

Probability from similarity*

Sergey Blok

Northwestern University
s-blok@northwestern.edu

Douglas Medin

Northwestern University
medin@northwestern.edu

Daniel Osherson

Rice University
osherson@rice.edu

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.

Introduction

For the probability calculus to serve as the “faithful guardian of common sense” (Pearl 1988), a great many estimates of chance are needed. Even when the domain can be structured by conditional independence (as assumed in medical diagnosis, Heckerman 1991), 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” (Russell & Norvig 1995). 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 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 (Smith 1995; Sattath & Tversky 1977), 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.¹

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 (Tversky 1977; Osherson 1987) rely on little more than counting the predicates that apply to objects. Likewise, geometrical models (Shepard 1980) involve the positions of objects along quantifiable dimensions like size. (See Krantz, Luce, Suppes & Tversky 1989, and Gärdenfors 2000, 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 $\Pr(p \mid q_1 \cdots q_n)$, 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 salaries after graduating from familiar colleges.

Helping ourselves to the probabilities of $p, q_1 \cdots q_n$ is not (mere) indolence. In Juslin & Persson (2002) the classification model of Medin & Schaffer (1978) is adapted to the construction of subjective probabilities for statements of the form “object o belongs to category C .” Implementing the method is shown to be psychologically plausible from the algorithmic point of view, and the method’s success in predicting elicited probabilities from college students is documented. The statements to be analyzed below likewise involve categorization; hence, our theory can be interpreted as building on Juslin & Persson (2002) by showing how similarity might be used to leverage absolute into conditional probability.

Let us stress at this point that our results are intended merely to illustrate the larger enterprise of culling useful

*Research supported by NSF grants 9978135 and 9983260. Copyright © 2002, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

¹See (Rips 1975; Osherson *et al.* 1990). An attempt to integrate rule-based and similarity-based reasoning within an overarching connectionist system is reported in (Sun 1995). For a broader analysis of commonsense knowledge (that does not assimilate it to probabilities over explicit statements), see (Narasimhan 1997).

probabilities from varied information available to the reasoner. Similarity is only one type of information that ultimately influences human estimates of chance.²

We conclude this introductory section with a remark about *probabilistic coherence*. An agent may endorse $\Pr(p \wedge q \wedge r) = .5$ and $\Pr(q) = .8$ and then use similarity to derive $\Pr(p | q) = .6$. Together, the three judgments are *incoherent* in the sense that no joint distribution generates all three.³ 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 fact, recent progress (Batsell *et al.* 2002; Deines *et al.* 2002) 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.

Theory

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

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) Conditional probabilities to be derived

- (a) $\Pr(Qc | Qa)$
- (b) $\Pr(Qc | \neg Qa)$
- (c) $\Pr(\neg Qc | Qa)$
- (d) $\Pr(\neg Qc | \neg Qa)$
- (e) $\Pr(Qc | Qa, Qb)$
- (f) $\Pr(\neg Qc | \neg Qa, \neg Qb)$
- (g) $\Pr(\neg Qc | Qa, Qb)$
- (h) $\Pr(Qc | \neg Qa, \neg Qb)$
- (i) $\Pr(Qc | \neg Qa, Qb)$ and $\Pr(Qc | Qa, \neg Qb)$
- (j) $\Pr(\neg Qc | \neg Qa, Qb)$ and $\Pr(\neg Qc | Qa, \neg Qb)$

²Obviously, causal theories are also central to commonsense reasoning. See (Ortiz 1999; Turner 1999).

³They imply that $\Pr(p \wedge q) = \Pr(q) \times (\Pr(p \wedge q) / \Pr(q)) = \Pr(q) \times \Pr(p | q) = .8 \times .6 = .48 < .5 = \Pr(p \wedge q \wedge r)$. No distribution assigns lower probability to $p \wedge q$ than to $p \wedge q \wedge r$.

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 $\text{sim}(x, y)$, and assume that it is scaled on the unit interval with 1 representing identity. That is, $\text{sim}(x, y) \in [0, 1]$ and $\text{sim}(x, y) = 1$ iff $x = y$. It is further assumed that sim is a symmetric function, that is, $\text{sim}(x, y) = \text{sim}(y, x)$. The symmetry assumption is consistent with the bulk of human judgment (Aguilar & Medin 1999) even though it may be violated in rare circumstances (Tversky 1977). The symmetry of $\text{sim}(\cdot, \cdot)$ ensures that similarity is not a disguised judgment of conditional probability inasmuch as $\Pr(\cdot | \cdot)$ is not symmetric in its arguments.

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

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

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.

Formulas

The appendix lists the formulas used to construct the conditional probabilities in (1). If the formulas seem complex, the reader is warmly invited to simplify them in a way compatible with the qualitative conditions of "reasonableness" discussed next. In fact, we only discuss cases (1)(a),(e),(i). The others are straightforward variants.

$\Pr(Qc | Qa)$. Any formula for constructing $\Pr(Qc | Qa)$ on the basis of $\Pr(Qa)$, $\Pr(Qc)$ and $\text{sim}(a, c)$ should meet certain qualitative conditions. Trivially, the formula must ensure that $0 \leq \Pr(Qc | Qa) \leq 1$. More substantively, as $\text{sim}(a, c)$ approaches 1, $\Pr(Qc | Qa)$ should approach 1. For if $\text{sim}(a, c) \approx 1$ then $\Pr(Qc | Qa) \approx \Pr(Qc | Qc) = 1$. (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 $\text{sim}(a, c)$ goes to 0, $\Pr(Qc | Qa)$ should go to $\Pr(Qc)$. This is because $\text{sim}(a, c) \approx 0$ signals the unrelatedness of a and c , rendering Qa irrelevant to the estimation of Qc . Further conditions arise from purely probabilistic considerations. For example, $\Pr(Qa) \approx 1$ should imply $\Pr(Qc | Qa) \approx \Pr(Qc)$. (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 $\Pr(Qa)$ decreases $\Pr(Qc | Qa)$ should increase. The formula must also respect the familiar fact that as $\Pr(Qc)$ goes to unity so does $\Pr(Qc | Qa)$, and similarly for zero.

It is easy to verify that Formula (4) of the appendix meets the foregoing conditions. For example, as $\text{sim}(a, c)$ goes to 1,

$$\frac{1 - \text{sim}(a, c)}{1 + \text{sim}(a, c)}$$

goes to 0, hence

$$\alpha = \left(\frac{1 - \text{sim}(a, c)}{1 + \text{sim}(a, c)} \right)^{1 - \text{Pr}(Qa)}$$

goes to 0, so $\text{Pr}(Qc)^\alpha$ 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 $\text{Pr}(Qc | Qa)$ in at least one experimental context.

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

$\text{Pr}(Qc | Qa, Qb)$. We propose ten qualitative conditions that should be satisfied by a formula for $\text{Pr}(Qc | Qa, Qb)$; five involve similarity and are discussed first. As either $\text{sim}(a, c)$ or $\text{sim}(b, c)$ approach unity, $\text{Pr}(Qc | Qa, Qb)$ 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 $\text{sim}(a, c)$ and $\text{sim}(b, c)$ go to 0 then $\text{Pr}(Qc | Qa, Qb)$ should go to $\text{Pr}(Qc)$ (since zero similarity signals irrelevance of the conditioning events). On the other hand, if just $\text{sim}(a, c)$ approaches 0, then $\text{Pr}(Qc | Qa, Qb)$ should approach $\text{Pr}(Qc | Qb)$; likewise, if just $\text{sim}(b, c)$ approaches 0 then $\text{Pr}(Qc | Qa, Qb)$ should approach $\text{Pr}(Qc | Qa)$. (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 $\text{sim}(a, b)$ goes to unity, $\text{Pr}(Qc | Qa, Qb)$ should go to $\text{Pr}(Qc | Qa)$ [equivalently, $\text{Pr}(Qc | Qa, Qb)$ should go to $\text{Pr}(Qc | Qb)$]. For $\text{sim}(a, b) \approx 1$ 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 $\text{Pr}(Qc | Qa, Qb) > \text{Pr}(Qc | Qa), \text{Pr}(Qc | Qb)$. (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 $\text{Pr}(Qc | Qa, Qb)$ include the following.

- (a) $0 \leq \text{Pr}(Qc | Qa, Qb) \leq 1$.
- (b) As $\text{Pr}(Qa)$ approaches unity, $\text{Pr}(Qc | Qa, Qb)$ approaches $\text{Pr}(Qc | Qb)$. Likewise, as $\text{Pr}(Qb)$ approaches unity, $\text{Pr}(Qc | Qa, Qb)$ approaches $\text{Pr}(Qc | Qa)$.
- (c) As $\text{Pr}(Qa)$ and $\text{Pr}(Qb)$ both go to unity, $\text{Pr}(Qc | Qa, Qb)$ goes to $\text{Pr}(Qc)$.
- (d) Other things equal, as $\text{Pr}(Qa)$ and $\text{Pr}(Qb)$ both decrease, $\text{Pr}(Qc | Qa, Qb)$ increases.
- (e) As $\text{Pr}(Qc)$ approaches unity, so does $\text{Pr}(Qc | Qa, Qb)$; as $\text{Pr}(Qc)$ approaches zero, so does $\text{Pr}(Qc | Qa, Qb)$.

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

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

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

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

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

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.

Experiment 1

Stimuli and procedure. The following institutions served as objects in the first experiment (playing the roles of a, b, c in the theoretical section above).

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 five statements give rise to 20 conditional probabilities of form (1)(a) [$\Pr(Qc | Qa)$] and to 60 of form (1)(e) [$\Pr(Qc | Qa, Qb)$].⁴ 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.

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.⁵ We computed three linear correlations, namely:

(3) Correlations computed

- (a) between the conditional probabilities of form (1)(a) versus their predicted values using Formula (4);
- (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 ($N = 60$).

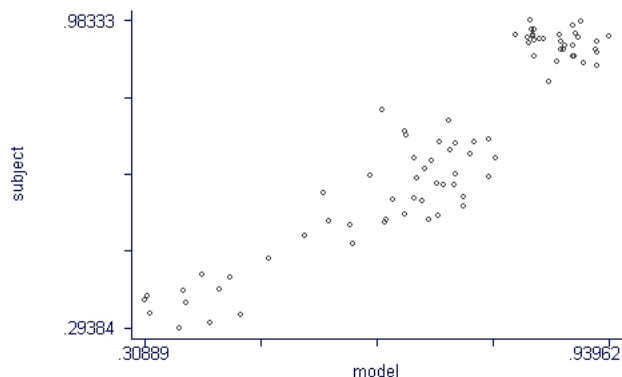
Experiment 2

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

⁴We distinguish the order of two conditioning events [otherwise, there would be only 30 probabilities of form $\Pr(Qc | Qa, Qb)$ based on 5 statements]. The order in which information is presented is an important variable in many reasoning contexts (Johnson-Laird 1983) although there is little impact in the present study.

⁵Student-by-student analyses of the data are reported in Blok (2003).



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

Figure 1: **Correlations in Experiment 1**

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

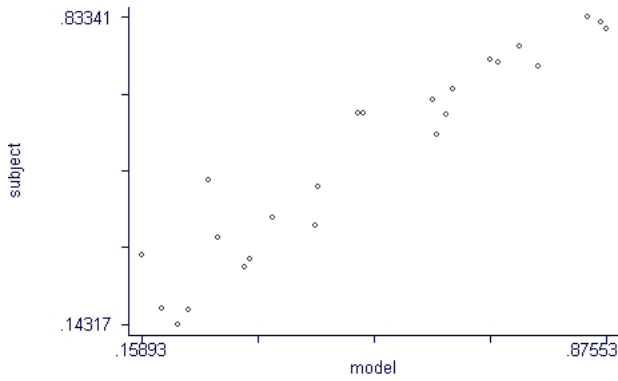
We conceived the second predicate as the negation of the first (ignoring the case of equal salaries). The resulting statements give rise to twelve conditional probabilities of form (1)(a) [$\Pr(Qc | Qa)$] and twelve of form (1)(b) [$\Pr(Qc | \neg Qa)$]. Forty-one students at Northwestern University evaluated 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.

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 $\Pr(Qc | Qa)$ is .98; for $\Pr(Qc | \neg Qa)$ it is .91.

Experiment 3

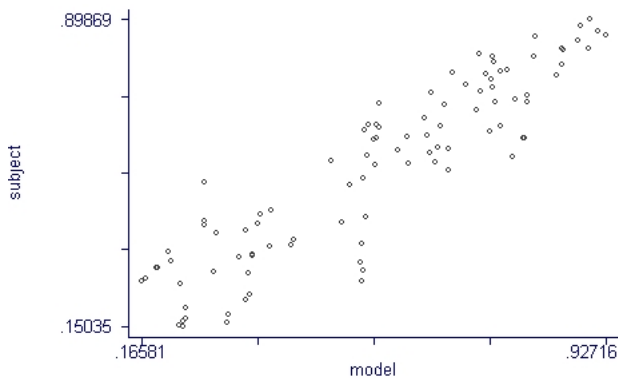
The stimuli of Experiment 2 also served in Experiment 3. The resulting statements give rise to 96 conditional probabilities of forms (1)(e) [$\Pr(Qc | Qa, Qb)$], (1)(h) [$\Pr(Qc | \neg Qa, \neg Qb)$], and (1)(i) [$\Pr(Qc | \neg Qa, Qb)$ and $\Pr(Qc | Qa, \neg Qb)$]. Forty-seven students at Northwestern University each evaluated half of these conditional probabilities, along with the six similarities and eight absolute probabilities needed to test formulas (8), (11) and (12). Each of the 96 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)(e),(h),(i) versus their pre-



Results for $\Pr(Q_c | Q_a)$ and $\Pr(Q_c | \neg Q_a)$. Correlation = .95 ($N = 24$). Slope = .922, Intercept = .053.

Figure 2: Correlation in Experiment 2



Results for $\Pr(Q_c | Q_a, Q_b)$, $\Pr(Q_c | \neg Q_a, \neg Q_b)$, $\Pr(Q_c | \neg Q_a, Q_b)$ and $\Pr(Q_c | Q_a, \neg Q_b)$. Correlation = .91 ($N = 96$). Slope = .893, Intercept = .042.

Figure 3: Correlation in Experiment 3

dicted values using Formulas (8), (11), and (12). For $\Pr(Q_c | Q_a, Q_b)$ the correlation is .93 ($N = 24$). For $\Pr(Q_c | \neg Q_a, \neg Q_b)$ it is .92 ($N = 24$), whereas for $\Pr(Q_c | \neg Q_a, Q_b)$ and $\Pr(Q_c | Q_a, \neg Q_b)$ it is .93 ($N = 48$).

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 (Keil *et al.* 1998; Osherson, Smith, & Shafir 1986; Medin, Goldstone, & Gentner 1993). 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. (Blum & Langley 1997; Goodman 1972). 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 covary in their application to objects familiar to the reasoner. Constructing $\Pr(Q_c | Q_a)$ would thus proceed by first calculating the relevance of predicates to Q , next calculating $\text{sim}(a, c)$ on the basis of weighted features or dimensions, then applying rule (4).⁶

In addition to relevance, there is another type of information that needs to be accessed prior to applying our formulas. To calculate $\Pr(Q_c | Q_a)$, 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), 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 $\Pr(Q_c | Q_a)$ 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, susceptibility to disease may normally be assumed to extend to species of the same genus and rarely beyond (Coley, Atran, & Medin 1997). Exceptions to such rules need to be marked in memory. Marked cases can be expected to include a range of special facts that resist convenient classification. Even the similarity of identical twins (which usually underwrites high probability of matching features) 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 Blok (2003). The theory can also be generalized to conditional probabilities like $\Pr(Q_c | Ra)$, involving different predicates in the conditioning and target events. For this purpose, similarity must be assessed between pairs of predicates.

⁶For further discussion of similarity, relevance, and induction, see (Medin *et al.* 2003). That similarity involves more than just counting features is shown in (Lassaline 1996).

Appendix: Formulas for predicting conditional probability

(4) $\Pr(Qc | Qa) = \Pr(Qc)^\alpha$, where

$$\alpha = \left(\frac{1 - \text{sim}(a, c)}{1 + \text{sim}(a, c)} \right)^{1 - \Pr(Qa)}$$

(5) $\Pr(Qc | \neg Qa) = 1.0 - (1.0 - \Pr(Qc))^\alpha$, where

$$\alpha = \left(\frac{1 - \text{sim}(a, c)}{1 + \text{sim}(a, c)} \right)^{\Pr(Qa)}$$

(6) $\Pr(\neg Qc | Qa) = 1.0 - \Pr(Qc)^\alpha$, where

$$\alpha = \left(\frac{1 - \text{sim}(a, c)}{1 + \text{sim}(a, c)} \right)^{1 - \Pr(Qa)}$$

(7) $\Pr(\neg Qc | \neg Qa) = (1.0 - \Pr(Qc))^\alpha$, where

$$\alpha = \left(\frac{1 - \text{sim}(a, c)}{1 + \text{sim}(a, c)} \right)^{\Pr(Qa)}$$

(8) $\Pr(Qc | Qa, Qb) = \beta M + (1 - \beta)S$, where

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

$M = \max\{\Pr(Qc | Qa), \Pr(Qc | Qb)\}$, $S =$

$\Pr(Qc | Qa) + \Pr(Qc | Qb) - \Pr(Qc | Qa) \times \Pr(Qc | Qb)$,

and $\Pr(Qc | Qa)$ and $\Pr(Qc | Qb)$ are defined by Equation (4).

(9) $\Pr(\neg Qc | \neg Qa, \neg Qb) = \beta M + (1 - \beta)S$, where

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

$M = \max\{\Pr(\neg Qc | \neg Qa), \Pr(\neg Qc | \neg Qb)\}$, $S =$

$\Pr(\neg Qc | \neg Qa) + \Pr(\neg Qc | \neg Qb) -$

$\Pr(\neg Qc | \neg Qa) \times \Pr(\neg Qc | \neg Qb)$,

and $\Pr(\neg Qc | \neg Qa)$ and $\Pr(\neg Qc | \neg Qb)$ are defined by Equation (7).

(10) $\Pr(\neg Qc | Qa, Qb) = \beta M + (1 - \beta)S$, where

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

$M = \max\{\Pr(\neg Qc | Qa), \Pr(\neg Qc | Qb)\}$, $S =$

$\Pr(\neg Qc | Qa) + \Pr(\neg Qc | Qb) -$

$\Pr(\neg Qc | Qa) \times \Pr(\neg Qc | Qb)$,

and $\Pr(\neg Qc | Qa)$ and $\Pr(\neg Qc | Qb)$ are defined by Equation (6).

(11) $\Pr(Qc | \neg Qa, \neg Qb) = \beta M + (1 - \beta)S$, where

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

$M = \max\{\Pr(Qc | \neg Qa), \Pr(Qc | \neg Qb)\}$, $S =$

$\Pr(Qc | \neg Qa) + \Pr(Qc | \neg Qb) -$

$\Pr(Qc | \neg Qa) \times \Pr(Qc | \neg Qb)$,

and $\Pr(Qc | \neg Qa)$ and $\Pr(Qc | \neg Qb)$ are defined by Equation (5).

(12) $\Pr(Qc | \neg Qa, Qb) =$

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

where

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

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

and $\Pr(Qc | \neg Qa)$ and $\Pr(Qc | Qb)$ are defined by Equations (5) and (4), respectively.

(13) $\Pr(\neg Qc | \neg Qa, Qb) =$

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

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

where

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

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

and $\Pr(\neg Qc | \neg Qa)$ and $\Pr(\neg Qc | Qb)$ are defined by Equations (7) and (6), respectively.

References

- Aguilar, C. M., and Medin, D. L. 1999. Asymmetries of comparison. *Psychonomic Bulletin & Review* 6:328–337.
- Batsell, R.; Brenner, L.; Osherson, D.; Tsavachidis, S.; and Vardi, M. Y. 2002. Eliminating incoherence from subjective estimates of chance. In *Proceedings of the 8th International Conference on the Principles of Knowledge Representation and Reasoning (KR 2002)*.
- Blok, S. 2003. *Probability intuitions derived from judgments of similarity*. Ph.D. Dissertation, Northwestern University.
- Blum, A. L., and Langley, P. 1997. Selection of relevant features and examples in machine learning. *Artificial Intelligence* 97:245–271.
- Coley, J. D.; Atran, S.; and Medin, D. L. 1997. Does rank have its privilege? inductive inference within folk biological taxonomies. *Cognition* 64:73–112.
- Deines, J.; Osherson, D.; Thompson, J.; Tsavachidis, S.; and Vardi, M. Y. 2002. Removing incoherence from subjective estimates of chance. Submitted.
- Gärdenfors, P. 2000. *Conceptual Spaces: The Geometry of Thought*. Cambridge MA: MIT Press.
- Goodman, N. 1972. Seven strictures on similarity. In *Problems and projects*. New York NY: Bobbs-Merrill.
- Heckerman, D. 1991. *Probabilistic Similarity Networks*. Cambridge MA: MIT Press.
- Johnson-Laird, P. N. 1983. *Mental Models*. Cambridge MA: Harvard University Press.
- Juslin, P., and Persson, M. 2002. PROBABILITIES from EXemplars (PROBEX): a ‘lazy’ algorithm for probabilistic inference from generic knowledge. *Cognitive Science* 26:563–607.
- Keil, F.; Smith, W. C.; Simons, D. J.; and Levin, D. T. 1998. Two dogmas of conceptual empiricism: implications for hybrid models of the structure of knowledge. *Cognition* 65:103 – 135.
- Krantz, D. H.; Luce, R. D.; Suppes, P.; and Tversky, A. 1989. *Foundations of Measurement, Volume II*. New York NY: Academic Press.
- Lassaline, M. E. 1996. Structural alignment in induction and similarity. *Journal of Experimental Psychology: Learning, Memory and Cognition* 22:754–770.
- Medin, D. L., and Schaffer, M. M. 1978. Context model of classification learning. *Psychological Review* 85:207–238.
- Medin, D. L.; Coley, J. D.; Storms, G.; and Hayes, B. K. 2003. A relevance theory of induction. *Psychonomic Bulletin & Review* (in press).
- Medin, D. L.; Goldstone, R. L.; and Gentner, D. 1993. Respects for similarity. *Psychological Review* 100:254–278.
- Narasimhan, R. 1997. Knowledge processing and commonsense. *Knowledge-Based Systems* 10(3):147–151.
- Ortiz, C. L. 1999. A commonsense language for reasoning about causation and rational action. *Artificial Intelligence* 111:73–130.
- Osherson, D.; Smith, E. E.; Wilkie, O.; López, A.; and Shafir, E. 1990. Category Based Induction. *Psychological Review* 97(2):185–200.
- Osherson, D.; Smith, E. E.; and Shafir, E. 1986. Some origins of belief. *Cognition* 24:197–224.
- Osherson, D. 1987. New axioms for the contrast model of similarity. *Journal of Mathematical Psychology* 93–103.
- Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems*. San Mateo, CA: Morgan Kaufmann.
- Rips, L. 1975. Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior* 14:665–681.
- Russell, S. J., and Norvig, P. 1995. *Artificial Intelligence: A Modern Approach*. Upper Saddle River NJ: Prentice Hall.
- Sattath, S., and Tversky, A. 1977. Additive Similarity Trees. *Psychometrika* 42(3):319–345.
- Shepard, R. N. 1980. Multidimensional scaling, tree-fitting, and clustering. *Science* 390–398.
- Smith, E. E. 1995. Concepts and Categorization. In Smith, E., and Osherson, D., eds., *Invitation to Cognitive Science: Thinking (2nd Edition)*. Cambridge MA: M.I.T. Press.
- Sun, R. 1995. Robust reasoning: integrating rule-based and similarity-based reasoning. *Artificial Intelligence* 75:241–295.
- Turner, H. 1999. A logic of universal causation. *Artificial Intelligence* 113:87–123.
- Tversky, A. 1977. Features of Similarity. *Psychological Review* 84:327–352.