REPORT DOCUMENTATION PAGE

Form Approved
OPM No. 0704-0188

AD-A250 601

...ated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources gathering and ...rmation. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions ... Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to ...ment and Budget, Washington, DC 20503.

	2. REPORT DATE	**3. REPORT TYPE AND DATES COVERED**
	1989	Unknown

4. TITLE AND SUBTITLE	**5. FUNDING NUMBERS**
Objective Probabilities	DAAB10-86-C-0567

DTIC ELECTE MAY 27 1992 S C D

6. AUTHOR(S)

Henry E. Kyburg, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
University of Rochester Departmenet of Philosophy Rochester, NY 14627	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
U.S. Army CECOM Signals Warfare Directorate Vint Hill Farms Station Warrenton, VA 22186-5100	92-TRF-0009

1. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT	**12b. DISTRIBUTION CODE**
Statement A; Approved for public release; distribution unlimited	

13. ABSTRACT *(Maximum 200 words)*

Almost everyone will agree that when our background statistical knowledge is extensive enough, and when the case with which we are concerned is a "repeatable event", then objective probabilities are appropriate, and these are the probabilities that should enter into computation of expectation and into our decision theory. A great many people will also agree that there is another whole class of cases, in which we are concerned with unique events, in which we lack statistical knowledge, and for which we must turn to subjective probability or one of its surrogates. I purpose here to argue against this distinction. Of course it is easy enough to argue this way in a purely philosophical vein; every event must be unique -- it has its own spatio-temporal locus; and every unique event must belong to some class of events about which, in principle, we could have statistical knowledge. But this is not my point. My point is that from a down-to-earth practical point of view, form the point of view that seeks to compute probabilities and expectations for making decisions, the distinction between "repeatable" and "unique" events is not only untenable, but seriously misleading.

1. SUBJECT TERMS		**15. NUMBER OF PAGES**
Artificial Intelligence, Data Fusion, Objective Probabilities Philosophy		10
		18. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Objective Probabilities

1. statistics and unique events.

Almost everyone will agree that when our background statistical knowledge is extensive enough, and when the case with which we are concerned is a "repeatable event", then objective probabilities are appropriate, and these are the probabilities that should enter into the computation of expectation and into our decision theory. A great many people will also agree that there is another whole class of cases, in which we are concerned with unique events, in which we lack statistical knowledge, and for which we must turn to subjective probability or one of its surrogates. I propose here to argue against this distinction. Of course it is easy enough to argue this way in a purely philosophical vein: every event must be unique -- it has its own spatio-temporal locus; and every unique event must belong to some class of events about which, in principle, we could have statistical knowledge. But this is not my point. My point is that from a down-to-earth practical point of view, from the point of view that seeks to compute probabilities and expectations for making decisions, the distinction between 'repeatable' and 'unique' events is not only untenable, but seriously misleading.

2. kinds of cases.

Let us consider some examples of these alleged distinctions. Consider a toss of a coin. There is a classical 'repeatable' event: not only can we toss a coin over and over again; coins have been tossed over and over again, and in the experience of each of us there is a large data-base of results of coin tosses (or an impressionistic resume of such a data base). And we have physical grounds (i.e., grounds stemming from the

laws of physics) for thinking that coins land heads about half the time. And so we can regard that toss of that coin as a member of a class of tosses. of which we have reason to believe that half yield heads. (Alternatively we might regard that toss as a _kind_ of trial that has a propensity of a half to yield heads.)

Now of course a particular toss, at a particular time and place cannot be repeated. We all know that. But the event _can_ be repeated in 'all relevant respects'. We don't have to make the toss at the same time or the same place; we don't have to use the same coin; we don't have to use the same kind of coin; we don't have to flip it in any particular way.

Consider another case. Suppose I am interested in whether or not my friend Sam will be at home tonight after supper. I know that on some week-nights he goes to the movies, but that he never goes anywhere else. But this doesn't give me a 'repeatable event'. Whether or not he goes to the movies depends on factors that I can specify. Unlike the flip of the coin, I can be quite sure that there are factors that influence his behavior -- and I can even specify some of them. Thus I might know that he likes Westerns -- so if there is a good Western in town he will be more likely to be at the movies. On the other hand, I might know that Sam is very conscientious about studying for his chemistry examinations -- so if there were to be a chemistry examination tomorrow, Sam would most likely be at home. In short, this seems like a perfect case for subjective probability. But if I have known Sam for a long time, I _do_ have a basis for knowing how often, in general, he goes to the movies on week-nights. My knowledge is neither so precise nor so secure as my knowledge about the coin, but it is surely not non-existent. Similarly, those factors I just named (Sam's conscientiousness about chemistry; his liking for Westerns),

2

so far from accentuating the contrast with the well-known coin, create a greater similarity. It is as if we _did_ know something special about tosses of coins of different denominations: quarters tend to land heads more often than other coins, dimes less often.

Of course, the better I know Sam, the more precise I can be about this probability. But this just means that I know more about the relative frequency with which he goes to the movies. If his life were to change a lot, I would regard my knowledge of this frequency as suspect. But if the coin were to have been run over by a trolley-car, I would regard my knowledge of it as suspect.

We must beware of allowing the variety of our knowledge about Sam to serve as an excuse for guessing wildly. Analogously, if we were to have detailed and microscopic data concerning the coin toss, we could perhaps predict with a better than 50% success rate. This possibility should not be allowed to undermine our sensible tendency to assign a probability of 1/2 to the occurrence of heads on the specified toss when we lack that microscopic data.

Finally, there are some circumstances under which my probability concerning Sam's being at home is just as exact as my probability concerning the flip of a coin. For example, I may know that he decided whether or not to go to the movies by flipping an ordinary coin. Suppose in addition that I know that he went out if and only if the coin landed heads. Then the probability that Sam went to the movies is one-half. This brings out an important point that I shall enshrine as an axiom:

Al If $\underline{S}$ and $\underline{T}$ are known to have the same truth-value then

they have the same probability.

This axiom does not require that $\underline{S}$ and $\underline{T}$ be equivalent in any strong

3

sense; all that is required is that we know that they have the same truth value.

This axiom already undermines the argument for subjectivity based on uniqueness. It is asked, "How can you find an objective probability *for the event* of an Indian plutonium plant suffering a melt-down, when there is no class of instances to generalize from: there is only one Indian plutonium plant, its design is unique, etc." The long answer is that there will be a melt-down if and only if there is a failure of the cooling system or a failure of the rod control system. In turn we can break these propositions into equivalent ones until we have quite small and quite general statements about which we <u>do</u> have statistical data. (Gate valves of the kind in the Indian plant's cooling system have a failure rate of about one per ten thousand hours of use.) But how do we use statistical data when we do have it?

Here is another example. I hold in my hand a newly minted coin. I will toss it once and then it will be melted down. What is the probability that this coin will yield heads when tossed? The relative frequency of heads among tosses of this coin is 0 or 1 - we have no statistical knowledge of the behavior of this coin. But we know that the toss of this coin I am about to perform will yield heads iff the next toss of a coin performed by me yields heads - and for tosses of coins in general we have lots of statistical evidence.

3. A Simple System

Here is a very simple example of how objective (frequency or chance) probability can be applied to "unique events". It is essentially due to Reichenbach (1949). It is a step backward from the discussion of the previous section, but we will regain our insights in the following

4

section.

Let $\underline{R} = \{\underline{r}_1...\underline{r}_n\}$ be a finite set of potential reference classes, let $\underline{P} = \{\underline{p}_1...\underline{p}_k\}$ be a finite set of properties (including such properties as being a member of a particular reference class), and let $\underline{I} = \{\underline{i}_1...\underline{i}_m\}$ be a set of distinct individuals (or individual events).

We can define a language on this basis in the usual way.

Add to this language enough mathematics to do statistics, and define an _item_ _of_ _possible_ _statistical_ _knowledge_ to be a sentence of the syntactical form: "$\%(\underline{r}_j,\underline{P}_s) = \underline{x}$", which we read: the proportion of objects in the reference class $\underline{r}_j$ that have the property $\underline{P}_s$ is $\underline{x}$. Proportions satisfy the classical probability axioms.

Let a body of knowledge $\underline{K}$ be a set of sentences. We impose few restrictions on $\underline{K}$: we want it to be consistent in the sense that there should be no sentence $\underline{S}$ in $\underline{K}$ for which $\sim\underline{S}$ is also in $\underline{K}$; we want some logical truths in $\underline{K}$ (we could essentially enumerate them); and we want the items of possible statistical knowledge to be consistent statements concerning the relations of the possible reference classes and properties. (For example, one would not want to have "$(\underline{x})(\underline{P}_1(x) \leftrightarrow \underline{P}_2(\underline{x}))$", "$\%(\underline{r}_1,\underline{P}_1) \neq \%(\underline{r}_1,\underline{P}_2)$" both in $\underline{K}$. Finally, we want the (finite number of) sentences of the form "$\underline{P}_a(\underline{i}_q)$" together with "$0 = 0$" and "$0 = 1$" to generate a partition of all the sentences of the language under the relation of having the same truth value.

These constraints are embodied in the following axioms:

> A2.1 $\underline{R}$ is closed under intersection.
>
> A2.2 If $\underline{i} \neq \underline{j}$, $\underline{r}_i \in \underline{R}$, $\underline{r}_j \in \underline{R}$, then $\not\vdash \underline{r}_i = \underline{r}_j$.
>
> A2.3 If $\vdash \underline{r}_i \subset \underline{r}_j$ then "$\underline{r}_i \subset \underline{r}_j$" $\in \underline{K}$.
>
> A3.1 $\underline{P}$ is closed under conjunction and negation.

A3.2 If $\underline{i} \neq \underline{j}$, $\underline{P}_i \in \underline{P}$, $\underline{P}_j \in \underline{P}$, then $\nvdash \bigwedge \underline{x}(\underline{P}_i(\underline{x}) \leftrightarrow \underline{P}_j \underline{x})$.

A3.3 If $\vdash \bigwedge \underline{x}(\underline{P}_i(\underline{x}) \rightarrow \underline{P}_j(\underline{x}))$ then "$\bigwedge \underline{x}(\underline{P}_i(\underline{x}) \rightarrow \underline{P}_j(\underline{x}))$" $\in \underline{K}$.

A4　If "$\underline{i}_x \in \underline{r}_y$" $\in \underline{K}$ and "$\underline{i}_x \in \underline{r}_\omega$" $\in \underline{K}$, then

　　"$\underline{i}_x \in \underline{r}_y \cap \underline{r}_\omega$" $\in \underline{K}$.

A5.1 If "$\underline{S}_1 \leftrightarrow \underline{S}_2$" $\in \underline{K}$ then "$\underline{S}_2 \leftrightarrow \underline{S}_1$" $\in \underline{K}$

A5.2 If "$\underline{S}_1 \leftrightarrow \underline{S}_2$" and "$\underline{S}_2 \leftrightarrow \underline{S}_3$" are in $\underline{K}$, then "$\underline{S}_1 \leftrightarrow \underline{S}_3$" $\in \underline{K}$.

A5.3 "$\underline{S}_1 \leftrightarrow \underline{S}_1$" $\in \underline{K}$.

A6　For every non-mathematical* sentence $\underline{S}$ in $\underline{L}$, there exists

　　a $\underline{P}_y$, and exactly one $\underline{i}_x$, such that "$\underline{S} \leftrightarrow \underline{P}_y(\underline{i}_x)$" $\in \underline{K}$.

A7　There exists a model of the sentences in $\underline{K}$, with "$\underline{Z}(\underline{X},\underline{Y})$"

　　construed as the proportion of $\underline{X}$'s that are $\underline{Y}$'s.

We can now define the probability of a sentence $\underline{S}$ relative to a body of knowledge $\underline{K}$ to be $\underline{x}$ just in case $\underline{S}$ is known in $\underline{K}$ to be equivalent to a sentence of the form "$\underline{P}_z(\underline{i}_y)$" -- this is just to say that the biconditional "$\underline{S} \leftrightarrow \underline{P}_z(\underline{i}_y)$" is in $\underline{K}$ -- and for some reference class $\underline{r}_\omega$ to which $\underline{i}_y$ is known to belong, "$\underline{Z}(\underline{r}_\omega,\underline{P}_z) = \underline{x}$" is in $\underline{K}$, and, finally, if $\underline{r}_{\omega'}$ is another reference class to which $\underline{i}_y$ is known to belong, and "$\underline{Z}(\underline{r}_{\omega'},\underline{P}_z) <> \underline{x}$" is in $\underline{K}$, then it is known that $\underline{r}_\omega$ is included in $\underline{r}_{\omega'}$. (This is just to say that "$\underline{r}_\omega \subset \underline{r}_{\omega'}$" $\in \underline{K}$.) Formally:

D1　$\underline{Prob}(\underline{S},\underline{K}) = \underline{x}$ iff there are $\underline{P}_z$, $\underline{i}_y$, and $\underline{r}_\omega$ such that

　　(1) "$\underline{S} \leftrightarrow \underline{P}_z(\underline{i}_y)$" is in $\underline{K}$.

　　(2) "$\underline{i}_y \in \underline{r}_\omega$" is in $\underline{K}$.

　　(3) "$\underline{Z}(\underline{r}_\omega,\underline{P}_z) = \underline{x}$" is in $\underline{K}$.

　　(4) If "$\underline{i}_y \in \underline{r}_{\omega'}$" is in $\underline{K}$, and "$\underline{Z}(\underline{r}_{\omega'},\underline{P}_z) = \underline{x}$" is in $\underline{K}$,

　　　then "$\underline{r}_\omega \subset \underline{r}_{\omega'}$" is in $\underline{K}$.

Thus $\underline{r}_\omega$ is the __smallest__ reference class about which we have statistical information to which $\underline{i}_y$ is known to belong. This is essentially

Reichenbach's idea except for the addition of axiom A1.

We can generate the probability more clearly by putting the fourth

condition as a constraint on a table. Let the first column of the table

contain a list of all the reference classes r_w to which i_y is known to

belong. Let the second column contain the value of $x(r_w)$ from the

corresponding item of possible statistical knowledge: "$%(r_w,P_t) = x(r_w)$."

Work down the table, deleting every row that fails condition (4). (Rule: if

$x(r_{w_k}) \neq x(r_{w_l})$, delete both rows unless "$r_{w_k} \subset r_{w_l}$" is in $\underline{K}$.) There may

be several rows left, but they will all mention the same value of x. There

may be no rows left.

4. Limitations.

This approach deals perfectly reasonably with tosses of coins and the

like. It also does what we want for Sam, in the case in which he decides

whether or not to go to the movies by tossing a coin. But it has serious

drawbacks. It fails to provide for the case in which we get the probabilty

of Sam going to the movies from a limited statistical basis. It gives us no

probability at all when we know of i_y that it belongs to two reference

classes, our knowledge about those reference classes doesn't agree, and we

don't know that either reference class is included in the other. Example:

I know that chemistry majors go to the movies half the time on week nights; I

know that computer science majors go to the movies 10% of the time; Sam is a

double major in both computer science and chemistry; but I have no knowledge

of how frequently such double majors go to the movies.

The remedy is simple and obvious, but it entails considerable

complication. We allow items of possible statistical knowledge to embody

7

__approximate__ knowledge. To stick to one syntacical form, let us write

$$"Z(r_\omega, P_\xi) \in [x_1, x_2]"$$

to mean that the proportion of objects (or the chance of an object) in the reference class r_ω having the property P_ξ is in the closed interval $[x_1, x_2]$.

Suddenly we have statistical knowledge about every property and every reference class: at the very least we will know that the proportion lies in $[0,1]$. And now what do we mean by "$\langle \rangle$"? These changes work: Say that two intervals "differ" if neither is included in the other, and rewrite (4) to say that if r_ω and $r_{\omega'}$ differ, then r_ω is known to be included in $r_{\omega'}$:

(4') If "$i_y \in r_{\omega'}$" is in $\underline{K}$ and "$Z(r_{\omega'}, P_\xi) \in [x_1', x_2']$ " is in $\underline{K}$, and $[x_1', x_2']$ differs from $[x_1, x_2]$, then "$r_\omega \subset r_{\omega'}$" is in $\underline{K}$.

And, finally, we must include another clause to single out for us the most informative interval that is not ruled out be conflict with another interval:

(5) If $r_{\omega'}$ has not been eliminated as a possible reference class by the earlier conditions, then the interval corresponding to r_ω is a subinterval of the interval corresponding to $r_{\omega'}$.

This definition of probability is still limited — it turns out that we would like to other relations, in addition to the subset relation, to excuse "difference". And we would like to to able to consider equivalence to statements concerning several different individuals. (A general definition along these lines is provided in Kyburg 1985.) But it is already quite powerful, and it has some rather interesting properties:

(1) __All__ probabilities are objective, in the sense that each probability is based on empirical knowledge about frequencies or chances in the world.

(2) __Every__ statement in the language __has__ a probability: there is no distinction between statements concerning "repeatable" events and statements

8

concerning "unique" events.

(3) <u>No</u> <u>a</u> <u>priori</u> probabilities are required; all probabilities can be based on experience. (But <u>how</u> they can be so based is another story.)

5.Conclusion

Probabilistic knowledge may be regarded as all of a piece. There is no need to distinguish between "statistical" probabilities that have objective warrant in the world, and "subjective" probabilities that merely reflect our subjective feelings. When we apply our knowledge <u>of</u> statistical facts to individual cases,(as we always do when it comes to decision theory) it is the probability of a unique event that is at issue (the <u>next</u> toss). When we offer a "subjective" probability for a unique event, it is, if it has any epistemological justification at all, based on some (possibly approximate) statistical knowledge. The difference between the two cases lies in the fact that in the former case it is easy to specify the reference class -- it may even be built into the problem through the use of the indefinite articles "a" and "an" -- and in the latter case, it may be quite difficult to put your finger on the reference class. But this is a difference of degree, and not of kind; and the procedures suggested here, and in more detail in (Kyburg 1985) can render both kinds computable within quite rich languages.

references:

Reichenbach, Hans (1949) _Theory of Probability_. University of California

Press, Berkeley and Los Angeles.

Kyburg, Henry E. Jr. (1985) "The Reference Class," _Philosophy of Science_

1985.)

Note:

*   A non-mathematical sentence is one whose truth-value does not depend on

the truth of a mathematical statement. Thus the conjunction (or

disjunction) of "Heads on toss #7" and Fermat's last theorem is a

mathematical sentence, and so is "the toss indexed by the fourth prime

yields heads." But the latter is equivalent to "Toss number seven yields

heads." The point is to avoid worrying about undecidable, and even semi-

decidable, sentences.