Probability from similarity*

Sergey Blok
Northwestern University

Douglas Medin
Northwestern University

Daniel Osherson
Rice University

September 24, 2002

Abstract

The probability calculus provides an attractive canonical form for reasoning but its use requires numerous estimates of chance. Some of the estimates needed in artificial systems can be recorded individually or via Bayesian networks. Others can be tabulated as relative frequencies from stored data. For the shifting contexts of commonsense reasoning, however, the latter sources are likely to prove insufficient. To help fill the gap, we show how sensible conditional probabilities can be derived from absolute probabilities plus information about the *similarity* of objects and categories. Experimental evidence from studies of human reasoning documents the naturalness of the numbers we derive.

1 Introduction

For the probability calculus to serve as the "faithful guardian of common sense" [15], a great many estimates of chance are needed. Even when the domain can be structured by conditional independence (as assumed in medical diagnosis [8]), numerous probabilities must often be recorded. Without such structure, the source of probability estimates becomes a yet more pressing issue, especially for autonomous agents exposed to novel situations in everyday life.

It has been observed that the needed estimates "can come either from the knowledge engineer's (or expert's) subjective experience, or from measurements of frequencies in a database of past experiences, or from some combination of the two" [17]. Little attention appears to have been devoted to forging probabilities from other kinds of information that may be available to a reasoning agent. The latter information includes the pairwise *similarity* between objects or

*Research supported by NSF grants 9978135 and 9983260. Contact information: D. Osherson, MS 25, Rice University, P.O. Box 1892, Houston TX 77251-1892 USA. Electronic mail: osherson@rice.edu.

categories, e.g., that a German Shepherd is more similar to a Labrador than to a Chihuahua. In a competent reasoner, some means of calculating similarity is likely to be present in any event. For similarity is essential to categorization [20, 18], and to evaluating substitutions in a search task (e.g., whether to retrieve a Labrador or a Chihuahua when a German Shepherd cannot be found). It also plays a role in naive inductive inference, at least at the qualitative level [16, 14].

Similarity is a particularly intriguing raw material for probability computations because it can often be derived from facts easily discovered by an autonomous agent. Thus, feature-based models of similarity [21, 12] rely on little more than counting the predicates that apply to objects. Likewise, geometrical models [19] involve the positions of objects along quantifiable dimensions like size. (See [11, 6] for further discussion.)

The goal of the present paper is to argue for the feasibility of extracting estimates of probability from similarity. Rather than treating the problem in full generality, we focus on a narrow class of conditional probabilities $\mathsf{Pr}\,(p \mid q_1 \cdots q_n)$, described in the next section. The derivations depend on no more than the (absolute) probabilities attached to $p, q_1 \cdots q_n$ and the similarity of objects (or categories) mentioned in the latter statements. Our proposals conform to qualitative conditions that must be met by sensible estimates of chance. It will also be shown that the numbers implied by our formulas are close to estimates of conditional probabilities produced by students asked to think about salary prospects after graduating from familiar colleges.

We stress that our results are intended merely to illustrate the larger enterprise of culling useful probabilities from varied information available to the reasoner. Similarity is only one type of information that ultimately influences human estimates of chance.

We conclude this introductory section with a remark about *probabilistic coherence*. An agent may endorse $\mathsf{Pr}\,(p \wedge q \wedge r) = .5$ and $\mathsf{Pr}\,(q) = .8$ and then use similarity to derive $\mathsf{Pr}\,(p \mid q) = .6$. Together, the three judgments are *incoherent* in the sense that no joint distribution generates all three.[1] Incoherent probabilities are a questionable basis for reasoning, so a similarity-based estimate of chance seems to be constrained to take into account the agent's pre-existing probabilities. This is a daunting prospect if the latter set is numerous. It might nonetheless be safe to generate new probabilities in isolation from earlier ones if there is an efficient way to "rectify" all the estimates in a second step, minimally adjusting them to achieve coherence. In

[1] They imply that $\mathsf{Pr}\,(p \wedge q) = \mathsf{Pr}\,(q) \times (\mathsf{Pr}\,(p \wedge q)) / \mathsf{Pr}\,(q) = \mathsf{Pr}\,(q) \times \mathsf{Pr}\,(p \mid q) = .8 \times .6 = .48 < .5 = \mathsf{Pr}\,(p \wedge q \wedge r)$. No distribution assigns lower probability to $p \wedge q$ than to $p \wedge q \wedge r$.

fact, recent progress [2, 5] in devising algorithms for rectifying incoherent judgment warrants exploring the two step strategy for expanding a corpus of probabilistic beliefs by similarity-derived estimates of chance. We proceed on this basis, thus ignoring most of the probability corpus already available to the reasoner prior to her exploitation of similarity.

2 Theory

First we specify the class of conditional probabilities to be derived. Then we present and discuss formulas for carrying out the derivation.

2.1 Cases to be considered

Let Q represent a given predicate like "has trichromatic vision." Letters a, b, c stand for objects or categories to which Q can be meaningfully applied, e.g., *foxes, wolves, goldfish*. We assume that a, b, c are all at the same conceptual/hierarchical level. Thus, if a is *foxes* then b cannot be a particular fox nor the class of canines. The intent of this requirement is to allow similarity to be naturally assessed among all pairs from a, b, c. The statements figuring in our analysis have the form Qk or $\neg Qk$, where k is one of a, b, c. We consider the following types of conditional probabilities.

(1)
- (a) $\mathsf{Pr}(Qc \mid Qa)$
- (b) $\mathsf{Pr}(Qc \mid \neg Qa)$
- (c) $\mathsf{Pr}(\neg Qc \mid Qa)$
- (d) $\mathsf{Pr}(\neg Qc \mid \neg Qa)$
- (e) $\mathsf{Pr}(Qc \mid Qa, Qb)$
- (f) $\mathsf{Pr}(\neg Qc \mid \neg Qa, \neg Qb)$
- (g) $\mathsf{Pr}(\neg Qc \mid Qa, Qb)$
- (h) $\mathsf{Pr}(Qc \mid \neg Qa, \neg Qb)$
- (i) $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ and $\mathsf{Pr}(Qc \mid Qa, \neg Qb)$
- (j) $\mathsf{Pr}(\neg Qc \mid \neg Qa, Qb)$ and $\mathsf{Pr}(\neg Qc \mid Qa, \neg Qb)$

To predict the conditional probabilities in (1), we allow ourselves no more than (i) the probability of each of the three statements Qa, Qb, Qc, and (ii) the pairwise similarities among each pair in $\{a, b, c\}$. We denote the similarity between objects x, y by $\mathsf{sim}(x, y)$, and assume that it is scaled on the unit interval with 1 representing identity. That is, $\mathsf{sim}(x, y) \in [0, 1]$ and $\mathsf{sim}(x, y) = 1$ iff $x = y$. It is further assumed that sim is a symmetric function, that is,

$\mathsf{sim}(x, y) = \mathsf{sim}(y, x)$. The symmetry assumption is consistent with the bulk of human judgment [1] even though it may be violated in rare circumstances [21]. The symmetry of $\mathsf{sim}(\cdot, \cdot)$ ensures that similarity is not a disguised judgment of conditional probability inasmuch as $\mathsf{Pr}(\cdot \mid \cdot)$ is not symmetric in its arguments.

For any pair of statements p, q, at least one of $\mathsf{Pr}(p \mid q) \geq \mathsf{Pr}(p)$ and $\mathsf{Pr}(p \mid \neg q) \geq \mathsf{Pr}(p)$ must hold. It simplifies notation to assume:

(2) Confirmation Assumption: $\mathsf{Pr}(Qc \mid Qa) \geq \mathsf{Pr}(Qc)$ and $\mathsf{Pr}(Qc \mid Qb) \geq \mathsf{Pr}(Qc)$.

The stimuli used in our experiments have been designed to satisfy (2). Their construction was guided by the authors' intuition rather than theory, however. We return to this theoretical gap in the discussion section.

2.2 Formulas

The appendix lists the formulas used to construct the conditional probabilities in (1). We discuss cases (1)(a),(e),(i). The others are straightforward variants.

2.2.1 $\mathsf{Pr}(Qc \mid Qa)$**.** Any formula for constructing $\mathsf{Pr}(Qc \mid Qa)$ on the basis of $\mathsf{Pr}(Qa)$, $\mathsf{Pr}(Qc)$ and $\mathsf{sim}(a, c)$ should meet certain qualitative conditions. Trivially, the formula must ensure that $0 \leq \mathsf{Pr}(Qc \mid Qa) \leq 1$. More substantively, as $\mathsf{sim}(a, c)$ approaches 1, $\mathsf{Pr}(Qc \mid Qa)$ should approach 1. For if $\mathsf{sim}(a, c) \approx 1$ then $\mathsf{Pr}(Qc \mid Qa) \approx \mathsf{Pr}(Qc \mid Qc) = 1$. (Thus, the conditional probability that pigs have trichromatic vision given that the hogs do is close to unity given the similarity of these creatures.) On the other hand, as $\mathsf{sim}(a, c)$ goes to 0, $\mathsf{Pr}(Qc \mid Qa)$ should go to $\mathsf{Pr}(Qc)$. For $\mathsf{sim}(a, c) \approx 0$ signals the unrelatedness of a and c, rendering Qa irrelevant to the estimation of Qc. Further conditions arise from purely probabilistic considerations. For example, $\mathsf{Pr}(Qa) \approx 1$ should imply $\mathsf{Pr}(Qc \mid Qa) \approx \mathsf{Pr}(Qc)$. (Consider the probability that newborn rats typically weigh at least one ounce assuming that the same is true for newborn elephants.) Conversely, and other things equal, as $\mathsf{Pr}(Qa)$ decreases $\mathsf{Pr}(Qc \mid Qa)$ should increase. The formula must also respect the familiar fact that as $\mathsf{Pr}(Qc)$ goes to unity so does $\mathsf{Pr}(Qc \mid Qa)$, and similarly for zero.

It is easy to verify that Formula (4) of the appendix meets the foregoing conditions. For example, as $\mathsf{sim}(a, c)$ goes to 1, $\frac{1-\mathsf{sim}(a,c)}{1+\mathsf{sim}(a,c)}$ goes to 0, hence $\alpha = \left(\frac{1-\mathsf{sim}(a,c)}{1+\mathsf{sim}(a,c)}\right)^{1-\mathsf{Pr}(Qa)}$ goes to 0, so $\mathsf{Pr}(Qc)^{\alpha}$ goes to 1. Of course, (4) is not unique with these properties, but it is the simplest formula that occurred to us, and reaches its limiting behavior monotonically. It will be seen in the next section that (4) provides a reasonable approximation to $\mathsf{Pr}(Qc \mid Qa)$ in at least one experimental context.

Note that Formula (4) satisfies the Confirmation Assumption (2) inasmuch as $\mathsf{Pr}(Qc)$ cannot exceed $\mathsf{Pr}(Qc \mid Qa)$.

2.2.2 $\mathsf{Pr}(Qc \mid Qa, Qb)$. We propose ten qualitative conditions that should be satisfied by a formula for $\mathsf{Pr}(Qc \mid Qa, Qb)$; five involve similarity and are discussed first. As either $\mathsf{sim}(a, c)$ or $\mathsf{sim}(b, c)$ approach unity, $\mathsf{Pr}(Qc \mid Qa, Qb)$ should also approach unity. (Thus, the conditional probability that pigs have trichromatic vision given that the hogs and squirrels do is close to unity given the similarity of pigs and hogs.) Next, if both $\mathsf{sim}(a, c)$ and $\mathsf{sim}(b, c)$ go to 0 then $\mathsf{Pr}(Qc \mid Qa, Qb)$ should go to $\mathsf{Pr}(Qc)$ (since zero similarity signals irrelevance of the conditioning events). On the other hand, if just $\mathsf{sim}(a, c)$ approaches 0, then $\mathsf{Pr}(Qc \mid Qa, Qb)$ should approach $\mathsf{Pr}(Qc \mid Qb)$; likewise, if just $\mathsf{sim}(b, c)$ approaches 0 then $\mathsf{Pr}(Qc \mid Qa, Qb)$ should approach $\mathsf{Pr}(Qc \mid Qa)$. (Thus, the probability that wolves are fond of garlic given that bears and bees are is close to the probability that wolves are fond of garlic given than bears are.) Next, as $\mathsf{sim}(a, b)$ goes to unity, $\mathsf{Pr}(Qc \mid Qa, Qb)$ should go to $\mathsf{Pr}(Qc \mid Qa)$ [equivalently, $\mathsf{Pr}(Qc \mid Qa, Qb)$ should go to $\mathsf{Pr}(Qc \mid Qb)$]. For $\mathsf{sim}(a, b) \approx 1$ indicates that Qa, Qb record virtually identical facts. (Thus, the probability that otters can hear ultrasounds given that porpoises and dolphins can should be close to the probability that otters can hear ultrasounds given that porpoises can.) Our formula also represents the converse tendency when neither similarities nor absolute probabilities are extreme. In this case, we typically expect $\mathsf{Pr}(Qc \mid Qa, Qb) > \mathsf{Pr}(Qc \mid Qa), \mathsf{Pr}(Qc \mid Qb)$. (Thus, the probability that geese have a magnetic sense given that sparrows and eagles do exceeds the probability that geese have a magnetic sense given that sparrows do, without reference to eagles.) Naturally, there are counterexamples to such generalizations; they will be discussed at the end.

Purely probabilistic conditions on the construction of $\mathsf{Pr}(Qc \mid Qa, Qb)$ include the following.

(a) $0 \leq \mathsf{Pr}(Qc \mid Qa, Qb) \leq 1$.

(b) As $\mathsf{Pr}(Qa)$ approaches unity, $\mathsf{Pr}(Qc \mid Qa, Qb)$ approaches $\mathsf{Pr}(Qc \mid Qb)$. Likewise, as $\mathsf{Pr}(Qb)$ approaches unity, $\mathsf{Pr}(Qc \mid Qa, Qb)$ approaches $\mathsf{Pr}(Qc \mid Qa)$.

(c) As $\mathsf{Pr}(Qa)$ and $\mathsf{Pr}(Qb)$ both go to unity, $\mathsf{Pr}(Qc \mid Qa, Qb)$ goes to $\mathsf{Pr}(Qc)$.

(d) Other things equal, as $\mathsf{Pr}(Qa)$ and $\mathsf{Pr}(Qb)$ both decrease, $\mathsf{Pr}(Qc \mid Qa, Qb)$ increases.

(e) As $\mathsf{Pr}(Qc)$ approaches unity, so does $\mathsf{Pr}(Qc \mid Qa, Qb)$; as $\mathsf{Pr}(Qc)$ approaches zero, so does $\mathsf{Pr}(Qc \mid Qa, Qb)$.

Formula (8) satisfies all the conditions we have posited for $\mathsf{Pr}(Qc \mid Qa, Qb)$. For example, if $\mathsf{sim}(a, b) \approx 1$ then $\beta \approx 1$ in (8), hence $\mathsf{Pr}(Qc \mid Qa) \approx \max\{\mathsf{Pr}(Qc \mid Qa),\ \mathsf{Pr}(Qc \mid Qb)\}$. Since $\mathsf{sim}(a, b) \approx 1$, the latter expression is close to $\max\{\mathsf{Pr}(Qc \mid Qa),\ \mathsf{Pr}(Qc \mid Qa)\} = \mathsf{Pr}(Qc \mid Qa)$.

2.2.3 $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$**.** Note first that (2) implies that $\mathsf{Pr}(Qc \mid Qb) \geq \mathsf{Pr}(Qc)$. It also follows that $\mathsf{Pr}(Qc \mid \neg Qa) \leq \mathsf{Pr}(Qc)$. In light of these inequalities, four constraints on the relation between similarity and $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ may be formulated. First, as $\mathsf{sim}(a, c)$ approaches unity, $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ approaches zero. Likewise, as $\mathsf{sim}(b, c)$ approaches unity, so does $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$. [Of course, by the transitivity of identity and Leibniz's law, it can't be the case that both $\mathsf{sim}(a, c)$ and $\mathsf{sim}(b, c)$ approach unity.] Next, as $\mathsf{sim}(a, c)$ and $\mathsf{sim}(b, c)$ both go to zero, $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ goes to $\mathsf{Pr}(Qc)$. Finally, if just $\mathsf{sim}(a, c)$ goes to zero then $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ goes to $\mathsf{Pr}(Qc \mid Qb)$; likewise, if $\mathsf{sim}(b, c)$ goes to zero then $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ goes to $\mathsf{Pr}(Qc \mid \neg Qa)$.

We also have familiar conditions involving only probability. As $\mathsf{Pr}(Qa)$ approaches zero, $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ approaches $\mathsf{Pr}(Qc \mid Qb)$; likewise, as $\mathsf{Pr}(Qb)$ approaches 1, $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ approaches $\mathsf{Pr}(Qc \mid \neg Qa)$. In the same way, as $\mathsf{Pr}(Qa)$ approaches zero while $\mathsf{Pr}(Qb)$ approaches unity, $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ approaches $\mathsf{Pr}(Qc)$. And as $\mathsf{Pr}(Qc)$ goes to unity (respectively, to zero), so does $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$.

Formula (12) satisfies the foregoing conditions. For example, $\mathsf{sim}(a, c) \approx 1$ implies that X is large relative to Y, hence that $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ is dominated by $\mathsf{Pr}(Qc \mid \neg Qa)$.

$\mathsf{Pr}(Qc \mid Qa, \neg Qb)$ is treated in parallel fashion.

3 Experimental test of the theory

Three experiments were performed to assess the psychological plausibility of our formulas. The vast array of potential predicates and objects renders definitive evaluation a long range project. As a preliminary test, we chose a domain about which college students were likely to have opinions and interest, namely, post-graduation salaries from different colleges and universities.

3.1 Experiment 1

3.1.1 Stimuli and procedure. The following institutions served as objects in the first experiment (playing the roles of a, b, c in Section 2.1).

Connecticut State University Oklahoma State University Harvard University
Arkansas State University Yale University

The predicate employed was:

> over 60% of the graduates from [a given institution] will earn more than $50,000 a year at their first job.

The resulting 5 statements give rise to 20 conditional probabilities of form (1)(a) [$\mathsf{Pr}\,(Qc \mid Qa)$] and to 60 of form (1)(e) [$\mathsf{Pr}\,(Qc \mid Qa, Qb)$].[2] Eighteen students at Northwestern University were recruited to evaluate probabilities and similarities. Each student estimated all 20 conditional probabilities of the first form, and half of the second form (50 judgments in all). In addition, each student estimated the (absolute) probabilities of the five statements, as well as the 10 similarities among the five institutions. Similarity was rated on a scale from 0 (perfect dissimilarity) to 1 (perfect similarity). Observe that 15 judgments of similarity and absolute probability were used to predict 80 judgments of conditional probability (or 50 judgments if the order of conditioning events is ignored). The predictions were made on the basis of formulas (4) and (8) of the appendix.

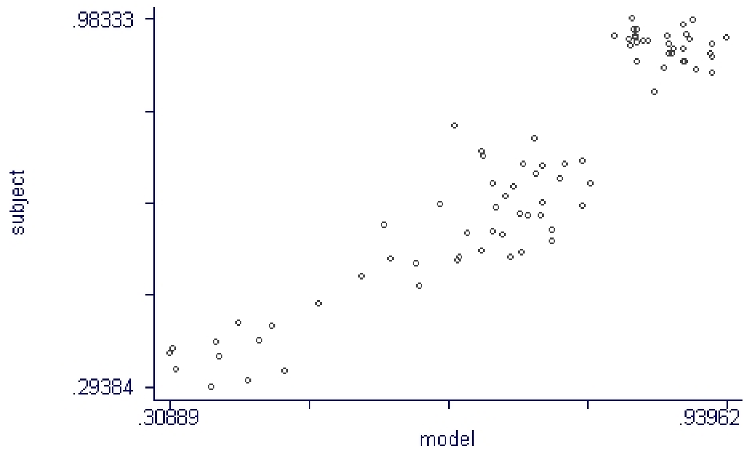
Data were collected using a computerized questionnaire. Similarity judgments were elicited first, followed by absolute probabilities, conditional probabilities of form (1)(a) then conditional probabilities of form (1)(e). Within these categories, stimuli were individually randomized.

3.1.2 Results. Data were averaged prior to analysis. Probabilities of form (1)(e) are thus the average of estimates by 9 students; all other averages involve 18 students.[3] We computed three linear correlations, namely:

(3) (a) between the conditional probabilities of form (1)(a) versus their predicted values using Formula (4);

[2]We distinguish the order of two conditioning events [otherwise, there would be only 30 probabilities of form $\mathsf{Pr}\,(Qc \mid Qa, Qb)$ based on 5 statements]. The order in which information is presented is an important variable in many reasoning contexts [9] although there is little impact in the present study.

[3]Student-by-student analyses of the data are reported in [3].



Results for $\Pr(Qc \mid Qa)$ **and** $\Pr(Qc \mid Qa, Qb)$**.** Correlation = .93 (N = 80). Slope = 1.13, Intercept = .099.

Figure 1: **Correlations in Experiment 1**

(b) between the conditional probabilities of form (1)(e) versus their predicted values using Formula (8); and

(c) between the conditional probabilities of both forms (1)(a) and (1)(e) versus their predicted values using Formulas (4) and (8).

Correlation (3)(c) is plotted in Figure 1, and reveals good fit between predicted and observed values. Note that the regression line does not lie precisely along the diagonal. This disparity seems inevitable given the arbitrary endpoints of the similarity scale, which may not be coterminous with the probability interval in the minds of human judges. The correlation for (3)(a) is .99 ($N = 20$). For (3)(b) it is .92 ($N = 60$).

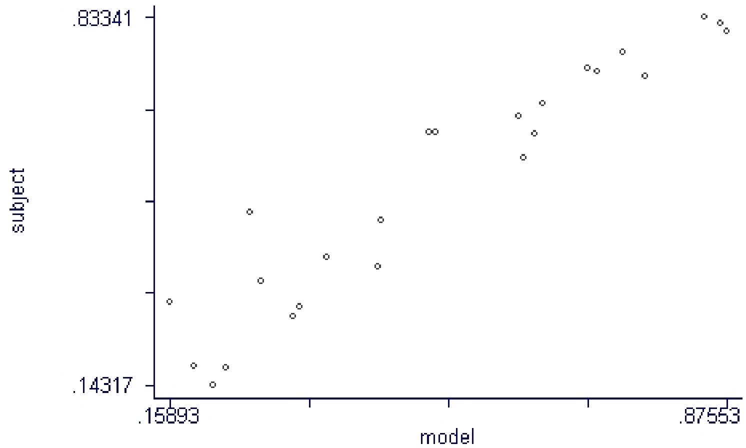
3.2 Experiment 2

3.2.1 Stimuli and procedure. The following four institutions served as objects in the second experiment.

Harvard University	Texas Technical Institute
Harvard Divinity School	Texas Bible College

Two predicates were employed:

- graduates [of a given institution] earned an average salary of *more* than $50,000 a year in their first job after graduation.



Results for $\mathsf{Pr}(Qc \mid Qa)$ and $\mathsf{Pr}(Qc \mid \neg Qa)$. Correlation = .95 ($N$ = 24). Slope = .922, Intercept = .053.

Figure 2: **Correlation in Experiment 2**

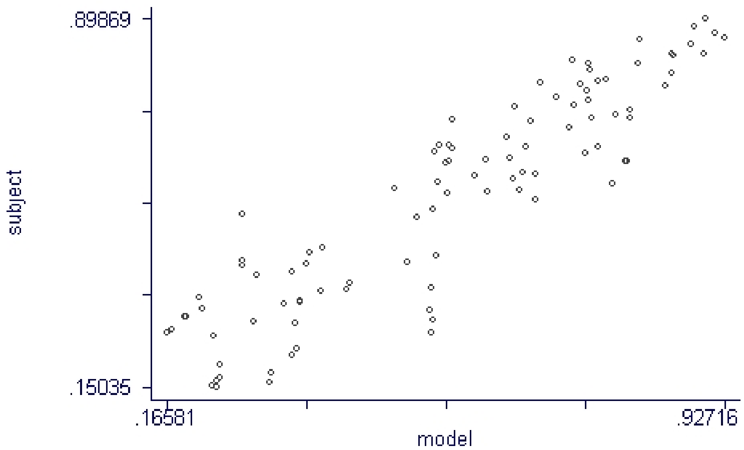
- graduates [of a given institution] earned an average salary of *less* than $50,000 a year in their first job after graduation.

We conceived the second predicate as the negation of the first (ignoring the case of equal salaries). The resulting statements give rise to twelve conditional probabilities of form (1)(a) [$\mathsf{Pr}(Qc \mid Qa)$] and twelve of form (1)(b) [$\mathsf{Pr}(Qc \mid \neg Qa)$]. Forty-one students at Northwestern University evaluated these 24 conditional probabilities along with the similarities and absolute probabilities needed to test formulas (4) and (5). Data collection proceeded as in Experiment 1.

3.2.2 Results. Data were averaged over the 41 students. Figure 2 shows the correlation between the conditional probabilities of forms (1)(a) and (1)(b) versus their predicted values using Formulas (4) and (5). The correlation for the 12 probabilities of form $\mathsf{Pr}(Qc \mid Qa)$ is .98; for $\mathsf{Pr}(Qc \mid \neg Qa)$ it is .91.

3.3 Experiment 3

The stimuli of Experiment 2 also served in Experiment 3. The resulting statements give rise to 96 conditional probabilities of forms (1)(e) [$\mathsf{Pr}(Qc \mid Qa, Qb)$], (1)(f) [$\mathsf{Pr}(Qc \mid \neg Qa, \neg Qb)$]. and (1)(i) [$\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ and $\mathsf{Pr}(Qc \mid Qa, \neg Qb)$]. Forty-seven students at Northwestern University each evaluated half of these conditional probabilities, along with the six similarities and eight absolute probabilities needed to test formulas (8), (9) and (12). Each of the 96



Results for $\mathsf{Pr}(Qc \mid Qa, Qb)$**,** $\mathsf{Pr}(Qc \mid \neg Qa, \neg Qb)$**,** $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ **and** $\mathsf{Pr}(Qc \mid Qa, \neg Qb)$**.** Correlation = .91 (N = 96). Slope = .893, Intercept = .042.

Figure 3: **Correlation in Experiment 3**

conditional probabilities was thus evaluated by 23 or 24 students. Data were collected in the same order as for Experiments 1 and 2. Note that in the present experiment 14 judgments of similarity and absolute probability are used to predict 96 judgments of conditional probability.

Figure 3 shows the correlation between the conditional probabilities of forms (1)(a),(f),(i) versus their predicted values using Formulas (8), (9), and (12). For $\mathsf{Pr}(Qc \mid Qa, Qb)$ the correlation is .93 ($N = 24$). For $\mathsf{Pr}(Qc \mid \neg Qa, \neg Qb)$ it is .92 ($N = 24$), whereas for $\mathsf{Pr}(Qc \mid \neg Qa, Qb)$ and $\mathsf{Pr}(Qc \mid Qa, \neg Qb)$ it is .93 ($N = 48$).

4 Discussion

Our experimental results illustrate the thesis that sensible estimates of chance can be extracted from nonprobabilistic information that may be available to reasoning agents for other purposes. In the present case, conditional probabilities are derived from similarity supplemented with absolute probability. The specific proposals embodied in formulas (4) - (13) should be considered tentative and preliminary.

Of particular importance to developing our theory is the context-dependent nature of similarity, often noted in psychology [10, 13]. Thus, the similarity of the Texas Technical Institute to the Texas Bible College should depend on the content of the statements whose probabilities are at issue (e.g., involving geography versus curriculum). Similarity is usually calculated on the basis of shared features or proximity in a feature space (as noted earlier). The context-dependence of similarity thus requires that some features be counted as more *relevant* than others, depending on which predicate Q is involved in the probability to be estimated. The issue of relevance is much discussed in philosophy and A.I. [4, 7]. For present purposes, it may suffice to measure relevance in terms of path-length in an ontological hierarchy of predicates. Alternatively, two predicates might be considered mutually relevant to the extent that they co-vary in their application to objects familiar to the reasoner. Constructing $\mathsf{Pr}(Qc \mid Qa)$ would thus proceed by first calculating the relevance of predicates to Q, next calculating $\mathsf{sim}(a, c)$ on the basis of weighted features or dimensions, then applying rule (4).

In addition to relevance, there is another type of information that needs to be accessed prior to applying our formulas. To calculate $\mathsf{Pr}(Qc \mid Qa)$, for example, it is necessary to know whether Qa confirms or disconfirms Qc. Suppose that a is Dell Computer Corporation, and c is HP/Compaq. If Q is *increases sales next year* then Qa will strike many reasoners as confirmatory of Qc, whereas if Q is *increases market share next year* then Qa will seem disconfirmatory of Qc. The similarity of a and c can be expected to have different effects in the two cases. To simplify the issue, we constructed stimuli that satisfy the confirmation assumption (2) in Section 2.1, thereby aligning polarity with confirmation; specifically, one of our experimental statements confirms another just in case both have the same number of negation signs. For greater generality, the computation of $\mathsf{Pr}(Qc \mid Qa)$ should proceed on the basis of (4) if Qa confirms Qc, but on the basis of a formula like (5) if Qa disconfirms Qc [and similarly for the other forms listed in (1)]. Default assumptions governing large sets of objects and predicates might suffice in many cases. For example, the susceptibility to a given disease by a given species may be assumed not to disconfirm the susceptibility to that disease by any other species, unless the case is marked otherwise. Marked cases can be expected to include a range of special facts that must be recorded separately. Even the similarity of identical twins can be trumped by information that just one of them is dating a millionaire.

Finally, we note that the theory discussed in the present paper can be extended to a broader class of probabilities, beyond those listed in (1). For example, the probability of universal statements is treated in [3]. The theory can also be generalized to conditional probabilities like $\mathsf{Pr}(Qc \mid Ra)$, involving different predicates in the conditioning and target events. For this purpose, similarity must be assessed between pairs of predicates.

Appendix: Formulas for predicting conditional probability

(4) $$\mathsf{Pr}\,(Qc \mid Qa) = \mathsf{Pr}\,(Qc)^{\alpha},\ \text{where } \alpha = \left(\frac{1-\mathsf{sim}(a,c)}{1+\mathsf{sim}(a,c)}\right)^{1-\mathsf{Pr}(Qa)}.$$

(5) $$\mathsf{Pr}\,(Qc \mid \neg Qa) = 1.0 - (1.0 - \mathsf{Pr}\,(Qc))^{\alpha},\ \text{where } \alpha = \left(\frac{1-\mathsf{sim}(a,c)}{1+\mathsf{sim}(a,c)}\right)^{\mathsf{Pr}(Qa)}.$$

(6) $$\mathsf{Pr}\,(\neg Qc \mid Qa) = 1.0 - \mathsf{Pr}\,(Qc)^{\alpha},\ \text{where } \alpha = \left(\frac{1-\mathsf{sim}(a,c)}{1+\mathsf{sim}(a,c)}\right)^{1-\mathsf{Pr}(Qa)}.$$

(7) $$\mathsf{Pr}\,(\neg Qc \mid \neg Qa) = (1.0 - \mathsf{Pr}\,(Qc))^{\alpha},\ \text{where } \alpha = \left(\frac{1-\mathsf{sim}(a,c)}{1+\mathsf{sim}(a,c)}\right)^{\mathsf{Pr}(Qa)}.$$

(8) $$\mathsf{Pr}\,(Qc \mid Qa, Qb) = \beta M + (1-\beta)S,\ \text{where:}$$

$$\beta = \max\left\{\begin{array}{l} \mathsf{sim}(a,b) \\ \mathsf{sim}(a,c) \\ \mathsf{sim}(b,c) \\ 1.0 - \mathsf{sim}(a,c) \\ 1.0 - \mathsf{sim}(b,c) \\ \mathsf{Pr}\,(Qa) \\ \mathsf{Pr}\,(Qb) \end{array}\right\}, \quad M = \max\{\mathsf{Pr}\,(Qc \mid Qa),\ \mathsf{Pr}\,(Qc \mid Qb)\},$$

$$S = \mathsf{Pr}\,(Qc \mid Qa) + \mathsf{Pr}\,(Qc \mid Qb) - \ \mathsf{Pr}\,(Qc \mid Qa) \times \mathsf{Pr}\,(Qc \mid Qb),$$

and $\mathsf{Pr}\,(Qc \mid Qa)$ and $\mathsf{Pr}\,(Qc \mid Qb)$ are defined by Equation (4).

(9) $$\mathsf{Pr}\,(\neg Qc \mid \neg Qa, \neg Qb) = \beta M + (1-\beta)S,\ \text{where:}$$

$$\beta = \max\left\{\begin{array}{l} \mathsf{sim}(a,b) \\ \mathsf{sim}(a,c) \\ \mathsf{sim}(b,c) \\ 1.0 - \mathsf{sim}(a,c) \\ 1.0 - \mathsf{sim}(b,c) \\ 1.0 - \mathsf{Pr}\,(Qa) \\ 1.0 - \mathsf{Pr}\,(Qb) \end{array}\right\}, \quad M = \max\{\mathsf{Pr}\,(\neg Qc \mid \neg Qa),\ \mathsf{Pr}\,(\neg Qc \mid \neg Qb)\},$$

$$S = \mathsf{Pr}(\neg Qc \mid \neg Qa) + \mathsf{Pr}(\neg Qc \mid \neg Qb) - \mathsf{Pr}(\neg Qc \mid \neg Qa) \times \mathsf{Pr}(\neg Qc \mid \neg Qb),$$

and $\mathsf{Pr}(\neg Qc \mid \neg Qa)$ and $\mathsf{Pr}(\neg Qc \mid \neg Qb)$ are defined by Equation (7).

(10) $\mathsf{Pr}(\neg Qc \mid Qa, Qb) = \beta M + (1 - \beta) S,$ where:

$$\beta = \max \left\{ \begin{array}{l} \mathsf{sim}(a,b) \\ \mathsf{sim}(a,c) \\ \mathsf{sim}(b,c) \\ 1.0 - \mathsf{sim}(a,c) \\ 1.0 - \mathsf{sim}(b,c) \\ \mathsf{Pr}(Qa) \\ \mathsf{Pr}(Qb) \end{array} \right\}, \quad M = \max\{\mathsf{Pr}(\neg Qc \mid Qa),\ \mathsf{Pr}(\neg Qc \mid Qb)\},$$

$$S = \mathsf{Pr}(\neg Qc \mid Qa) + \mathsf{Pr}(\neg Qc \mid Qb) - \mathsf{Pr}(\neg Qc \mid Qa) \times \mathsf{Pr}(\neg Qc \mid Qb),$$

and $\mathsf{Pr}(\neg Qc \mid Qa)$ and $\mathsf{Pr}(\neg Qc \mid Qb)$ are defined by Equation (6).

(11) $\mathsf{Pr}(Qc \mid \neg Qa, \neg Qb) = \beta M + (1 - \beta) S,$ where:

$$\beta = \max \left\{ \begin{array}{l} \mathsf{sim}(a,b) \\ \mathsf{sim}(a,c) \\ \mathsf{sim}(b,c) \\ 1.0 - \mathsf{sim}(a,c) \\ 1.0 - \mathsf{sim}(b,c) \\ 1.0 - \mathsf{Pr}(Qa) \\ 1.0 - \mathsf{Pr}(Qb) \end{array} \right\}, \quad M = \max\{\mathsf{Pr}(Qc \mid \neg Qa),\ \mathsf{Pr}(Qc \mid \neg Qb)\},$$

$$S = \mathsf{Pr}(Qc \mid \neg Qa) + \mathsf{Pr}(Qc \mid \neg Qb) - \mathsf{Pr}(Qc \mid \neg Qa) \times \mathsf{Pr}(Qc \mid \neg Qb), \text{ and}$$

and $\mathsf{Pr}(Qc \mid \neg Qa)$ and $\mathsf{Pr}(Qc \mid \neg Qb)$ are defined by Equation (5).

(12) $\mathsf{Pr}(Qc \mid \neg Qa, Qb) =$

$$\left(\mathsf{Pr}(Qc \mid \neg Qa) \times \frac{X}{X+Y}\right) + \left(\mathsf{Pr}(Qc \mid Qb) \times \frac{Y}{X+Y}\right)$$

where:

$$X = \left(\frac{\mathsf{sim}(a, c)}{1 - \mathsf{sim}(a, c)} \right) \times \mathsf{Pr}\,(Qa)$$

$$Y = \left(\frac{\mathsf{sim}(b, c)}{1 - \mathsf{sim}(b, c)} \right) \times [1.0 - \mathsf{Pr}\,(Qb)], \quad \text{and}$$

$\mathsf{Pr}\,(Qc \mid \neg Qa)$ and $\mathsf{Pr}\,(Qc \mid Qb)$ are defined by Equations (5) and (4), respectively.

(13) $\mathsf{Pr}\,(\neg Qc \mid \neg Qa, Qb) =$

$$\left(\mathsf{Pr}\,(\neg Qc \mid \neg Qa) \times \frac{X}{X + Y} \right) + \left(\mathsf{Pr}\,(\neg Qc \mid Qb) \times \frac{Y}{X + Y} \right)$$

where:

$$X = \left(\frac{\mathsf{sim}(a, c)}{1 - \mathsf{sim}(a, c)} \right) \times \mathsf{Pr}\,(Qa)$$

$$Y = \left(\frac{\mathsf{sim}(b, c)}{1 - \mathsf{sim}(b, c)} \right) \times [1.0 - \mathsf{Pr}\,(Qb)], \quad \text{and}$$

$\mathsf{Pr}\,(\neg Qc \mid \neg Qa)$ and $\mathsf{Pr}\,(\neg Qc \mid Qb)$ are defined by Equations (7) and (6), respectively.

References

[1] C. M. Aguilar and D. L. Medin. Asymmetries of comparison. *Psychonomic Bulletin & Review*, 6:328–337, 1999.

[2] R. Batsell, L. Brenner, D. Osherson, S. Tsavachidis, and M. Y. Vardi. Eliminating incoherence from subjective estimates of chance. In *Proceedings of the 8th International Conference on the Principles of Knowledge Representation and Reasoning (KR 2002)*, 2002.

[3] Sergey Blok. *Probability intuitions derived from judgments of similarity.* PhD thesis, Northwestern University, 2003.

[4] A. L. Blum and P. Langley. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97:245–271, 1997.

[5] J. Deines, D. Osherson, J. Thompson, S. Tsavachidis, and M. Y. Vardi. Removing incoherence from subjective estimates of chance. Submitted, 2002.

[6] P. Gärdenfors. *Conceptual Spaces: The Geometry of Thought.* MIT Press, Cambridge MA, 2000.

[7] Nelson Goodman. Seven strictures on similarity. In *Problems and projects.* Bobbs-Merrill, New York NY, 1972.

[8] D. Heckerman. *Probabilistic Similarity Networks.* MIT Press, Cambridge MA, 1990.

[9] P. N. Johnson-Laird. *Mental Models.* Harvard University Press, Cambridge MA, 1983.

[10] F. Keil, W. C. Smith, D. J. Simons, and D. T. Levin. Two dogmas of conceptual empiricism: implications for hybrid models of the structure of knowledge. *Cognition*, 65:103 – 135, 1998.

[11] D. H. Krantz, R. D. Luce, P. Suppes, and A. Tversky. *Foundations of Measurement, Volume II.* Academic Press, New York NY, 1989.

[12] D. Osherson. New axioms for the contrast model of similarity. *Journal of Mathematical Psychology*, pages 93–103, 1987.

[13] D. Osherson, E. E. Smith, and E. Shafir. Some origins of belief. *Cognition*, 24:197–224, 1986.

[14] D. Osherson, E. E. Smith, O. Wilkie, A. López, and E. Shafir. Category Based Induction. *Psychological Review*, 97(2):185–200, 1990.

[15] J. Pearl. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufamann, San Mateo, CA, 1988.

[16] L. Rips. Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, 14:665–681, 1975.

[17] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Upper Saddle River NJ, 1995.

[18] S. Sattath and A. Tversky. Additive Similarity Trees. *Psychometrika*, 42(3):319–345, 1977.

[19] R. N. Shepard. Multidimensional scaling, tree-fitting, and clustering. *Science*, pages 390–398, 1980.

[20] E. E. Smith. Concepts and Categorization. In E.E. Smith and D. Osherson, editors, *Invitation to Cognitive Science: Thinking (2nd Edition)*. M.I.T. Press, Cambridge MA, 1995.

[21] A. Tversky. Features of Similarity. *Psychological Review*, 84:327–352, 1977.