Extracting the coherent core of human probability judgement: a research program for cognitive psychology

Daniel Osherson[*,a], Eldar Shafir[b], Edward E. Smith[c]

[a]*IDIAP, C.P. 609, CH-1920 Martigny, Valais, Switzerland*
[b]*Department of Psychology, Green Hall, Princeton University, Princeton, NJ 08544-1010, USA*
[c]*Department of Psychology, University of Michigan, 330 Packard Road, Ann Arbor, MI 48104, USA*

Abstract

Human intuition is a rich and useful guide to uncertain events in the environment but suffers from probabilistic incoherence in the technical sense. Developing methods for extracting a coherent body of judgement that is maximally consistent with a person's intuition is a challenging task for cognitive psychology, and also relevant to the construction of artificial expert systems. The present article motivates this problem, and outlines one approach to it.

1. Introduction

Human assessment of chances provides a guide to objective probabilities in a wide variety of circumstances. The survival of the species in diverse and rapidly evolving environments is testimony to this fact, as is the adequacy of our choices and judgements in most contexts encountered in daily life. At the same time, our assessments of chance are subject to systematic errors and biases that render them incompatible with the elementary axioms of probability. The character and origin of these errors have been the topic of intense scrutiny over several decades (for a partial review, see Osherson, in press).

How can the power and wisdom of human judgement be exploited while avoiding its weaknesses? One approach is to formulate principles of reasoning that simulate human judgement in large measure but correct it in view of probabilistic coherence. Such is the research program advocated and elaborated in our recent work (Osherson, Biolsi, Smith, Shafir, & Gualtierotti, 1994), and proposed in this article as a research program for cognitive psychology. The fundamental idea of extracting elements of human judgement for use within a normatively correct system of reasoning has already been articulated by Judea Pearl (1986, 1988). However, as explained below, our approach is different from, and complementary to, the tradition spawned by Pearl's studies. In a word, earlier approaches are "extensional" in character, assigning probabilities to unanalyzed statements and their logical combinations; in contrast, our approach is "intensional" inasmuch as it relies on a representation of the semantic content of the statements to which probabilities must be attached.

A specific implementation of our approach is presented in Osherson et al. (1994) but a somewhat different one will be summarized below. The goal of the present article is not to insist upon the details of any one approach but rather to highlight a research question that we find challenging and important. The question is: how can orderly and plausible judgement about uncertain events be extracted from the turmoil of human intuition?

We begin with general remarks on the difficulty of reasoning about probabilities in coherent fashion.

2. Coherent reasoning

2.1. The attraction of probabilistic reasoning

A system that recommends action in the face of uncertainty should quantify its estimates of chance in conformity with the axioms of probability.[1] Such is the burden of classical analyses of betting and prediction, which highlight the risks of reasoning non-probabilistically (see Cox, 1946; de Finetti, 1964, 1972; Jeffrey, 1983; Lindley, 1982; Resnik, 1987; Savage, 1972, for extended discussion). Alternative principles have been proposed to govern situations in which probabilities cannot generally be determined, as in Shafer (1976, 1986) and Shortliffe and Buchanan (1975). However, close examination of these principles (e.g., Fagin & Halpern, 1991; Heckerman, 1986) reinforces the conviction that

---

[1] Presentation and discussion of the elementary axioms of probability is provided in Resnik (1987, section 3.2). We do not rely here on the additional axiom of countable additivity (Ross, 1988, section 1.3), which is more controversial (see Kelly, 1993, section 13.4]).

probabilistic reasoning remains the standard of rigorous thinking in the face of uncertainty.

Although the foregoing conclusion remains the subject of lively debate and reflection (as in Dubois & Prade, 1991; Shafer & Pearl, 1990) it will be adopted without further justification in what follows.

2.2. The computational difficulty of probabilistic reasoning

Whatever its normative attractiveness, probabilistic reasoning poses difficult computational challenges. If probabilities must be distributed over the sentences of an expressive language, these difficulties can become insurmountable.[2] However, even when probabilities must be attributed to sentences without complex internal structure, certain manipulations are known to be intractable, for example updating Bayesian belief networks (see Cooper, 1987).

The root difficulty is the large number of events that must be kept in view in order to ensure probabilistic coherence. To explain the problem we now review some standard terminology. The remaining discussion bears exclusively on finite event spaces (the domain for most probabilistic expert systems). Moreover, instead of the term "event" we use the equivalent terminology of "statements" and "propositions". (For additional discussion, see Neapolitan, 1990, section 5.1; Pearl, 1988, section 2.1.)

2.3. Probability over statements

To establish a domain of discourse for reasoning, a finite number of statements $S_1, S_2, \ldots, S_N$ are fixed in advance. Each S_i is a determinate claim whose truth value may not be known with certainty, for example:

(1) Tulsa will accumulate more than 10 inches of rain in 1995.

The N statements give rise to 2^N *state descriptions*, each of the form:

$$\pm S_1 \wedge \cdots \wedge \pm S_N$$

where $\pm S_i$ means that S_i may or may not be negated. A state description is the logically strongest claim that can be made about the domain since it consists of the conjunction of every statement or its negation. A *proposition* is a subset of state

---

[2]This is shown for first-order arithmetical languages in Gaifman and Snir (1982, Theorem 3.7); the case of weak monadic second order logic is discussed in Stockmeyer, (1974, Theorem 6.1).

descriptions. Propositions are often denoted by the apparatus of propositional logic. For example, "$S_1 \vee S_2$" denotes the set of state descriptions in which at least one of S_1, S_2 occur positively. There are 2^{2^N} propositions.

A *distribution* is an assignment of non-negative numbers to the state descriptions, whose sum is one. A given distribution induces a probability for each proposition, obtained by summing the numbers associated with the state descriptions that make it up.

Let a collection **C** of propositions be given, and suppose that M is a mapping of **C** to real numbers. M is called *coherent* just in case there is some distribution **P** such that for all $X \in C$, $M(X) = P(X)$. Otherwise, M is *incoherent*.

2.4. Maintaining coherence in large domains

Now suppose that the domain in question embraces 500 statements $S_1, S_2 \ldots S_{500}$ concerning, let us say, five meteorological events in 100 American cities (we assume the events to be non-exclusive, referring to different months, for example). Suppose as well that some human agent $\mathscr{H}$ is asked to assign probabilities to a growing list $\mathscr{L}$ of propositions built from those statements. Each proposition is relatively simple in form, but the same statement may occur across different propositions. Thus, $\mathscr{L}$ might start off like this:

$$(2) \qquad \mathscr{L} = \left\{ \begin{array}{c} S_2 \vee \neg S_3 \\ S_{14} \to (S_2 \wedge S_1) \\ S_8 \wedge \neg S_{14} \\ \ldots \end{array} \right\}$$

Question: How can $\mathscr{H}$ assign coherent probabilities to the successive propositions in $\mathscr{L}$? In other words, as $\mathscr{H}$ associates numbers with more and more propositions, what procedure can be employed to ensure that the numbers are always extendible to a genuine probability distribution?

To achieve coherency, $\mathscr{H}$ could in principle proceed as follows.

Stage 1: Faced with the first proposition $S_2 \vee \neg S_3$, $\mathscr{H}$ writes down the four state descriptions based on statements S_2 and S_3, namely: $S_2 \wedge S_3$, $\neg S_2 \wedge S_3$, $S_2 \wedge \neg S_3$, $\neg S_2 \wedge \neg S_3$. Then $\mathscr{H}$ chooses a distribution over these state descriptions that reflects her beliefs about their respective probabilities. By summing over the first, third and fourth state descriptions she arrives at her probability for $S_2 \vee \neg S_3$. Her probabilities are coherent at the end of this stage.

Stage 2: Faced with the second proposition $S_{14} \to (S_2 \wedge S_1)$, she writes down the

16 state descriptions based on statements S_1, S_2, S_3 and S_{14}, that is, based on the four statements that appear in stages 1 and 2. Then $\mathcal{H}$ chooses a distribution over these state descriptions that meet two conditions, namely:

    (i) it is consistent with the distribution chosen for the four state descriptions of stage 1 (this is possible since her probabilities are coherent at the end of stage 1);

    (ii) it reflects her beliefs about the 16 state descriptions now in play.

By taking the appropriate sum, $\mathcal{H}$ arrives at her probability for $S_{14} \rightarrow (S_2 \wedge S_1)$. Because of property (i) her probability for the first proposition $S_2 \vee \neg S_3$ may be recovered by adding the relevant state descriptions from among the current 16; the probability assigned to $S_2 \vee \neg S_3$ will not have changed from stage 1. Consequently, the totality of her probability attributions are still coherent at the end of this stage.

Stage 3: Faced with the third proposition $S_8 \wedge \neg S_{14}$, she writes down the 32 state descriptions based on statements S_1, S_2, S_3, S_8, S_{14}, ... etc.

The disadvantage of this procedure is that it soon requires excessive space to write down the list of needed state descriptions. Eventually, 2^{500} state descriptions need be written down at once, an enormous number.

It is not immediately obvious what procedure $\mathcal{H}$ can substitute for this one. Suppose, for example, that $\mathcal{H}$ attempts at each stage to limit attention to just the state descriptions needed in the evaluation of the current proposition. Thus, in the first stage, $\mathcal{H}$ would attribute probabilities only to the four state descriptions based on S_2 and S_3; in the second stage, $\mathcal{H}$ would attribute probabilities only to the eight state descriptions based on S_1, S_2 and S_{14}; and so forth. Let us assume, furthermore, that $\mathcal{H}$ chooses her probabilities in coherent fashion at each stage. This procedure is nonetheless insufficient to guarantee the coherence of $\mathcal{H}$'s judgement since it ignores logical dependencies among propositions at different stages. To take the simplest example, suppose that the state descriptions of (I) show up in $\mathcal{L}$ at one stage, and those of (II) show up later:

$$(I) \left\{ \begin{array}{c} S_1 \wedge S_2 \\ \neg S_1 \wedge S_2 \\ S_1 \wedge \neg S_2 \\ \neg S_1 \wedge \neg S_2 \end{array} \right\} \qquad (II) \left\{ \begin{array}{c} S_2 \wedge S_3 \\ \neg S_2 \wedge S_3 \\ S_2 \wedge \neg S_3 \\ \neg S_2 \wedge \neg S_3 \end{array} \right\}$$

Then overall coherence requires that the sum of the probabilities assigned to the first two state descriptions of (I) equal the sum of the probabilities assigned to the first and third state descriptions of (II). Otherwise, the two distributions imply different values for the statement S_2, violating overall coherence. It is thus clear that the revised procedure suggested above is inadequate without some means of keeping track of the combinations of state descriptions seen in $\mathscr{L}$ up to the current point, and this entails the same combinatorial explosion as before.

The difficulty of maintaining coherence over large domains is widely recognized. For example, Clark Glymour (1992, p. 361) summarizes the matter this way:

> To represent an arbitrary probability distribution, we must specify the value of the probability function for each of the state descriptions. So with 50 atomic sentences, for many probability distributions we must store 2^{50} numbers to represent the entire distribution. ... We cannot keep 2^{50} parameters in our heads, let alone 2 raised to the power of a few thousand, which is what would be required to represent a probability distribution over a realistic language. ... For such a language, therefore, there will be cases in which our beliefs are inconsistent and our degrees of belief incoherent.

Finally, note that the problem is aggravated by the expressiveness of the language used for ordinary thought. For example, we can easily express 100 predicates that apply sensibly to any of 100 grammatical subjects. There result 10,000 statements concerning which an agent might wish to reason probabilistically. Is there any hope of carrying out such reasoning coherently?

2.5. Tractability via conditional independence

Some distributions have special properties that allow coherent reasoning within modest computational bounds. For example, consider a distribution **P** over statements $S_1, \ldots, S_N$ such that for all subsets $\{S_{1'}, \ldots S_{r'}\}$ of $\{S_1, \ldots S_N\}$, $\mathbf{P}(S_{1'} \wedge \cdots \wedge S_{r'}) = \mathbf{P}(S_{1'}) \times \cdots \times \mathbf{P}(S_{r'})$, that is, in which the underlying statements are mutually stochastically independent. Reasoning according to **P** does not require storing the probability of each state description. It suffices to store only the probabilities associated with $S_1, \ldots S_N$ since the probability of any needed state description can be recovered through multiplication (relying where needed on the fact that $\mathbf{P}(\neg S_i) = 1 - \mathbf{P}(S_i)$). It follows that if the beliefs of our agent $\mathscr{H}$ are mutually independent in the foregoing sense then she can confront the list $\mathscr{L}$ with no fear of incoherence. For each proposition X of $\mathscr{L}$, $\mathscr{H}$ need only carry out the following steps: (a) list the statement letters occurring in X; (b) decide which state descriptions over the foregoing list imply X; and (c) take the sum of the probabilities of the latter state descriptions as the final answer. Using this strategy, not only can $\mathscr{H}$ assign coherent probabilities to all the propositions that

might arise in $\mathscr{L}$ (assuming that each proposition remains reasonably short). In addition, her judgement exhibits "path independence" in the sense that reordering $\mathscr{L}$ will not change $\mathscr{H}$'s probability for any proposition.

The mutual independence of $S_1, \ldots S_N$ is an unrealistic assumption in most situations. Weaker but still useful forms of independence can be defined (see Whittaker, 1990, for extended discussion). For example, if $\mathbf{P}(S_1 \mid S_2 \wedge S_3) = \mathbf{P}(S_1 \mid S_3)$ then S_1 is said to be "conditionally independent" of S_2 given S_3 (according to $\mathbf{P}$). If $\mathbf{P}$ exhibits a felicitous pattern of conditional independence, then its underlying state descriptions can be factored in such a way as to render their manipulation computationally tractable. A variety of schemes for exploiting conditional independence have been devised (e.g., Geiger, Verma, & Pearl, 1990; Heckerman, 1990b; Lauritzen & Spiegelhalter, 1988; Andreassen, Woldbye, Falck, & Andersen, 1989; Olesen et al., 1989; Long, Naimi, Criscitiello, & Jayes, 1987). Unfortunately, even with the right configuration of conditional independencies, many of these systems require manual entry of an excessive number of probabilities and conditional probabilities, often minute in magnitude. Usually the probabilities cannot be assessed in actuarial fashion, so must be based on the judgement of experts (e.g., doctors). It is well known that experts are apt to provide incoherent estimates of probability (see Casscells, Schoenberger, & Grayboys, 1978; Kahneman & Tversky, 1972; Tversky & Kahneman, 1983; Winterfeld & Edwards, 1986, Ch. 4.5), and the numbing task of making thousands of judgements no doubt aggravates this tendency.

Several responses may be envisioned to the foregoing problem. First, the interrogation of experts can be rationalized and simplified (as in Heckerman, 1990a). Second, procedures can be devised to reduce the effect of judgemental biases that lead to incoherence (as discussed in Kahneman, Slovic, & Tversky, 1982, Chs. 30–32; Winterfeld & Edwards, 1986). Third, techniques can be implemented for finding a probability distribution that is maximally close to a set of possibly incoherent judgements (see Osherson et al., 1994). Fourth, methods can be invented for constructing human-like distributions on the basis of judgements that are psychologically more natural than assessments of the probability of arbitrary propositions.

The fourth response has been raised in Szolovits and Pauker (1978). It is the one advocated here, though not to the exclusion of the other three. The essential innovation of our approach is to attempt to derive probabilities on the basis of the "semantic" (really, "informational") content of the grammatical constituents that compose statements. The potential benefit is reduction in the amount of information that must be culled from a human informant (and later stored). The reduction is based on the combinatorial mechanisms of grammar, which allow a large number of statements to be generated from a small number of constituents. Let us now examine this idea.

3. Statement semantics

How can the meaning of statements best be represented in view of deriving human-like probabilities from their representations? Surely we are far from having the answer to this question. In the hope of stimulating discussion, the present section describes a simple approach that analyzes grammatical constituents along a fixed shock of dimensions. We begin by specifying the kind of statement to be treated.

3.1. Subjects, predicates, objects

A statement like

(3)   Lawyers seldom blush

decomposes into a grammatical subject (namely, "Lawyers") and a grammatical predicate (namely, "seldom blush"). Henceforth we employ the term "object" in place of "grammatical subject" in order to prevent confusion with the "subjects" participating in psychological experiments. Thus, "Lawyers" and "seldom blush" are the object and predicate, respectively, of statement (3). We limit attention to statements of this simple object–predicate form. Given object o and predicate p, we use $[o, p]$ to denote the statement formed from them.

A domain of reasoning is established by fixing a (finite) list **obj** of objects and a (finite) list **pred** of predicates and then considering the set of statements $S = \{[o, p] \mid o \in \textbf{obj} \text{ and } p \in \textbf{pred}\}$. For simplicity in what follows we assume that all the statements in S are meaningful, neither contradictory nor analytic, and logically independent from each other.

3.2. Vectorial representations

Our approach associates each object and predicate with a real vector in n dimensions, for some fixed value of n. Such a vector may be conceived as a rating of the object or predicate along n dimensions (e.g., for $n = 3$, a rating of the object TIGER along the dimensions *size*, *speed* and *ferocity*). The vector is intended to code a given person's knowledge (or mere belief) about the item in question (see Smith & Medin, 1981, for discussion).

Vector representations might seem impoverished compared to "frames" or other elaborate schemes for knowledge representation (Bobrow & Winograd, 1976; Minsky, 1981, 1986). It is thus worth noting the considerable representational power of real vectors. Suppose that person **P** is chosen, and let us say that

predicate p "fits" object o just in case the probability of $[o, p]$ according to **P** exceeds .5 (any other threshold would serve as well). This may be abbreviated to: $\mathbf{P}([o, p]) > .5$. We would like to represent the fit-relation in terms of vectors. For this purpose, suppose that n-dimensional vectors are assigned to **obj** ∪ **pred**, one per object and one per predicate. Given such an assignment, let us say that o "dominates" p just in case the coordinates of o's vector are at least as great as the corresponding coordinates of p's vector. We have the following fact, proved in Doignon, Ducamp, and Falmagne (1984):

(4) Let **P** embody any fit-relation whatsoever. Then for some n, there is an assignment of n-dimensional vectors to **obj** ∪ **pred** such that for all $o \in$ **obj** and all $p \in$ **pred**, $\mathbf{P}([o, p]) > .5$ if and only if o dominates p. Moreover, n can be chosen to not exceed the smaller of:

> the cardinality of **obj**,
> the cardinality of **pred**.

Intuitively, we can think of the vector assigned to a predicate as establishing criteria for membership in the associated category. (For example, the predicate "can learn a four choice-point maze in three trials" might have a requirement of .75 in the coordinate corresponding to *intelligent*.) For o to have greater than .5 probability of possessing p, o's values at each coordinate must exceed the criterion established by p. Fact (4) shows that such a scheme is perfectly general for representing probability thresholds and it renders plausible the idea that real vectors might also serve to predict continuous assessments of the probability of statements.

3.3. To and from statement representations

Recall that our goal is to capture the coherent core of a person's judgement about chances. Call the person at issue $\mathcal{H}$. Having decided to use vectors to represent **obj** ∪ **pred**, two questions remain to be answered. These are:

(a) Which particular object and predicate are attributed to $\mathcal{H}$?
(b) How are object and predicate vectors translated into a probability distribution?

Once answers are offered to (a) and (b), a third question may be addressed, namely:

(c) If $\mathcal{H}$'s vectors are fixed in accordance with the answer to (a), and if

probabilities are subsequently assigned to propositions in accordance with the answer to (b), how well do the resulting probabilities accord with $\mathcal{H}$'s judgement about chances? Does the processed and regimented output of our system retain any of the insight that characterizes $\mathcal{H}$'s understanding about probabilities in the environment?

Let us now briefly consider (a)–(c).

3.4. Fixing object and predicate vectors

One means of obtaining vectors is to request the needed information directly from $\mathcal{H}$ via feature ratings (as in Osherson, Stern, Wilkie, Stob, & Smith, 1991). A less direct approach is to infer the vectors from similarity ratings among objects and predicates. In this case, we work backwards from a suitable vector-based model of similarity (e.g., those discussed in Osherson, 1987; Suppes, Krantz, Luce, & Tversky, 1989; Tversky, 1977), looking for vectors that best predict $\mathcal{H}$'s similarity data. Another strategy is to postulate a model of simple probability judgements based on the needed vectorial representations, and then work backwards to vectors from such judgements. In this case, our system carries out "extrapolation", extending a small set of probability judgements to a more complete set (see Osherson, Smith, Meyers, Shafir, & Stob, in press).

3.5. Vectors to probabilities

Turning to question (b) above, we describe one procedure for synthesizing probabilities from the vectors underlying **obj** and **pred**. It rests upon a scheme for constructing three-dimensional Venn diagrams. Specifically, the pair of vectors associated with object o and predicate p is translated into a subregion $\mathcal{R}$ of the unit cube.[3] The volume of $\mathcal{R}$ represents the probability of $[o, p]$. The position of $\mathcal{R}$ determines its intersection with subregions assigned to other statements. The probability of a complex proposition (e.g., the intersection or the union of two statements) may then be determined by calculating the volume of the corresponding region. It is easy to see that use of the diagram guarantees probabilistic coherence.[4]

Let us now outline a simple scheme for selecting the particular region assigned

---

[3] The unit cube has sides of length 1. It is used for convenience in what follows; various other kinds of solids would serve as well.

[4] There is no mathematical reason to limit the diagram to three dimensions. Volumes in the n-dimensional unit cube for any positive n yield *bona fide* distributions. So far our experiments indicate that three dimensions are enough.

to a given statement $[o, p]$. Let O, P be the vectors underlying o, p, respectively, and suppose them to be suitably normalized so that all coordinates fall into the interval $[0, 1]$. Define the "O, P-box" to be the (unique) rectangular solid $\mathscr{R}$ with the following properties:

(5) (a) O falls within $\mathscr{R}$;
    (b) for $i \leq 3$ the length of the ith side of $\mathscr{R}$ is 1 minus the absolute difference between O_i and P_i;
    (c) within the foregoing constraints, P is as close as possible to the geometrical center of $\mathscr{R}$.

It may be seen that the volume of $\mathscr{R}$ varies directly with the match of P's coordinates to those of O; statements based on compatible objects and predicates are accorded higher probability thereby. Moreover, $\mathscr{R}$'s position in the cube represents aspects of the semantics of $[o, p]$. For example, if p and q are complementary predicates with contrasting vectors then (5) assigns $[o, p]$ and $[o, q]$ boxes with little or no intersection. This reflects the low probability that must sensibly be assigned to $[o, p] \wedge [o, q]$ in view of the incompatible contents of p and q.

Many alternatives to (5) are possible. To serve as a computationally tractable means of coherent reasoning in large domains it suffices to meet the following condition:

(C) Given a point x in the unit cube, and vectors O, P underlying object o and predicate p, it must be computationally easy to determine whether x lies in the region associated with $[o, p]$.

In this case it is straightforward to calculate the volumes associated with any Boolean combination of statements, hence with any proposition.[5] It is clear that (5) satisfies **C**.

Observe that within any Venn diagram scheme that conforms to **C**, coherent probabilistic reasoning can proceed without storing an entire distribution. It is enough to store the vectors underlying objects and predicates since the volume associated with any given proposition (of reasonable length) can be easily retrieved from the vectors. Thus, given 10 objects and 10 predicates, only 20 vectors need be stored. This is easily achieved even for vectors of considerable size. In contrast, 10 objects and 10 predicates give rise to 100 statements and thus to a distribution with 2^{100} state descriptions. A potential solution to the problem of coherent reasoning, posed in section 2.4, is offered thereby.

It must be emphasized that not every distribution can be represented by a Venn

---

[5] A more careful formulation of **C** would refer to ϵ-spheres in place of points x, etc.

diagram that meets **C** (just as not every distribution manifests conditional independencies of a computationally convenient kind). The question thus arises: do distributions that conform to **C** approximate human intuition about chance in a wide variety of domains? We are thus led to question (c) above, namely, whether the distribution delivered by our method resembles the original intuitions of subject $\mathscr{H}$. Sticking with the simple scheme in (5) – henceforth called the "Venn model" – let us now address this matter.

3.6. Accuracy of the method

We summarize one experimental test of our method. By an *elementary argument* (over **obj** ∪ **pred**) is meant a non-empty set of statements, one of which is designated as "conclusion", the remainder (if any) as "premises". Statements are considered special cases of elementary arguments, in which the premise set is empty. An argument may be conceived as an invitation to evaluate the probability of its conclusion while assuming the truth of its premises. Thirty college students evaluated 80 elementary arguments based on four mammals (which served as objects) and two predicates (e.g., "are more likely to exhibit 'fight' than 'flight' posture when startled"). For each subject, an individually randomized selection of 30 arguments was used to fix vectors representing his objects and predicates. This was achieved by working backwards from the Venn model, seeking vectors that maximize its fit to the subject's judgement about the 30 input arguments. The Venn model was then applied to the resulting vectors to produce probabilities for the remaining 50 arguments. For each subject we calculated the average, absolute deviation between the Venn model's predictions for the 50 arguments and the probabilities offered directly by the subject. Pearson correlations between the two sets of numbers were also calculated.

The median, average absolute deviation between the observed probabilities assigned to a subject's 50 predicted arguments and the probabilities generated by the Venn model is .11.[6] The correlation between the two sets of numbers is .78. The results suggest that the Venn method can extrapolate a coherent set of probabilities from a small input set, and do this in such a way that the extrapolated distribution provides a reasonable approximation to the judgement of the person providing input. The input set of probabilities need not be coherent.

---

[6] This deviation can be compared to the following statistic. Consider the mean of the probabilities assigned to the 30 arguments used to fix the object and predicate vectors of a given subject. We may use this single number as a predictor of the probabilities assigned to the remaining 50 arguments. In this case the median, average absolute deviation between the observed and predicted probabilities is .20.

4. Concluding remarks

The problem of recovering the coherent core of human probability judgement strikes us as an important project for cognitive psychology. It unites theorizing about the mental mechanisms of reasoning with a practical problem for expert systems, namely, finding an exploitable source of Bayesian priors. The system sketched above is preliminary in character, and serves merely to suggest the feasibility of the research program we advocate.

Psychological research in recent years has produced considerable understanding of the character and causes of incoherent reasoning, even if debate continues about its scope and interpretation (see Gigerenzer & Murray, 1987; Osherson, 1990; Shafir & Tversky, 1992; Tversky & Shafir, 1992, and references cited there). It was noted in section 2.1 that probabilistic coherence has non-trivial justification as a standard – however incomplete – of normatively acceptable reasoning. We thus take there to be good empirical evidence, plus great computational plausibility, in favor of the thesis that human judgement is imperfect from the normative point of view. This thesis does not, however, impugn every aspect of ordinary reasoning. Indeed, the merits of human judgement have often been emphasized by the very researchers who investigate its drawbacks (e.g., Nisbett & Ross, 1980, p. 14). A challenge is posed thereby, namely, to devise methods that distill the rational component of human thought, isolating it from the faulty intuition that sometimes clouds our reason. Such appears to have been the goal of early inquiry into probability and utility (Gigerenzer et al., 1989, Ch. 1). It remains a worthy aim today.

References

Andreassen, S., Woldbye, M., Falck, B., & Andersen, S. (1989). Munin: A causal probabilistic network for interpretation of electromyographic findings. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*.

Bobrow, D., & Winograd, T. (1976). An overview of KRL, a knowledge representation language. *Cognitive Science, 1*, 3–46.

Casscells, W., Schoenberger, A., & Grayboys, T. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine, 299*, 999–1000.

Cooper, G.F. (1987). Probabilistic inference using belief networks is np-hard. *Memo KSL-87-27*, Knowledge Systems Laboratory, Stanford University, May 1987.

Cox, R. (1946). Probability, frequency, and reasonable expectation. *American Journal of Physics, 14*, 1–13.

de Finetti, B. (1964). La prévision: Ses lois logiques, ses sources subjectives [transl. into English]. In H. Kyburg & P. Smokler (Eds.), *Studies in subjective probability*. New York: Wiley.

de Finetti, B. (1972). *Probability, induction and statistics*. New York: Wiley.

Doignon, J.-P., Ducamp, A., & Falmagne, J.-C. (1984). On realizable biorders and the biorder dimension of a relation. *Journal of Mathematical Psychology, 28*, 73–109.

Dubois, D., & Prade, H. (1991) Updating with belief functions, ordinal conditional functions and

possibility measures. In P.P. Bonissone, M. Henrion, L.N. Kanal, & J.F. Lemmer (Eds.), *Proceedings of the Sixth Workshop on Uncertainty in Artificial Intelligence* (pp. 311–329). Amsterdam: Elsevier.

Fagin, R., & Halpern, J. (1991). A new approach to updating beliefs. In P.P. Bonissone, M. Henrion, L.N. Kanal, & J.F. Lemmer (Eds.), *Proceedings of the Sixth Workshop on Uncertainty in Artificial Intelligence* (pp. 347–374). Amsterdam: Elsevier.

Gaifman, H. & Snir, M. (1982). Probabilities over rich languages. *Journal of Symbolic Logic, 47,* 495–548.

Geiger, D., Verma, T., & Pearl, J. (1990). d-Separation: From theorems to algorithms. In R.D. Shachter, M. Henrion, L.N. Kanal, & J.F. Lemmer (Eds.), *Uncertainty in artificial intelligence 5.* Amsterdam: North-Holland.

Gigerenzer, G., & Murray, D.J. (1987). *Cognition as intuitive statistics.* Hillsdale, NJ: Erlbaum.

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Kruger, L. (1989). *The empire of chance.* Cambridge, UK: Cambridge University Press.

Glymour, C. (1992). *Thinking things through.* Cambridge, MA: MIT Press.

Heckerman, D. (1986). Probabilistic interpretations for MYCIN's certainty factors. In L.N. Kanal & J.F. Lemmer (Eds.), *Uncertainty in artificial intelligence.* Amsterdam: North-Holland.

Heckerman, D. (1990a). Probabilistic similarity networks.

Heckerman, D. (1990b). A tractable inference algorithm for diagnosing multiple diseases. In R.D. Schachter, M. Henrion, L.N. Kanal, & J.F. Lemmer (Eds.), *Uncertainty in artificial intelligence 5.* Amsterdam: North-Holland.

Jeffrey, R. (1983). *The logic of decision* (2nd edn.). Chicago: University of Chicago Press.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge, UK: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgement of representativeness. *Cognitive Psychology, 3,* 430–454.

Kelly, K.T. (1993). *The logic of reliable inquiry.* MS, Department of Philosophy, Carnegie Mellon University.

Lauritzen, S., & Spiegelhalter, D. (1988). Local computations with probabilities on graphical structures and their applications to expert systems. *Journal of the Royal Statistical Society, B, 50,* 157–224.

Lindley, D. (1982). Scoring rules and the inevitability of probability. *International Statistical Review, 50,* 1–26.

Long, W., Naimi, S., Criscitiello, M., & Jayes, R. (1987). The development and use of a causal model for reasoning about heart failure. *IEEE Symposium on Computer Applications in Medical Care* (pp. 30–36).

Minsky, M. (1981). A framework for representing knowledge. In J. Haugeland (Ed.), *Mind design.* Cambridge, MA: MIT Press.

Minsky, M. (1986). *The society of mind.* New York: Simon & Schuster.

Neapolitan, R. (1990). *Probabilistic reasoning in expert systems.* New York: Wiley.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Olesen, K.G., Kjaerulff, U., Jensen, F., Jensen, F.V., Falck, B., Andreassen, S., & Andersen, S.K. (1989). A munin network for the median nerve: A case study on loops. *Applied Artificial Intelligence, Special Issue: Towards Causal AI Models in Practice.*

Osherson, D. (1987). New axioms for the contrast model of similarity. *Journal of Mathematical Psychology, 31,* 93–103.

Osherson, D. (1990). Judgment. In D. Osherson & E.E. Smith (Eds.), *Invitation to cognitive science: Thinking* (pp. 55–88). Cambridge, MA: MIT Press.

Osherson, D. (in press). Probability judgment. In E.E. Smith, & D. Osherson (Eds.), *Invitation to cognitive science: Thinking* (2nd edn.). Cambridge, MA: MIT Press.

Osherson, D., Biolsi, K., Smith, E., Shafir, E., & Gualtierotti, A. (1994). A source of Bayesian priors. IDIAP Technical Report 94-03.

Osherson, D., Smith, E.E., Meyers, T.S., Shafir, E., & Stob, M. (in press). Extrapolating human probability judgment. *Theory and Decision.*

Osherson, D., Stern, J., Wilkie, O., Stob, M., & Smith, E. (1991). Default probability, *Cognitive Science, 15*, 251–270.

Pearl, J. (1986). Fusion, propogation, and structuring in belief networks. *Artificial Intelligence, 29*, 241–288.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference.* San Mateo, CA: Morgan-Kaufmann.

Resnik, M. (1987). *Choices: An introduction to decision theory.* Minneapolis: University of Minnesota Press.

Ross S. (1988). *A first course in probability.* New York: Macmillan.

Savage L. (1972). *The foundations of statistics.* New York: Dover.

Shafter, G. (1976). *A mathematical theory of evidence.* Princeton, NJ: Princeton University Press.

Shafer, G. (1986). Constructive probability. *Synthese, 48*, 1–60.

Shafer, G., & Pearl, J. (Eds.) (1990). *Readings in uncertain reasoning.* San Mateo, CA: Morgan-Kaufmann.

Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology, 24*, 449–474.

Shortliffe, E., & Buchanan, B. (1975). A model of inexact reasoning in medicine. *Mathematical Biosciences, 23*, 351–379.

Smith, E.E., & Medin, D. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Stockmeyer, L. (1974). *The complexity of decision problems in automata theory and logic.* Ph.D. thesis, MIT.

Suppes, P., Krantz, D.H., Luce, R.D., & Tversky, A. (1989). *Foundations of measurement* (Vol. II). San Diego: Academic Press.

Szolovits, P., & Pauker, S. (1978). Categorical and probabilistic reasoning in medical diagnosis. *Artificial Intelligence, 11*, 115–144.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327–362.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. *Psychological Review, 90*, 293–315.

Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science, 5*, 305–309.

Whittaker, J. (1990). *Graphical models in applied multivariate statistics.* New York: Wiley.

Winterfeld, D.V., & Edwards, W. (1986). *Decision analysis and behavioral research.* New York: Cambridge University Press.