A CALCULUS OF POSSIBILITY, EVENTUALITY AND PROBABILITY

B.R.Gaines

University of Essex,
Colchester, Essex, England.

Abstract: Three distinct explicata of uncertain system behaviour are developed and it is shown that they each give rise to different phenomena that are confounded if only probability theory is used to represent them. It is also shown that a conventional binary representation of possible transitions in non-deterministic automata cannot support certain legitimate arguments about the resultant behaviour. A weakened logic of probability is developed as a precise explicatum of all that may be inferred about non-deterministic, but also non-probabilistic, behaviour. This is extended to cover all three forms of uncertain behaviour, and their combinations, leading to a rigorous calculus of possibility, eventuality and probability. based on a vector representation of uncertainty.

1 INTRODUCTION

The commonly used tools for analysing systems whose behaviour is uncertain are those of probability theory. However, the assignment of a non-zero, non-unity, probability (a proper probability) to an event has more connotations than that the occurrence of the event is uncertain (i.e. it may, or may not, occur). It implies that in a sufficiently long sequence of events this one is eventually bound to occur. It also implies the even stronger result that the relative frequency of such events in a sequence will tend to converge to the given probability with increasing length of sequence. Either, or both, of these additional connotations may be too strong in practical situations where the concepts of probability theory are being used to express the effects of uncertain behaviour.

For example, in the analysis of system stability or reliability we are often faced with situations where an event, E, may occur, but there is no guarantee that E actually will occur, no matter how long we wait. If we ascribe some arbitrary, non-zero probability to E then we certainly express that it is a possible event. However we are then in a position to derive totally unjustified results based on the certainty of some eventual occurrence of E, or meaningless numeric results based on the actual 'probability' of occurrence of E.

The danger of deriving profound results that have no justification other than an unwarranted strength in the theory is a real one. For example, Gaines (1971,1974) has shown that a two-state stochastic automaton can solve a class of control problems otherwise requiring a recursive automaton (Gold 1971) and not soluble by any finite automaton (Gaines 1971, Gold 1971). This significant result is dependent on a source of uncertain behaviour that is properly probabilistic, but whose probability does not have to be known. It cannot be derived if the behaviour is merely possibilistic. There is no way, however, of preventing the consequences of this result appearing in the analysis of a system in which uncertainties have been represented by probabilities rather than possibilities.

A similar problem arises in the practical aplication of linear systems theory. There are many results which may be derived from the assumption of linearity (such as the complete extension of knowledge of local behaviour to that of global behaviour) which are false in most practical systems. The engineer resolves these problems in practice through a set of 'rules-of-thumb' based on commonsense and experience which constrain the deductions he is prepared to assume valid. Such a resolution is however extremely difficult to implement in an automated, or computer-aided, design system, and becomes increasingly difficult to apply as the system involved becomes more complex.

This paper analyses the problem of describing precisely and quantitatively the structure of systems whose behaviour is uncertain. A formal calculus is developed for the three connotations of uncertainty outlined above, allowing possibilistic, eventualistic and probabilistic behaviour, and any mixture of these three, to be taken into account and only legitimate deductions to be drawn.

2 POSSIBLE, EVENTUAL, AND PROBABLE EVENTS

It appears that there are three distinct explicata of uncertainty, each of which has its own consequences that require clear separation:-

(i) <u>Possible</u> Event E is possible - no reliance may be placed upon the occurrence or the non-occurrence of E. This corresponds to an interpretation of E as an event whose negative consequences must be taken into account, but whose positive consequences cannot be relied upon. The modal operator of 'possibility', M, in alethic modal logic (Hughes and Creswell 1968, Snyder 1971) represents this concept, but conventional probability theory provides no explicatum for it.

(ii) Eventual Event E will eventually occur in that it frequent in the sense of the theory of infinite sequences, i.e in a series of events E(i), for any n, there exists m>n, such that E(m)=e. This corresponds to the interpretation of E as an event whose eventual occurrence may be relied upon, but whose relative frequency of occurrence is not necessarily stable or known. A suitable explicatum in probability theory is that p(E)>0, the event has a non-zero probability of occurrence.

(iii) Probable Event E is frequent and its relative frequency of occurrence in a sequence of events converges to a definite value, p(E), its probability of occurrence. This is the type of event with which we are most used to dealing using the methods of probability theory.

Gaines and Kohout (1975) have shown that it is possible to take these three types of event and add to them two further types, necessary and impossible events (always or never occur, respectively), to form a multi-valued logic. The logic is mixed discrete-continuous since probable events are represented by a number in the semi-open interval (0,1]. Without probable events the logic (in terms of conjunction and disjunction) is a 4-value Post algebra and may also be regarded as a fuzzy logic (Zadeh 1976, Lakoff 1973). With probable event types included idempotency fails and even the generalization of fuzzy logic to a distributive lattice (Brown 1971) is inadequate illustrating the need for more general truth sets discussed by Goguen (1974).

The multivalued logic proposed by Gaines and Kohout (1975) provides an improved account of possibility and probability, and their mixtures, in that it does not allow false conclusions to be drawn about possibilistic events, and yet it contains a full account of truly probabilistic events. However, it suffers from what appears to be a fundamental defect of all attempts to account for possibilistic, or non-deterministic, behaviour in terms of a finite- valued logic. It is unable to sustain certain forms of deduction leading to deterministic conclusions about non-deterministic behaviour. This problem is analysed through an example in the next section and a new logic of possibility is developed that is similar to that of probability and supports all the legitimate, but only the legitimate, derivations of results about possibilistic behaviour.

3 THE PROBLEM OF POSSIBILITY NORMALIZATION

The problem of drawing conclusions about possible events is best seen in terms of an example. Consider the indeterministic automaton of Figure 1 - starting in state A ~~S0~~, its future states are indeterminate. However, even if we know only that the transitions are possible, it is clear that ~~the~~ state B S2 will certainly be entered at some time. If we know also that the transitions are eventual then it is also certain that the ultimate state will be S5. If, in addition, the transition probabilities are well-defined then we may also derive the expected time for this state to be reached. This last conclusion is a numeric result readily represented in probabilistic terms, but what of the weaker results ? They are not in themselves quantitative but they do seem to be based on an underlying quantitative argument - when the state will be S2 is uncertain but the 'total uncertainty' about that state sums to a certainty that it will occur.

The normal representation (Santos and Wee 1968 p.7) of a possible, or non-deterministic, transition by a binary logical variable taking the values 0 (impossible) and 1 (possible) cannot be used to support this form of reasoning. For example, Table 1 shows the possibility of each state of the automaton of Fig 1 at successive clock times. It can be seen that the pattern of behaviour for S2 is identical to that for S3, and yet we can see that S2 must occur whilst S3 may only possibly occur. Clearly an exhautive enumeration of all possible paths from S0 to S5 will show that S2 is on all of them whilst S3 is not, but such combinatorial searches become difficult when the system is complex and contains loops (leading to an inifinite number of possible paths).

If the transitions were probabilistic the argument could be based on a simple numeric calculation of the total probability of each of the states, S2 and S3. What appears to be lacking in the binary representation of possible transitions is the normalization possible with probabilities that expresses that the automaton is actually in one, and only one, state. As discussed in Gaines and Kohout (1975) the normalization of the columns of Table 1 is appropriate to a non-deterministic automaton in that at least one of the states has the value 1, but there is also the auxiliary rule that if only one of the states has the value 1 then the automaton is definitely in that state.

It is in the form of this auxiliary rule that the weakness of expressing possibility in a finite-valued logic seems to lie. To find out if the automaton is definitely in a state we have to examine the possibilities of all other states and show that they are zero. This global argument contrasts sharply with the local reasoning in the probabilistic case that the automaton is definitely in a state because the probability of that state is 1. There seems no reason, however, why we should not retain this 'conservation law' so readily expressed in probabilities without giving the actual numeric probabilities anything more than a possibilistic interpretation, i.e.:-

$p(E)=0$	E is impossible
$0<p(E)\leq 1$	E is possible
$p(E)=1$	E is necessary.

A calculus of possibility based on these definitions is quite simply developed and in fact gives non-deterministic automata the structure of probabilistic automata with the weakened semantics that, apart from 0 and 1, the values of 'probability' have no greater significance than that an event is possible.

4 A LOGIC OF POSSIBILITY AND EVENTUALITY

Rescher (1969 section 27) has given a set of postulates for what he calls a 'probability logic' over a domain of statements. The logic is defined in terms of a valuation over the lattice of conjunction and disjunction of statements that assigns some real value, P(A), to every member, A, of the universe of statements. This assignment has to satisfy the postulates:-

(P1) $0 \leq P(A)$, for any statement, A.

(P2) $P(A \vee \bar{A}) = 1$

(P3) $P(A \vee B) = P(A) + P(B)$, provided A and B are mutually exclusive

(P4) $P(A) = P(B)$, if A is logically equivalent to B

(P5) $P(A \wedge B) = P(A) + P(B) - P(A \vee B)$, defs conjt

(P~~5~~6) $P(A \supset B) = P(\bar{A} \wedge B)$, defining implication

(P~~6~~7) $P(A \equiv B) = P(A \supset B \wedge B \supset A)$, defining equivalence

These are the normal basic requirements for a probability measure, but they may also be regarded as a set of postulates for an infinite-valued logic. The logic is not truth-functional but if the value 1 only is designated then the truth tables for the operations of negation, conjunction and disjunction are those of the classical propositional calculus (PC). Conversely, the axioms that define PC may be shown to be tautologies of probability logic (Rescher 1969 p.187). Hence the system coincides completely with PC in its tautologies.

Rescher (1969 section 28.2) introduces modalities into the logic by the stipulations:

(P~~7~~8) Necessity: LA = 1 or 0 according as P(A') is, or is not, uniformly 1 for every substitution instance, A', of A.

(P~~8~~9) Possibility: MA = 0 or 1 ccording as P(A') is, or is not, uniformly 0 for every substitution instance, A', of A.

The use of the concept of substitution instances is necessary because the logic is not itself truth-functional. Rescher (1963) has demonstrated that the logic with these modalities is characteristic of Lewis' system S5 of modal logic (Hughes and Creswell 1968) in that its tautologies are precisely those of S5, and vice versa. Thus, whilst there is no finite-valued logic that represents precisely the alethic modal logic of necessity and possibility, this (infinite-valued) 'probability logic' does so.

If we consider only mutually exclusive events, such as an automaton being in one or another of its states, then it may be seen from P3 that the logic becomes truth-functional. Valuations are then just additive over the disjunction of events. Hence also ~~P8~~ P9 may be interpreted as, "an event is possible if and only if its valuation is non-zero", which may be seen as a binary evaluation similar to the 0/1 representation of impossibility/possible in non-deterministic automata. However we also have the new rule based on ~~P7~~ P8 that, "an event is necessary if its valuation is unity". This corresponds to our previous additional rule that if the disjunction of a set of mutually exclusive events is necessary, and only one of the events is possible, then that event must be necessary. This is now derivable from the purely arithmetic effect of the additivity of positive valuations, i.e. if the sum of a set of numbers is 1, and all but one of those numbers is zero, then that number must be 1.

It is interesting to compare this with the corresponding rule of the modal logic S5 (T28 in Hughes and Creswell 1968 p.51) that:

$$L(A \vee B) \supset (LA \vee MB)$$

which clearly extends to multiple events:

$$L(A \vee B \vee C \vee \ldots) \supset (LA \vee MB \vee MC \vee \ldots)$$

i.e. if it is necessary that at least one of a set of events occur then either one of the events is necessary or some of the others are possible. Hence, from the impossibility of all but one event we can infer the necessity for that event. It can be seen that what the 'probability logic' of S5 does is replace a process of logical deduction with one of arithmetic. The failure of a binary representation of possibility to do this may itself be seen as a demonstration of the impossibility of characterizing S5 with a finite-valued logic

(Dugundji 1940).

Let us now apply this logic to the previous problem of modelling a non-deterministic automaton. Consider the set of all posssible statements of the form, "the automaton was in a particular state, x, after the n'th transition", where $x \varepsilon X$, the state set of the automaton. For each n, these statements are mutually exclusive and clearly a set of numbers may be assigned to them which sum to 1 to express that the automaton is necessarily in some state, and are such that zero is assigned to an impossible state. Such an assignment of a value x(n) to state x is consistent with postulates P1 through P~~8~~, P9 and satisfies the weaker interpretation of 'probabilities' given at the end of the previous section. It will be called a 'normalized' distribution and satisfies:-

$$\sum_{x \varepsilon X} x(n) = 1$$

$$x(n) \begin{cases} = 0 & \text{if state x is impossible after transition n} \\ > 0 & \text{if state x is possible after transition n} \\ = 1 & \text{if state x is necessary after transition n} \end{cases}$$

A state transition now corresponds to a transformation of one normalized distribution into another, and to make the semantics correct it is necessary only to ensure that a state is possible after a transition if and only if there is a possible path to it from a state that was possible before the transition. The normal representation (Arbib 1969 Ch.9) of probabilistic transitions by a "stochastic matrix" transforming a state distribution "vector" by matrix/vector multiplication has precisely this property. It is not unique in this for the weaker cases of possibility and eventuality, but there is no reason to prefer any other choice. Thus a transition from x to y may be represented by a number, T(x,y), such that:

$$T(x,y) \geq 0$$

$$T(x,y) \begin{cases} = 0 & \text{if the transition is impossible} \\ > 0 & \text{if the transition is possible} \\ = 1 & \text{if the transition is necessary} \end{cases}$$

$$\sum_{y \in X} T(x,y) = 1$$

so that T(x,y) is a normalized distribution over $y \in X$. The next state distribution is then given by:

$$x(n+1) = \sum_{y \in X.} T(x,y)y(n)$$

A re-analysis of the automaton of Fig.1 shows that the difference between states ~~S2~~ C and ~~S3~~ D that was previously obscured is now apparent. Table II is the new version of Table I - to show the generality of the result symbols have been used rather than numbers - a, b, c, etc. are any numbers in the open interval, (0,1). The final column gives the sums of the elements in each row. For S0 through S3, since the automaton being, for example, in S2 at time 1 and at time 2 are mutually exclusive possibilities, the sum properly represents the total possibility of the automaton being in the state. It can be seen that S1 and S3 are only possibly entered but that S2, for which the total is 1, will be necessarily entered. The sums for S4 and S5 are not meaningful because the loops in the state diagram rule out mutual exclusion and hence the additivity of possibilities.

The penultimate column of Table II shows the final possibility of the automaton being in each of its states. Whilst that for S4 is asymptotic to 0 and that for S5 is asymptotic to 1, both are essentially non-zero for all time and hence, if the transitions are possibilistic, the most we can say is that both states are ultimately possible. This serves to illustrate the essential distinction between the analysis of possible and eventual behaviour since, if the transitions are eventual, we may show that an asymptotic approach of the possibility of an event to unity indicates that that event must ultimately necessarily occur.

To give this statement a rigorous interpretation we may say that an event is ultimately necessary if, no matter what the actual values of the transition possibilities provided that they are eventual and conform to the semantics of impossibility/possibility/necessity given previously (i.e. the distinction between zero and non-zero possibilities is preserved), the possibility of that event is asymptotic to unity. The idea behind this definition is that we are dealing with a probabilistic situation in which the actual values of probabilities are vague and may fluctuate provided the possibilistic logic of the situation is preserved. Any numeric result which is indepedent of the actual values is significant in the eventual case, whilst only exact summation (rather than asymptotic approach) to unity is significant in the possibile case.

Thus, in summary, one can take a model of automaton structure and behaviour which is identical to that for the conventional probabilistic automaton and by weakening the interpretation of the numeric 'probabilities' to that of 'possibilities' one can obtain precise accounts of the behaviour of automata with either possible, or eventual, transitions and state distributions. This result is a function of the equivalence between the modal logic, S5, and 'probability logic', an equivalence in which the frequentist interpretation of numeric 'probabilities' plays no part.

5 UNIFICATION OF POSSIBILITY, EVENTUALITY, AND PROBABILITY

It has been shown that a 'probability logic' with one of two weaker interpretations than usual of the numeric results gives an adequate and complete explicatum of possible and eventual behaviour in non-deterministic automata. However only the pure cases have been treated so far and it has already been argued that the treatment of a mixed case of more than one type of behaviour as a uniform example of a pure case will lead to deductions which are either too strong or too weak. In practice the extension to the mixed case may be made quite simply by considering generalized possibility to be a 3-vector, and the resultant calculus has some interesting further properties.

Consider the possibility of an event x now to be represented by a 3-vector of positive numbers, $(x1,x2,x3)$, whose components are: $x1$, probability; $x2$, eventuality; $x3$, possibility. this is to interpreted that: the true probability of the event is at least $x1$ (exactly $x1$ if $x2+x3=0$); that the event will eventually occur if $x1+x2>0$; and that the event is possible if $x1+x2+x3>0$. This last term will be defined as a norm on the three vector:

$$x0 = x1 + x2 + x3$$

and it will be postulated that this norm conforms to P1 through ~~P8~~ P9 for a probability logic. Hence it is clear that the uniform restriction of two of the components of the 3-vectors to zero gives rise to exactly the pure cases so far discussed.

For the mixed case one must ensure that the rules of combination are consistent with the semantics that probability can never be generated from eventuality and neither may be generated from possibility. For the addition of vector possibilities normal vector addition suffices - if x and y are mutually exclusive events whose joint occurrence is called z, then:

$$(z1,z2,z3) = (x1+y1,x2+y2,x3+y3)$$

so that also:

$$z0 = x0+y0.$$

Thus the minimum probability of z is the sum of those of x and y, and z is possible (eventual) if and only if either of x and y is possible (eventual).

Multiplication of the 3-vectors is more complex because it corresponds to interactions between the different types of possibility, e.g. a state is possible if it arises from a state which was probable through a possible transition. The definition that, if x is a vector possibility corresponding to a state and y that corresponding to a transition, then the resultant is z such that:

$$(z1,z2,z3) = (x1y1, x2y2+x2y1+x1y2, x3y3+x3y1+x3y2+x1y3+x2y3)$$

has the correct semantics and preserves the norm so that:

$$z0 = x0y0.$$

If the result for z is written in the form:

$$(z1,z2,z3) = (x1y1, (x1+x2)(y1+y2)-x1y1, x0y0-(x1+x2)(y1+y2))$$

it is more clearly apparent that powers of a 3-vector, x, by this definition of multiplication are of the simple form, if $z=x^n$:

$$(z1,z2,z3) = (x1^n, (x1+x2)^n-x1^n, x0^n-(x1+x2)^n)$$

$$z0 = x0^n$$

so that for any function, f, which may be expressed as a power series with these definitions of addition and multiplication we have, if z=f(x):

$$(z1,z2,z3) = (f(x1), f(x1+x2)-f(x1), f(x0)-f(x1+x2))$$

$$z0 = f(x0)$$

The mixed calculus of probability, eventuality and possibility thus appears the same as any of the pure cases with possibility represented as a 3-vector and addition and multiplication defined as above. The new properties of interest show up when we consider functionals of distributions since the eventualistic and possibilistic components of, for example, the entropy of a distribution have no obvious interpretation. It is meaningful however to consider the minimum and maximum of the norm of a functional for all assignments of the second and third components of the vectors in a distribution which preserve the semantics, i.e. are such that the overall norm of the distribution is constant. This corresponds to treating the eventual and possible components as being a residual probabilistic distribution to be assigned for best possible, or worst possible, effect. Thus such functionals as the entropy of a possibilistic distribution do not have a unique value but rather a range of values defined by its maximum and minimum.

6 CONCLUSIONS

The concepts advanced in this paper may be seen as a straightforward extension of probability theory to cope with richer forms of uncertain behaviour found in practical systems. The fact that the pure cases of possibilistic and eventualistic behaviour may be treated computationally with a normal probability logic whose semantic interpretation is weakened is itself a formal justification in engineering studies for treating all types of uncertainty as probabilistic but taking no notice of some of the computed results. This is not effective however in the mixed case where a system exhibits more than one type of non-deterministic behaviour, and the simple extension given to a vector of possibility with appropriate rules of addition and multiplication is necessary to cope with this case. It has the advantage that the techniques developed for the analysis of probabilistic automata may be carried over directly to the mixed case.

ACKNOWLEDGEMENT

I wish to thank Ladislav Kohout for many discussions of modality and probability.

REFERENCES

Arbib, M.A. (1968) "Theories of abstract automata", New Jersey: Prentice-Hall.

Brown, J.G. (1971) A note on fuzzy sets, Information and Control 18, 32-39.

Dugundji, J. (1940) Note on a property of matrices for Lewis and Langford's calculi of propositions, Journal of Symbolic Logic 5, 150-151.

Gaines, B.R. (1971) Memory minimization in control with stochastic automata, Electronics Letters 7, 710-711.

Gaines, B.R. (1974) The role of randomness in cybernetic systems, Proc. Cybernetics Society Conf., London, September 1974.

Gaines, B.R., and Kohout, L. (1975) Possible automata, Proc. Symp. Multi-Valued Logic, Bloomington, Indiana, May 1975.

Goguen, J.A. (1974) Concept representation in natural and artificial languages: axioms, extensions and applications for fuzzy sets, Int. J. Man-Machine Studies 6, 513-561.

Gold, E.M. (1971) Universal goal seekers, Information and Control 18, 395-403.

Hughes, G.E., and Creswell, M.J. (1968) "An introduction to modal logic", London: Methuen.

Lakoff, G. (1973) Hedges: a study in meaning criteria and the logic of fuzzy concepts, Journal of Philosophical Logic 2, 43-59.

Snyder, D.P. (1971) "Modal logic and its applications", New York: Van Nostrand Reinhold.

Rescher, N. (1963) A probabilistic approach to modal logic, Acta Philosophica Fennica 16, 215-226.

Rescher, N. (1969) "Many-valued logic", New York: McGraw Hill.

Santos, E.S., and Wee, W.G. (1968) General formulation of sequential machines, Information and Control 12, 5-10.

Zadeh, L.A. (1976) A Fuzzy-algorithmic approach to the definition of complex or imprecise concepts, International Journal of Man-Machine Studies 8, **-**.