, Santa Barbara.