From Similarity to Chance[1]

Sergey Blok
University of Texas at Austin

Daniel Osherson
Princeton University

Douglas L. Medin
Northwestern University

March 22, 2006

[1]Thanks to Robert Rehder for careful reading and helpful suggestions. Research supported by NSF grants 9978135 and 9983260. Contact information: osherson@princeton.edu, blok@psy.utexas.edu, medin@northwestern.edu.

> "In reality, all arguments from experience are founded on the similarity which we discover among natural objects, and by which we are induced to expect effects similar to those which we have found to follow from such objects. ... From causes which appear similar we expect similar effects."
>
> David Hume, *An Enquiry Concerning Human Understanding* (1772)

1 Similarity and inference

1.1 The property problem

Ever since Goodman's (1972) warnings about the "false friend" similarity, psychologists have been cautious about explaining inductive inference in terms of resemblance. Innumerable shared properties unite any pair of objects, inviting their dubious certification as similar along with the inference that some further property is also shared. Genuine resemblance appears to be a three-place predicate, relating two objects only in the context of a set $\mathcal{S}$ of properties (often implicit) that determine the *respects* in which objects may be compared.[1]

When it is plausible that the same set $\mathcal{S}$ is brought to mind in judging similarity and inductive strength, it can be revealing to explain the latter in terms of the former. Thus, famous Linda resembles the average feminist bank teller more than the average bank teller; and whatever properties support this judgment are likely also in place when people judge her chance of being one or the other.[2] Indeed, the probabilities people assign in Linda-like problems are accurately predicted from the relevant similarities rated independently (Shafir, Smith & Osherson, 1990).

A radical means of inclining $\mathcal{S}$ to remain constant between similarity and inference is to limit attention to arguments involving "blank" predicates like *requires biotin for hemoglobin synthesis*. The blankness consists in the difficulty of evaluating the relative credibility of the predicate's application to different objects.[3] The blank-predicate argument

Bears require biotin for hemoglobin synthesis.

Wolves requires biotin for hemoglobin synthesis.

evokes a vaguely physiological context but not much more. Thus, one may have the intuition that cameras and computers do not require biotin but no *a priori* sense of whether bears and wolves do. Nonetheless, the assertion that bears require biotin, coupled with the similarity of wolves to bears, accords strength to the foregoing argument. Theories that incorporate similarity as a basis for induction — such as the Similarity-Coverage model — have been able to account for a range of reasoning phenomena.[4]

Similarity-Coverage is unsuited to non-blank predicates like that figuring in the following arguments.

(1) (a)
$$\frac{\text{Fieldmice often carry the parasite Floxum.}}{\text{Housecats often carry the parasite Floxum.}}$$

(b)
$$\frac{\text{Tigers often carry the parasite Floxum.}}{\text{Housecats often carry the parasite Floxum.}}$$

Many people find (1)a to be stronger than (1)b, no doubt because the relational property *characteristically ingests* readily comes to mind (probably evoked by mention of parasites). In contrast, ingestion is unlikely to be salient when judging the similarity of fieldmice to housecats, preventing similarity from predicting strength. The root of the problem is the difference in causal theories that may occur to the reasoner when diverse sets of properties are evoked in the two settings.[5]

For another limitation of Similarity-Coverage, consider the conclusion that Labradors can bite through wire (of a given thickness and composition) given the alternative premises (a) Collies can, versus (b) Chihuahua's can. More people think that (b) provides better evidence for the conclusion than does (a), despite the fact that Labradors are more similar to Collies than to Chihuahuas. (Smith, Shafir & Osherson, 1993). This intuition likely derives from the belief that Chihuahuas are less powerful than Collies, hence (b) is less likely to be true. The example therefore suggests that a model of induction based solely on similarity will fail when the probability of predicate-application is not uniform across the objects in play.

The aim of the chapter is to show how the prior probabilities of statements might be incorporated into a similarity-based model of induction.[6] We attempt to achieve this by elaborating a new version of the "Gap Model" advanced in Smith et al. (1993). We also suggest how probabilities and similarities can be exploited to construct joint distributions over sets of propositional variables. The first step in our endeavor is to characterize a class of predicates that are neither blank nor epistemically too rich.

1.2 Towards non-blank but manageable predicates

In the space between blank and Floxum-like predicates are those that are meaningful yet mentally evoke the same set $\mathcal{S}$ of properties when people judge similarity compared to argument strength. Let us use Qo to abbreviate the attribution of a predicate Q to an object or category o, and $\neg Qo$ for the denial of this attribution. An argument composed of statements like these will be termed *stable* just in case the same set of properties come to mind when judging the strength of the argument compared to judging the similarity of the objects figuring in it. To test the stability of an argument, the mental potentiation of various predicates would need to be measured during similarity judgment and inference. We do not offer a recipe for such measurement but observe that stability is not circularly defined in terms that guarantee the success of any particular model of inference (like ours, below). Call a predicate "stable" if it gives rise to stable arguments.

For a further distinction, 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 mark the difference, we qualify a predicate Q as *monotonically increasing* [respectively *decreasing*] for objects $O = \{o_1 \cdots o_n\}$ just in case $\mathsf{Prob}\,(Qx : Qy) \geq \mathsf{Prob}\,(Qx)$ [respectively, $\mathsf{Prob}\,(Qx : Qy) \leq \mathsf{Prob}\,(Qx)$] for all $x, y \in O$; and call Q *monotonic* for O if Q is either monotonically increasing or decreasing for O.

Let us illustrate these ideas. The authors expect that for many people the following predicates are monotonically increasing, and yield stable arguments over the class of mammal species.[7]

(2)
has trichromatic vision
can suffer muscle damage through contact with poliomyelitis
brain/body mass ratio is 2 percent or more
requires at least 5 hours of sleep per day for normal functioning
sex drive varies seasonally

We are now in a position to define a circumscribed yet vast class of arguments.

(3) Definition: An argument A is *elementary* just in case:

(a) A has the form $\pm Qa \ / \ \pm Qc$ or $\pm Qa, \pm Qb \ / \ \pm Qc$, where a, b, c occupy the same hierarchical level;

(b) A is stable; and

(c) Q is monotonic for the objects appearing in A.

The notation $\pm Qo$ represents either Qo or $\neg Qo$.

How often do elementary arguments embody a person's reasoning? The authors refrain from exaggerated appraisal. We are uncertain, for example, of the stability of the predicate "can bite through wire," discussed above. Let it merely be noted that few explicit, testable theories of inductive strength for non-trivial classes of arguments are presently on offer. The stable predicates in (2) suggest the richness of the class of elementary arguments. Success in predicting their strength is thus a challenging endeavor, and may point the way to more general models.

1.3 Theoretical goals

The next section advances a theory of the strength of elementary arguments. Given such an argument, let $\mathsf{Prob}\,(\pm Qc : \pm Qa)$ or $\mathsf{Prob}\,(\pm Qc : \pm Qa, \pm Qb)$ be the conditional probability assigned by a given judge to the argument's conclusion given its premises. We attempt to predict these numbers from (a) the absolute probabilities the judge assigns to premises and conclusion individually, and (b) the similarities among the objects appearing in the argument. Letting $\mathsf{sim}(x, y)$ represent the judge's perceived similarity of x to y, our theory takes the form of a function with the following inputs and outputs.

(4)

type of elementary argument	*inputs*	*output*
$\pm Qa \ / \ \pm Qc$	$\mathsf{Prob}\,(\pm Qc)$ $\mathsf{Prob}\,(\pm Qa)$ $\mathsf{sim}(a, c)$	$\mathsf{Prob}\,(\pm Qc : \pm Qa)$
$\pm Qa, \pm Qb \ / \ \pm Qc$	$\mathsf{Prob}\,(\pm Qc)$ $\mathsf{Prob}\,(\pm Qa)$ $\mathsf{Prob}\,(\pm Qb)$ $\mathsf{sim}(a, c)$ $\mathsf{sim}(b, c)$ $\mathsf{sim}(a, b)$	$\mathsf{Prob}\,(\pm Qc : \pm Qa, \pm Qb)$

Similarity judgments are assumed to lie in the unit interval (identity corresponding to 1), and to be symmetric.[8] Beyond this, our model makes no assumptions about how similarity assessments are generated; in particular, complex processes of "feature alignment" may be involved (as in Gentner & Markman, 1997; Goldstone, 1994). Nor do we require that reasoners know whether an argument is elementary, specifically, whether its predicate is stable. Our theory applies when stability holds; other cases lie outside the boundary conditions. The reliance on absolute probabilities as inputs to the theory [$\mathsf{Prob}\,(\pm Qc)$, etc.] underlines the fundamental character of such judgments. For analysis of their provenance, see Juslin & Persson (2002), which adapts the classification model of Medin & Schaffer (1978).

The "Gap" model described by Smith, Shafir & Osherson (1993) was designed for stable predicates, but required strong hypotheses about the featural decomposition of objects and predicates. The present theory is free of such assumptions, and captures the insights of the Gap model in a different way. Given the family history of these ideas, we refer to the present model as "GAP2."

After presentation of the theory, we describe an initial experimental test and its results. A second approach to exploiting similarity for predicting inductive judgments is then advanced.

2 A theory of elementary arguments

Our theory takes the form of equations that determine the outputs of (4) from the inputs. For simplicity in what follows, we consider only predicates that are monotonically increasing. The decreasing case is presumed to be parallel. As a preliminary, we list some qualitative conditions that any candidate theory seems required to satisfy. The conditions express aspects of common sense for a typical reasoner.

2.1 Qualitative requirements: the one-premise case

Let an elementary argument Qa / Qc be given, and consider $\mathsf{Prob}\,(Qc : Qa)$. We expect the latter quantity to approach 1 as $\mathsf{sim}(a, c)$ does. For if $\mathsf{sim}(a, c) \approx 1$ then $\mathsf{Prob}\,(Qc : Qa) \approx \mathsf{Prob}\,(Qc : Qc) = 1$. To illustrate, the conditional probability that pigs have trichromatic vision given that the hogs do is close to 1 given the similarity of these creatures.

On the other hand, $\mathsf{Prob}\,(Qc : Qa)$ should go to $\mathsf{Prob}\,(Qc)$ as $\mathsf{sim}(a, c)$ goes to 0, This is because $\mathsf{sim}(a, c) \approx 0$ signals the unrelatedness of a and c, rendering Qa irrelevant to the estimation of Qc.

Further conditions arise from purely probabilistic considerations. Thus, $\mathsf{Prob}\,(Qa) \approx 1$ should imply $\mathsf{Prob}\,(Qc : Qa) \approx \mathsf{Prob}\,(Qc)$. (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 $\mathsf{Prob}\,(Qa)$ decreases $\mathsf{Prob}\,(Qc : Qa)$ should increase. Our equations should also respect the familiar fact that as $\mathsf{Prob}\,(Qc)$ goes to 1 so does $\mathsf{Prob}\,(Qc : Qa)$, and similarly for zero.

Related requirements apply to arguments of the forms (i) $\neg Qa \;/\; Qc$, (ii) $Qa \;/\; \neg Qc$, and (iii) $\neg Qa \;/\; \neg Qc$. For example, $\mathsf{Prob}\,(Qc : \neg Qa)$ and $\mathsf{Prob}\,(\neg Qc : Qa)$ should both approach 0 as $\mathsf{sim}(a, c)$ approaches 1, whereas $\mathsf{Prob}\,(\neg Qc : \neg Qa)$ should approach 1.

2.2 Qualitative requirements: the two-premise case

Now let a two-premise elementary argument $Qa, Qb \;/\; Qc$ be given. Common sense suggests a large number of constraints on $\mathsf{Prob}\,(Qc : Qa, Qb)$. To begin, as either $\mathsf{sim}(a, c)$ or $\mathsf{sim}(b, c)$ approach 1, $\mathsf{Prob}\,(Qc : Qa, Qb)$ should also approach 1. To illustrate, the conditional probability that pigs have trichromatic vision given that the hogs and squirrels do is close to 1 in view of the similarity of pigs and hogs.

If both $\mathsf{sim}(a, c)$ and $\mathsf{sim}(b, c)$ go to 0 then $\mathsf{Prob}\,(Qc : Qa, Qb)$ should go to $\mathsf{Prob}\,(Qc)$ (since zero similarity signals irrelevance of the conditioning events). On the other hand, if just $\mathsf{sim}(a, c)$ approaches 0, then $\mathsf{Prob}\,(Qc : Qa, Qb)$ should approach $\mathsf{Prob}\,(Qc : Qb)$; likewise, if just $\mathsf{sim}(b, c)$ approaches 0 then $\mathsf{Prob}\,(Qc : Qa, Qb)$ should approach $\mathsf{Prob}\,(Qc : Qa)$. For example, the probability that wolves are fond of garlic given that the same is true of bears and bees is close to the probability that wolves are fond of garlic given than bears are.

Next, as $\mathsf{sim}(a, b)$ goes to 1, $\mathsf{Prob}\,(Qc : Qa, Qb)$ should go to $\mathsf{Prob}\,(Qc : Qa)$ [equivalently, $\mathsf{Prob}\,(Qc : Qa, Qb)$ should go to $\mathsf{Prob}\,(Qc : Qb)$]. For $\mathsf{sim}(a, b) \approx 1$ indicates that Qa, Qb record nearly 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. On the other hand, as $\mathsf{sim}(a, b)$ approaches 0, the strength of $Qa, Qb \;/\; Qc$ should increase (the "diversity effect"). We therefore expect $\mathsf{Prob}\,(Qc : Qa, Qb) > \mathsf{Prob}\,(Qc : Qa), \mathsf{Prob}\,(Qc : Qb)$ (except at the extremes). To illustrate, 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.

Additional conditions on $\mathsf{Prob}\,(Qc : Qa, Qb)$ involve only probability. They include the following.

(a) As $\mathsf{Prob}\,(Qa)$ approaches 1, $\mathsf{Prob}\,(Qc : Qa, Qb)$ approaches $\mathsf{Prob}\,(Qc : Qb)$. Likewise, as $\mathsf{Prob}\,(Qb)$ approaches 1, $\mathsf{Prob}\,(Qc : Qa, Qb)$ approaches $\mathsf{Prob}\,(Qc : Qa)$.

(b) As $\mathsf{Prob}\,(Qa)$ and $\mathsf{Prob}\,(Qb)$ both approach 1, $\mathsf{Prob}\,(Qc : Qa, Qb)$ approaches $\mathsf{Prob}\,(Qc)$.

(c) Other things equal, as $\mathsf{Prob}\,(Qa)$ and $\mathsf{Prob}\,(Qb)$ both decrease, $\mathsf{Prob}\,(Qc : Qa, Qb)$ increases.

(d) As $\mathsf{Prob}\,(Qc)$ approaches 1, so does $\mathsf{Prob}\,(Qc : Qa, Qb)$; as $\mathsf{Prob}\,(Qc)$ approaches zero, so does $\mathsf{Prob}\,(Qc : Qa, Qb)$.

Similar constraints apply to the seven types of two-premise arguments with negated premises or conclusion; their formulation is left to the reader.

2.3 Formulas for one-premise arguments

Consider again the elementary argument $Qa \;/\; Qc$. We propose that $\mathsf{Prob}\,(Qc : Qa)$ is governed by:

(5) $\mathsf{Prob}\,(Qc : Qa) = \mathsf{Prob}\,(Qc)^{\alpha}$, where

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

The reader can verify that (5) satisfies the qualitative conditions reviewed above for $Qa \;/\; Qc$. To illustrate, as $\mathsf{sim}(a,c)$ goes to 1, $\frac{1-\mathsf{sim}(a,c)}{1+\mathsf{sim}(a,c)}$ goes to 0, hence α goes to 0, so $\mathsf{Prob}\,(Qc)^{\alpha}$ goes to 1 [hence, $\mathsf{Prob}\,(Qc : Qa)$ goes to 1]. For another example, as $\mathsf{Prob}\,(Qa)$ goes to 1, α goes to 1, so $\mathsf{Prob}\,(Qc)^{\alpha}$ goes to $\mathsf{Prob}\,(Qc)$ hence $\mathsf{Prob}\,(Qc : Qa)$ goes to $\mathsf{Prob}\,(Qc)$ as desired. Of course, (5) is not unique with these properties, but it is the simplest formula that occurs to us [and restricts $\mathsf{Prob}\,(Qc : Qa)$ to the unit interval]. In the next section, (5) will be seen to provide a reasonable approximation to $\mathsf{Prob}\,(Qc : Qa)$ in an experimental context. It can be seen that our proposal is meant for use with monotonically increasing predicates inasmuch as it guarantees that $\mathsf{Prob}\,(Qc : Qa) \geq \mathsf{Prob}\,(Qc)$.

The following formulas are assumed to govern one-premise elementary arguments with negations. They satisfy commonsense requirements corresponding to those discussed above.

(6) $\mathsf{Prob}\,(Qc : \neg Qa) = 1.0 - (1.0 - \mathsf{Prob}\,(Qc))^{\alpha}$, where

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

(7) $\mathsf{Prob}\,(\neg Qc : Qa) = 1.0 - \mathsf{Prob}\,(Qc)^{\alpha}$, where

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

(8) $\mathsf{Prob}\,(\neg Qc : \neg Qa) = (1.0 - \mathsf{Prob}\,(Qc))^{\alpha}$, where

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

The formulas also guarantee that $\mathsf{Prob}\,(Qc : Qa) + \mathsf{Prob}\,(\neg Qc : Qa) = 1$ and $\mathsf{Prob}\,(Qc : \neg Qa) + \mathsf{Prob}\,(\neg Qc : \neg Qa) = 1$.

We do not assume that people perform the calculations corresponding to Equations (5) - (8). Our claim is merely that the formulas approximate whatever underlying process gives rise to judgments about conditional probability.

2.4 Formulas for two-premise arguments

Two premise arguments likewise require distinguishing multiple cases, all variations on the same theme. We only consider arguments with positive conclusion; negated conclusions are handled similarly (and none of our experiments involve them). We rely on the following concept.

(9) Definition: The *confirmation* exhibited by an argument of form $Qa \;/\; Qc$ is

$$\frac{\mathsf{Prob}\,(Qc : Qa) - \mathsf{Prob}\,(Qc)}{1 - \mathsf{Prob}\,(Qc)}.$$

The *confirmation* exhibited by an argument of form $\neg Qa \ / \ Qc$ is

$$\frac{\mathsf{Prob}\,(Qc) - \mathsf{Prob}\,(Qc : \neg Qa)}{\mathsf{Prob}\,(Qc)}.$$

Given an argument $\pm Qa, \pm Qb \ / \ Qc$, the *dominant premise* is $\pm Qa$ if the confirmation exhibited by $\pm Qa \ / \ Qc$ exceeds the confirmation exhibited by $\pm Qb \ / \ Qc$; otherwise, $\pm Qb$ is the dominant premise.

Confirmation is the impact of an argument's premise on the credibility of its conclusion. The measure suggested above normalizes by the potential impact allowed by the prior credibility of the conclusion (namely, the distance to unity for positive premises and the distance to zero for negative). A variety of confirmation measures are analyzed in Eells & Fitelson (2002). In Tentori, Crupi, Bonini & Osherson (2005), they are compared for their ability to predict shifts of opinion in an experimental setting involving urns. The dominant premise in a two-premise argument is the one that yields the one-premise argument of greatest confirmation. The one-premise probabilities are derived from the theory of one-premise arguments offered above. We now present our theory of two-premise arguments.

(10) Case 1. Arguments of form $Qa, Qb \ / \ Qc$ with Qa dominant.

$$\mathsf{Prob}\,(Qc : Qa, Qb) = \mathsf{Prob}\,(Qc : Qa) +$$

$$[(1 - \mathsf{sim}(a, b)) \times (1 - \mathsf{Prob}\,(Qc : Qa)) \times (\mathsf{Prob}\,(Qc : Qb) - \mathsf{Prob}\,(Qc))].$$

Arguments of form $Qa, Qb \ / \ Qc$ with Qb dominant are treated similarly.

In words, (10) claims that $\mathsf{Prob}\,(Qc : Qa, Qb)$ is given by the probability of the dominant argument increased by a fraction of the probability $1 - \mathsf{Prob}\,(Qc : Qa)$ that the dominant argument "leaves behind." The size of this fraction depends on two factors, namely, the similarity between a and b (to avoid redundancy), and the impact of the nondominant premise on the credibility of Qc.

The constraints outlined earlier are satisfied by (10). For example, the formula implies that $\mathsf{Prob}\,(Qc : Qa, Qb) \approx 1$ if $\mathsf{sim}(a, c) \approx 1$. Note that our proposal ensures that strength increases with extra premises, that is, $\mathsf{Prob}\,(Qc : Qa, Qb) \geq \mathsf{Prob}\,(Qc : Qa), \mathsf{Prob}\,(Qc : Qb)$ This feature is plausible given the restriction to monotonically increasing predicates.

We now list the other two-premise cases. They are predictable from (10) by switching the direction of similarity and "the probability left behind" as appropriate.

(11) Case 2. Arguments of form $\neg Qa, Qb \;/\; Qc$ with $\neg Qa$ dominant.

$$\mathsf{Prob}\,(Qc : \neg Qa, Qb) = \mathsf{Prob}\,(Qc : \neg Qa) +$$

$$[\mathsf{sim}(a,b) \times (\mathsf{Prob}\,(Qc : \neg Qa)) \times (\mathsf{Prob}\,(Qc : Qb) - \mathsf{Prob}\,(Qc))].$$

Arguments of form $Qa, \neg Qb \;/\; Qc$ with $\neg Qb$ dominant are treated similarly.

(12) Case 3. Arguments of form $\neg Qa, Qb \;/\; Qc$ with Qb dominant.

$$\mathsf{Prob}\,(Qc : \neg Qa, Qb) = \mathsf{Prob}\,(Qc : Qb) +$$

$$[\mathsf{sim}(a,b) \times (1 - \mathsf{Prob}\,(Qc : Qb)) \times (\mathsf{Prob}\,(Qc : \neg Qa) - \mathsf{Prob}\,(Qc))].$$

Arguments of form $Qa, \neg Qb \;/\; Qc$ with Qa dominant are treated similarly.

(13) Case 4. Arguments of form $\neg Qa, \neg Qb \;/\; Qc$ with $\neg Qa$ dominant.

$$\mathsf{Prob}\,(Qc : \neg Qa, \neg Qb) = \mathsf{Prob}\,(Qc : \neg Qa) +$$

$$[(1 - \mathsf{sim}(a,b)) \times (\mathsf{Prob}\,(Qc : \neg Qa)) \times (\mathsf{Prob}\,(Qc : \neg Qb) - \mathsf{Prob}\,(Qc))].$$

Arguments of form $\neg Qa, \neg Qb \;/\; Qc$ with $\neg Qb$ dominant are treated similarly.

In summary, Gap2 — our theory of elementary arguments — consists of formulas (5) - (8), and (10) - (13). For a blank predicate Q, we set $\mathsf{Prob}\,(Qo)$ to a fixed constant for any object o (.5 will often be a reasonable choice). Then Gap2 still makes intuitively sound predictions, allowing it to be applied to elementary arguments with blank predicates as a limiting case.

3 Experimental test of the theory

3.1 Stimuli and procedure

To test GAP2 quantitatively we chose a domain about which undergraduates were likely to have opinions and interest, namely, post-graduation salaries from different colleges and universities. The following institutions served as objects.

(14)

(a) Connecticut State University	(b) Oklahoma State University
(c) Arkansas State University	(d) Yale University
(e) Harvard University	(f) Harvard Divinity School
(g) Texas Technical Institute	(h) Texas Bible College

The following predicate was employed.

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

Its negation was taken to be:

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

Considering just objects (14)a - e, there are 20 arguments of form $Qa \ / \ Qc$, and 60 of form $Qa, Qb \ / \ Qc$. Nine undergraduates were recruited to assess the inputs and outputs in (4) for all 20 arguments of form $Qa \ / \ Qc$, and half of the 60 arguments of form $Qa, Qb \ / \ Qc$. Another nine did the same for all 20 arguments of form $Qa \ / \ Qc$, and the other half of the 60 arguments of form $Qa, Qb \ / \ Qc$.[9]

Considering just objects (14)e - h, there are 12 arguments of form $Qa \ / \ Qc$ and 12 of form $\neg Qa \ / \ Qc$. Forty-one additional students evaluated the inputs and outputs for all 24 arguments of these two forms. The same objects give rise to 96 two-premise arguments of form $Qa, Qb \ / \ Qc$ or $\neg Qa, Qb \ / \ Qc$ or $Qa, \neg Qb \ / \ Qc$ or $\neg Qa, \neg Qb \ / \ Qc$. A third group of forty-seven students evaluated inputs and outputs for different halves of these 96 arguments (each argument was evaluated by either 23 or 24 people).

All participants were undergraduates at Northwestern University. Data were collected using a computerized questionnaire. Similarity judgments were elicited first, followed by absolute probabilities, then conditional probabilities. Within these categories, stimuli were individually randomized.

3.2 Results

Overall, the inputs and outputs for 200 different arguments were evaluated, each by 9, 18, 23, 24, or 41 undergraduates. The analysis that follows is based on the mean estimates for each input and output.

We computed the Pearson correlation between the values obtained for $\mathsf{Prob}\,(Qc : \pm Qa)$ or $\mathsf{Prob}\,(Qc : \pm Qa, \pm Qb)$ with the values predicted by GAP2. The result is $r = 0.94$. The regression line has slope $= 0.874$, and intercept $= 0.094$. See Figure 1. To gauge the role of the prior probability $\mathsf{Prob}\,(Qc)$ in GAP2's performance, we computed the correlation between the observed probabilities $\mathsf{Prob}\,(Qc : \pm Qa)$ and $\mathsf{Prob}\,(Qc : \pm Qa, \pm Qb)$ versus the predictions of GAP2 with $\mathsf{Prob}\,(Qc)$ subtracted out. The correlation remains substantial at $r = 0.88$ (slope $= 0.811$, intercept $= 0.079$).

Twenty arguments of form $Qa \;/\; Qc$ were evaluated using objects (14)a - e, and another 12 using (14)e - h. In 31 of these 32 cases, the average probabilities conform to the monotonicity principle $\mathsf{Prob}\,(Qc : Qa) \geq \mathsf{Prob}\,(Qc)$ thereby agreeing with GAP2. The situation is different in the two-premise case. An argument of form $Qa, Qb \;/\; Qc$ exhibits monotonicity if the average responses yield:

(15) $\mathsf{Prob}\,(Qc : Qa, Qb) \geq \max\{\mathsf{Prob}\,(Qc : Qa), \mathsf{Prob}\,(Qc : Qb)\}$.

In our experiment, objects (14)a - e figured in 60 arguments of form $Qa, Qb \;/\; Qc$, with the same participants also evaluating $Qa \;/\; Qc$ and $Qb \;/\; Qc$. In these 60 cases, (15) was violated 51 times ($p < .001$ via a binomial test).

These results suggest modifying our theory by *averaging* the impact of individual premises in a two-premise argument. Let us be cautious, however, about backing off GAP2's monotonicity in this way. Averaging depicts the reasoner as strangely insensitive to the accumulation of evidence, which becomes an implausible hypothesis for larger premise sets. Also, the prevalence of nonmonotonic responding may be sensitive to procedural details. A preliminary study that we have conducted yields scant violation of monotonicity when premises are presented sequentially rather than two at a time.

3.3 Second test of the model

To assess the generality of the foregoing results, we constructed 16 one-premise and 16 two-premise arguments using the following objects and predicate.

Objects:

bears	wolves	cows
sheep	cougars	lions

Predicate: "have at least 18% of their cortex in the frontal lobe"

The sixteen one-premise arguments were equally divided among the forms Qa / Qc, Qa / $\neg Qc$, $\neg Qa$ / Qc, and $\neg Qa$ / $\neg Qc$. The sixteen two-premise arguments were equally divided among Qa, Qb / Qc and $Qa, \neg Qb$ / Qc. Objects (i.e., the six mammals) were assigned to the thirty-two arguments so as balance their frequency of occurrence. Twenty college students at Northwestern University rated the similarity of each of the fifteen pairs of objects drawn from the list above. A separate group of twenty students estimated the conditional probabilities of the 32 arguments plus the probabilities of the six statements Qa. Data were collected via computer interface.

Since the similarity and probability estimates were collected from separate groups of subjects there could be no contamination of one kind of judgment by the other. The data from each group were averaged. The 15 similarity averages plus 6 averages for unconditional probabilities were then used to predict the 32 average conditional probabilities via Gap2. The correlation between predicted and observed values was .89. The results are plotted in Figure 2.

More thorough test of Gap2 (and comparison to rivals) requires experimentation with a range of predicates and objects. The arguments described in here are *homogeneous* in the sense that the same predicate figures in each. We expect Gap2 to apply equally well to *heterogeneous* collections of elementary arguments (involving a multiplicity of predicates).

4 Extensions of Gap2

It is easy to envision extensions of Gap2 beyond elementary arguments [as defined in (3)]. For example, the conditional probability associated with a general-conclusion argument like

Rats have retractable claws.
Squirrels have retractable claws.

All rodents have retractable claws.

might be computed as the minimum of the conditional probabilities for

Rats have retractable claws.
Squirrels have retractable claws.

X have retractable claws.

according to Gap2, where X ranges over the rodent species that come to the reasoner's mind.[10] Another extension is to arguments in which neither the object nor predicate match between premise and conclusion. The goal would be to explain the strength of arguments like the following.

Howler Monkeys will eat cheddar cheese.

Spider Monkeys will eat Swiss cheese.

Both these extensions still require that the argument's predicate(s) be stable, and monotonic for the single-premise case. Escaping this limitation requires understanding how the same predicate can evoke different associations when assessing strength compared to similarity. Specific causal information may often be responsible for such a state of affairs. We acknowledge not (presently) knowing how to incorporate causal judgments into an explicit model of induction.

5 Constructing joint probability distributions using similarity

Theories like Gap2 might help to automate the collection of subjective probabilities for artificial expert systems, especially since it is conditional probabilities that are typically needed (Pearl, 1988, Russell & Norvig, 2003). A more general approach to generating large sets of probability estimates for use in A.I. would construct an entire joint distribution from simple inputs involving similarities and probabilities. The remainder of the chapter advances one method for this purpose.

We limit attention to the following kind of context. Let n propositions $Q(o_1) \cdots Q(o_n)$ be given, where the o_i are n objects at the same hierarchical level, and Q is a stable predicate (in the sense of Section 1.2, above). In the present setting the requirement that predicates be monotonic is no longer needed. Thus, for $n = 6$, the $Q(o_i)$ might correspond to:

(16)
Hawks have muscle-to-fat ratio at least 10-to-1.
Eagles have muscle-to-fat ratio at least 10-to-1.
Parakeets have muscle-to-fat ratio at least 10-to-1.
Cardinals have muscle-to-fat ratio at least 10-to-1.
Geese have muscle-to-fat ratio at least 10-to-1.
Ducks have muscle-to-fat ratio at least 10-to-1.

Closing the six statements under boolean operators yields logically complex statements like:

(17) (a) Eagles and Cardinals have muscle-to-fat ratio at least 10-to-1.[11]
(b) Ducks but not Geese have muscle-to-fat ratio at least 10-to-1.
(c) Either Hawks or Parakeets (or both) have muscle-to-fat ratio at least 10-to-1.

Pairs of statements represent conditionals such as:

(18) (a) Cardinals have muscle-to-fat ratio at least 10-to-1 assuming that Parakeets do.
(b) Eagles have muscle-to-fat ratio at least 10-to-1 assuming that Geese don't.
(c) Ducks and Geese have muscle-to-fat ratio at least 10-to-1 assuming that Hawks do.

Judgments about chance begin to fade in the face of increasing logical complexity, but many people have a rough sense of probability for hundreds of complex and conditional structures like those in (17) and (18). We claim:

(19) Estimates of the chances of complex and conditional events defined over $Q(o_1) \cdots Q(o_n)$ are often mentally generated from no more information than:

(a) probability estimates for each of $Q(o_1) \cdots Q(o_n)$; and

(b) the pairwise similarity of the n objects o_i to each other.

Just as for Gap2, we here make no claim about the probabilities of $Q(o_1) \cdots Q(o_n)$. The focus is rather on sentences involving logical connectives and conditionalization, predicting their perceived chance on the basis of (19)a,b. The next section describes an algorithm whose inputs are the probabilities of $Q(o_1) \cdots Q(o_n)$ along with the pairwise similarities of the o_i, and whose output is the probability of every complex and conditional statement over $Q(o_1) \cdots Q(o_n)$. In support of (19), experimental data will be presented to show that the algorithm's output is a fair approximation of human intuition about statements simple enough to be evaluated.

The algorithm is not intended as a performance model, that is, as an account of the processes whereby people convert the algorithm's inputs into outputs. The algorithm lends credence to (19) only by demonstrating the sufficiency of its inputs as a basis for generating its output. The human method for effecting the same transformation is doubtless different.

Finally, consider a person whose estimates of probability are to be predicted. We use $Ch(\cdot)$ and $Ch(\cdot : \cdot)$ to denote the chances that the person assigns to statements and conditional statements. It is not assumed that *Ch* is a total function since the person may have no opinion about the probability of highly complex statements. As before, we let $\mathsf{sim}(\cdot, \cdot)$ denote the person's similarity function over pairs of objects. It is assumed that $\mathsf{sim}(o_i, o_j) \in [0, 1]$, with 1 representing identity and 0 maximal dissimilarity. For simplicity, we again assume that sim is symmetric in its two arguments.

The goal of our algorithm may now be stated concisely. It is to predict the values of *Ch* from sim along with just the values of *Ch* on $Q(o_1) \cdots Q(o_n)$. We begin by reviewing elements of subjective probability (for more ample discussion, see Jeffrey, 1983; Nilsson, 1986; and Skyrms, 2000, among other sources). Then an overview of the algorithm is presented, followed by details.

Subjective probability

Let $Q(o_1) \cdots Q(o_n)$ be given, where each statement $Q(o_i)$ is called a *variable*. By a *complete conjunction* we mean a sentence of the form $\pm Q(o_1) \wedge \cdots \wedge \pm Q(o_n)$, where each $\pm$ is either blank or the negation symbol $\neg$. For $n = 3$, one complete conjunction is $\neg Q(o_1) \wedge Q(o_2) \wedge$

$\neg Q(o_3)$. A complete conjunction relative to (16) might be expressed in English as:

> Hawks and Eagles and Parakeets but neither Cardinals nor Geese nor Ducks have muscle-to-fat ratio at least 10-to-1.

When there are n variables, there are 2^n complete conjunctions. By a *probability distribution over* $Q(o_1) \cdots Q(o_n)$ is meant a function that maps each complete conjunction to a nonnegative number so that all the numbers sum to one. Such a distribution is extended as follows to the class of boolean formulas constructed from $Q(o_1) \cdots Q(o_n)$. Given a distribution *Prob* and a formula φ, $\mathsf{Prob}\,(\varphi)$ is defined to be

$$\sum\{\mathsf{Prob}\,(\alpha) : \alpha \text{ is a complete conjunction that logically implies } \varphi\}.$$

Prob is extended again to pairs (φ, ψ) of formulas such that $\mathsf{Prob}\,(\psi) > 0$; specifically:

$$\mathsf{Prob}\,(\varphi, \psi) = \frac{\mathsf{Prob}\,(\varphi \wedge \psi)}{\mathsf{Prob}\,(\psi)}.$$

Conforming to the usual convention, we write $\mathsf{Prob}\,(\varphi, \psi)$ as $\mathsf{Prob}\,(\varphi : \psi)$ (the conditional probability of φ assuming ψ). To summarize, any procedure mapping the complete conjunctions into nonnegative numbers that sum to one also assigns probabilities and conditional probabilities to formulas of arbitrary complexity.

The function *Ch* is called *(probabilistically) coherent* if there is some distribution *Prob* such that (a) for all formulas φ, if $Ch(\varphi)$ is defined then $Ch(\varphi) = \mathsf{Prob}\,(\varphi)$, and (b) for all pairs φ, ψ of formulas, if $Ch(\varphi : \psi)$ is defined then $Ch(\varphi : \psi) = \mathsf{Prob}\,(\varphi : \psi)$. It is widely recognized that human estimates of chance are easily led into incoherence.[12] Since our algorithm makes its predictions through construction of a (coherent) distribution, its accuracy is limited by whatever incoherence is manifested in *Ch*.

Overview of the algorithm

The algorithm is based on a function f, defined below, that maps binary conjunctions of form $\pm Q(o_i) \wedge \pm Q(o_j)$ $(i < j)$ into the unit interval. Given a complete conjunction α over n variables, $f(\alpha)$ is defined to be the average of $f(\pm Q(o_i) \wedge \pm Q(o_j))$ for all binary conjunctions $\pm Q(o_i) \wedge \pm Q(o_j)$ whose conjuncts occur in α $(i < j \leq n)$. To illustrate, if $n = 3$, then the complete conjunction $\neg Q(o_1) \wedge Q(o_2) \wedge \neg Q(o_3)$ includes three binary conjunctions, namely, $\neg Q(o_1) \wedge Q(o_2)$, $\neg Q(o_1) \wedge \neg Q(o_3)$, and $Q(o_2) \wedge \neg Q(o_3)$. f maps each of them into $[0, 1]$; and f maps $\neg Q(o_1) \wedge Q(o_2) \wedge \neg Q(o_3)$ into the average of the latter three numbers. Use of average reflects the difficulty of evaluating long conjunctions directly; the binary conjunctions serve as estimates of the compatibility of all the conjuncts in a given complete conjunction.

We conceive of f as assigning "provisional chances" to the set of complete conjunctions generated from $Q(o_1) \cdots Q(o_n)$. Such provisional chances may not define a probability distribution since there is no guarantee that:

$$\sum \{f(\alpha) : \alpha \text{ is a complete conjunction}\} = 1.$$

After the provisional chances are generated via f, the goal of the algorithm is to build a genuine distribution that comes as close as possible to respecting them, subject to the constraints imposed by the rated probability $Ch(Q(o_i))$ of each variable. We therefore solve the following *quadratic programming problem.*

(20) Find a probability distribution *Prob* such that

$$\sum \{(f(\alpha) - \mathsf{Prob}\,(\alpha))^2 : \alpha \text{ is a complete conjunction}\}$$

is minimized subject to the constraints:

(a) $\mathsf{Prob}\,(Q(o_i)) = Ch(Q(o_i))$ for $i \leq n$,

(b) $\mathsf{Prob}\,(\alpha) \geq 0$ for all complete conjunctions α, and

(c) $\sum \{\mathsf{Prob}\,(\alpha) : \alpha \text{ is a complete conjunction}\} = 1$.

The last two constraints embody the requirement that *Prob* be a probability distribution. It can be shown (Luenberger, 1984) that the solution *Prob* to (20) is unique. It is taken to be the output of our algorithm.

We note that (20)a can be supplemented with any other equalities $\mathsf{Prob}\,(\varphi) = Ch(\varphi)$ or $\mathsf{Prob}\,(\varphi : \psi) = Ch(\varphi : \psi)$ where φ, ψ are arbitrary formulas; the only requirement is that the set of equalities appearing in (20)a be coherent. The resulting quadratic program yields a distribution that honors these constraints while minimizing discrepancy with similarity-based estimates of the chances of complete conjunctions. Thus, our method may be applied to situations in which the agent has particular information about the probabilities of various events and conditional events.[13]

The function f

Let $\boldsymbol{c}$ be the conjunction $\pm Q(o_i) \wedge \pm Q(o_j)$ $(i < j \leq n)$. We define $f(\boldsymbol{c})$ via four cases, depending on the polarity of the two conjuncts of $\boldsymbol{c}$.

Case 1: $\boldsymbol{c} = Q(o_i) \wedge Q(o_j)$. As $\mathsf{sim}(o_i, o_j)$ approaches unity, $Q(o_i)$ and $Q(o_j)$ become the same proposition, hence $Ch(\boldsymbol{c})$ should approach $\min\{Ch(Q(o_i)), Ch(Q(o_j))\}$.[14] As $\mathsf{sim}(o_i, o_j)$ approaches zero, $Q(o_i)$ and $Q(o_j)$ bear no relation to each other, hence should be probabilistically independent, so that $Ch(\boldsymbol{c})$ approaches $Ch(Q(o_i)) \times Ch(Q(o_j))$. Linear interpolation between these extremes yields:

$$f(Q(o_i) \wedge Q(o_j)) = [\mathsf{sim}(o_i, o_j) \times \min\{Ch(Q(o_i)), Ch(Q(o_j))\}]$$

$$+ \quad [(1 - \mathsf{sim}(o_i, o_j)) \times Ch(Q(o_i)) \times Ch(Q(o_j))].$$

Case 2: $\boldsymbol{c} = \neg Q(o_i) \wedge \neg Q(o_j)$. Substituting $1 - Ch(Q(o_i))$ for $Ch(Q(o_i))$, and similarly for $Q(o_j)$, transforms this case into the prior one. We therefore define:

$$f(\neg Q(o_i) \wedge \neg Q(o_j)) = [\mathsf{sim}(o_i, o_j) \times \min\{1 - Ch(Q(o_i)), 1 - Ch(Q(o_j))\}]$$

$$+ \quad [(1 - \mathsf{sim}(o_i, o_j)) \times (1 - Ch(Q(o_i))) \times (1 - Ch(Q(o_j)))].$$

Case 3: $\boldsymbol{c} = Q(o_i) \wedge \neg Q(o_j)$. As $\mathsf{sim}(o_i, o_j)$ approaches unity, $Q(o_i)$ and $Q(o_j)$ become the same proposition, hence $\boldsymbol{c}$ becomes a contradiction and $Ch(\boldsymbol{c})$ should approach zero. As $\mathsf{sim}(o_i, o_j)$ approaches zero, $Q(o_i)$ and $\neg Q(o_j)$ bear no relation to each other, hence should be probabilistically independent. Linear interpolation yields:

$$f(Q(o_i) \wedge \neg Q(o_j) = (1 - \mathsf{sim}(o_i, o_j)) \times Ch(Q(o_i)) \times (1 - Ch(Q(o_j))).$$

Case 4: $c = \neg Q(o_i) \wedge Q(o_j)$. This case is parallel to the preceding one. Hence:

$$f(\neg Q(o_i) \wedge Q(o_j) = (1 - \mathsf{sim}(o_i, o_j)) \times (1 - Ch(Q(o_i))) \times Ch(Q(o_j)).$$

The foregoing definition of f seems to be among the simplest ways to assign reasonable probabilities to conjunctions, based on just similarity and the chances attributed to variables.

Let us denote by $\mathsf{QP}f$ the algorithm just described, involving quadratic programming and the use of f to assign provisional chances to the set of complete conjunctions generated from $Q(o_1) \cdots Q(o_n)$.

Test of the algorithm QP*f*

Materials.

As objects o_i we chose the following sets of avian categories.

(21) (a) hawk, eagle, parakeet, cardinal, goose, duck
(b) robin, sparrow, chicken, bluejay, pigeon, parrot

For predicates Q we chose:

(22) (a) have muscle-to-fat ratio at least 10-to-1
(b) have detectable testosterone blood levels throughout the year

Based on no more than intuition, we expected that the predicates would be stable and monotonic for our subjects. Applying each predicate to each group yields four sets of six variables $Q(o_1) \cdots Q(o_6)$.

Similarity was assessed by measuring the distance between each kind of bird in Figure 1 of Rips, Shoben & Smith (1973, p. 10). The latter figure displays the multidimensional scaling solution among the twelve categories in (21),a,b.[15] These distances are shown in Table 1. To compute $\mathsf{sim}(o_i, o_j)$, we divided each distance by the largest value in the associated matrix, and subtracted this number from 1. The most separated pair in a given subset is thus assigned 0 whereas 1 is reserved for self-identity.

Note that the same similarity function is used for all participants in the study. Since similarity questions were not posed, there is no possibility of respondents interpreting similarity queries as disguised requests for conditional probability.

Statements.

Each of the four combinations available from (21) and (22) gives rise to a set of variables $Q(o_1) \cdots Q(o_6)$. Thirty-six complex statements were generated from the variables. The logical forms of the complex statements were the same across the four sets of stimuli; they are shown in Table 2. To illustrate, relative to (21)a and (22)a, the complex statement of form $Q(o_6) \wedge \neg Q(o_5)$ is (17)b, above. For the same set of stimuli, the form $Q(o_2)|\neg Q(o_5)$ is (18)b. In this way, each of the four stimulus-sets was associated with 42 statements (six variables plus 36 complex statements).

Procedure.

Forty Princeton undergraduates assessed the probabilities of the 42 statements associated with one of the four sets of stimuli. It was explained that "probability" was to be interpreted in the sense of personal conviction. Brief explanation was also provided for the logical connectives. Probabilities were collected via a computer interface, which presented the 42 statements in individualized random order.

Three data sets were discarded without further analysis; two manifested stereotyped answers (one or two values chosen for almost all questions), the other included a majority of extreme responses (either zero or one). Of the remaining 37 participants, 10 worked with the stimuli (21)a, (22)a, 10 with (21)a, (22)b, 9 with (21)b, (22)a, and 8 with (21)b, (22)b.

Results

The average estimates for the variables are shown in Table 3. The algorithm (20) was used to derive a separate distribution *Prob* for each participant. We then compared *Prob* to the subject's 36 assignments of chance to complex statements.[16] The median correlations obtained for the four sets of stimuli appear in the first column of Table 4. Next, for a given stimulus set we averaged the probabilities assigned to each of the 42 statements, and repeated the analysis for this "average subject." (There was thus one average subject for each stimulus-set; its 42 "judgments" were the mean responses from the 8, 9, or 10 subjects who worked with those stimuli.) The correlations between predicted and observed values for the four average subjects are shown in the second column of Table 4; scatter plots are provided in Figure 3. When random numbers are substituted for the provisional chances provided by f, the quadratic program (20) yields correlations reliably lower than those shown in Figure 3 ($p < .05$); see the third column of Table 4.

Recall that our algorithm returns a genuine distribution of probability. Its predictive success thus requires a modicum of coherence in the input estimates of chance. Allowing a 10% margin of error, we counted the number of times the inputs violated the following consequences of the probability calculus.[17]

(23) (a) $\mathsf{Prob}\,(\varphi) + \mathsf{Prob}\,(\psi) - 1.05 \leq \mathsf{Prob}\,(\varphi \wedge \psi) \leq \mathsf{Prob}\,(\varphi) + .05, \mathsf{Prob}\,(\psi) + .05$

(b) $\mathsf{Prob}\,(\varphi) - .05, \mathsf{Prob}\,(\psi) - .05 \leq \mathsf{Prob}\,(\varphi \vee \psi) \leq \mathsf{Prob}\,(\varphi) + \mathsf{Prob}\,(\psi) + .05$

(c) $\mathsf{Prob}\,(\varphi \wedge \psi)/\mathsf{Prob}\,(\psi) - .05 \leq \mathsf{Prob}\,(\varphi : \psi) \leq \mathsf{Prob}\,(\varphi \wedge \psi)/\mathsf{Prob}\,(\psi) + .05$

All 24 complex, absolute statements (see Table 2) are either conjunctions or disjunctions, hence accountable to one of (23)a,b. Only four of the 12 conditional statements could be tested against (23)c since in 8 cases the required conjunction $\varphi \wedge \psi$ does not figure among the absolute statements. For the averaged data, there are few violations of (23)ab — 6, 0, 1, and 1, respectively, for the four stimulus sets. In contrast, (23)c was violated all four times in each stimulus set (averaged data), confirming the well documented incomprehension of conditional probability by college undergraduates (see, e.g., Dawes, Mirels, Gold & Donahue, 1993). Incoherence was greater when tabulated for each subject individually. Participants averaged 10.3 violations of (23)a,b out of 24 possible, and 3.8 violations of (23)c out of 4 possible.

Our results support Thesis (19) inasmuch as the only inputs to the algorithm are similarities and the chances of variables. The algorithm's predictions are not perfect, however, and suggest that revised methods might prove more successful. In particular, accuracy may improve when similarities are elicited from respondents to the probability questions, instead of derived globally from an ancient text (Rips et al., 1973).

The construction of a (coherent) probability distribution over the set of complete conjunctions renders our algorithm unrealistic psychologically, but enhances its potential role in the design of autonomous, intelligent agents. Indeed, the probability calculus provides an attractive canonical form for reasoning, and is often considered to be the "faithful guardian of common sense" (Pearl, 1988).

Application to the design of autonomous intelligent agents

Our algorithm allows numerous probabilities to be computed from a small reservoir of stored values. For example, a distribution for thirty variables of form $Q(o_i)$ requires more than a billion probabilities but is specified by only 465 numbers within the scheme described here (namely, 30 probabilities for variables plus 435 pairwise similarities). The empirical results described above suggest that the 465 numbers can be chosen in a way that generates reasonable — that is human-like — estimates of chance. Moreover, the bulk of these numbers (namely, the similarity coefficients), will be of potential use for other purposes, e.g., when the agent must decide whether one object can be substituted for another in pursuit of a larger goal.

Although 465 numbers suffice to define a distribution involving 30 variables, recovering the distribution via our algorithm requires quadratic programming with a billion-dimensional matrix [see (20), above]. To avoid such computations, a small random sample of complete conjunctions can be chosen to carry all positive probabilities, and quadratic programming undertaken just with them. At least one sampled complete conjunction must imply each variable in order for the variables' estimated chances to be preserved in the solution; but this condition on samples is easy to check.

To assess the potential of the foregoing scheme, for each set of stimuli we drew six samples of 32 complete conjunctions, out of the 64 that are generated by 6 variables.[18] The resulting quadratic program — which is one quarter the size of the original — was then used to predict the subjects' average estimates of chance. The last column in Table 4 shows the mean of the six correlations produced for each set of stimuli. Comparing these values to the second column reveals some loss of predictive accuracy; substantial correlations nonetheless remain. In a sophisticated implementation of our algorithm, loss of accuracy could be partially remedied by applying quadratic programming to several samples of complete conjunctions and averaging the results.

6 Concluding remarks

Both Gap2 and QPf rely on similarity to predict the probabilities assigned to complex statements. We hope that the performance of the models demonstrates that similarity can play an explanatory role in theories of induction provided that attention is limited to stable predicates. Extension beyond the confines of stability is not trivial but might be facilitated by comparison to the simpler case.

The importance of similarity to estimates of probability was already articulated in the 17th and 18th century (see Cohen, 1980; Coleman, 2001). Bringing quantitative precision to this idea would be a significant achievement of modern Psychology.

Notes

[1]See also Goodman (1955), Quine (1960). In the psychology literature, the matter has been rehearsed by Tversky & Gati (1978), Osherson, Smith & Shafir (1986), Medin, Goldstone & Gentner (1993), and Keil, Smith, Simons & Levin (1998), among others.

[2]Linda was born in Tversky & Kahneman (1983). For recent debate about the fallacy she illustrates, see Kahneman & Tversky (1996), Gigerenzer (1996), and Tentori, Bonini & Osherson (2004), along with references cited there.

[3]A more cautious formulation would relativize blankness to a particular judge, and also to the objects in play. Degrees of blankness would also be acknowledged, in place of the absolute concept here invoked. In what follows, we'll rely on rough formulations (like the present one) to convey principal ideas.

[4]See Osherson, Smith, Wilkie, López & E. Shafir (1990). An alternative account is offered in Sloman (1993). Blank predicates first appear in Rips (1975). For examination of Similarity-Coverage type phenomena in the context of natural category hierarchies, see Coley, Atran and Medin (1997) Medin, Coley, Storms and Hayes (2003). Developmental perspectives on Similarity-Coverage are available in López, Gelman, Gutheil & Smith (1992), Heit & Hahn (1999), and Lo, Sides, Rozelle & Osherson (2002). Similarity among biological categories may be more resistant to contextual shifts than is the case for other categories (Barsalou & Ross, 1986). This would explain their popularity in many discussions of inductive inference.

[5]The Floxum example is based on López, Atran, Coley, Medin and Smith (1997). It is exploited in Lo et al. (2002) to argue that the "diversity principle" does not have the normative status often assumed by psychologists. For the role of causal theories in commonsense reasoning, see Ahn, Kalish, Medin & Gelman (1995) and Lassaline (1996), along with the theoretical discussions in Ortiz (1993) and Turner (1993).

[6]Reservations about the role of similarity in theories of reasoning are advanced in Sloman & Rips (1998). Overviews of theories of inductive reasoning are available in Heit (2000), Sloman & Lagnado (2005).

[7]For another example, susceptibility to disease is usually perceived to be shared for species of the same genus (Coley, Atran & Medin, 1997).

[8]Claims for asymmetry appear in Tversky (1977), but seem to arise only in unusual circumstances; see Aguilar and Medin (1999). The assumption $\mathsf{sim}(o_i, o_j) \in [0, 1]$ is substantive; in particular, dissimilarity might be unbounded. This possibility is ignored in what follows. For a review of ideas concerning asymmetries in inductive reasoning, see the chapter by Medin & Waxman in the present volume.

[9]The predicate was slightly (but inessentially) different for these 18 participants. Note that we distinguish the order of two conditioning events; otherwise, there would be only 30 probabilities of form $\mathsf{Prob}\,(Qc : Qa, Qb)$ based on 5 objects. No student received two arguments differing just on premise order. The order in which information is presented is an important variable in many reasoning contexts (Johnson-Laird, 1983) but there was little impact in the present study.

[10]A "coverage" variable could also be added to the theory, in the style of the Similarity-Coverage Model. Coverage for specific arguments is motivated by asymmetry phenomena such as the greater strength of 'Qcats / Qbats' compared to 'Qbats / Qcats' (see Carey, 1985; Osherson et al., 1990). GAP2 does not predict this phenomenon for any predicate Q such that $\mathsf{Prob}\,(Q\text{cats}) = \mathsf{Prob}\,(Q\text{bats})$ (e.g., when Q is blank). Asymmetry is known to be a weak effect, however (see Hampton & Cannon, 2004). If it is robust enough to merit modeling, this could also be achieved by positing asymmetry in sim.

[11]A more literal rendition of closure under boolean conjunction would be "Eagles have muscle-to-fat ratio at least 10-to-1 and Cardinals have muscle-to-fat ratio at least 10-to-1." We assume throughout that the official logical structure of statements is clear from the abbreviations that make them more colloquial. Note also that negation of predicates may often be expressed without "not." Thus, "have muscle-to-fat ratio below 10-to-1" serves as the negation of "have muscle-to-fat ratio at least 10-to-1."

[12]See the references in Note 2. Incoherence can be "fixed" via a method described in Batsell, Brenner, Osherson, Vardi & Tsvachidis (2002).

[13]More generally, the constraints may be weak inequalities rather than equalities.

[14]As $\mathsf{sim}(o_i, o_j)$ approaches unity, $Ch(\boldsymbol{c})$ should also approach $\max\{Ch(Q(o_i)), Ch(Q(o_j))\}$. Prior to reaching the limit, however, the maximum might exceed $Ch(\boldsymbol{c})$, which is probabilistically incoherent. No such incoherence is introduced by the minimum.

[15]The raw data for the multidimensional scaling algorithm were rated semantic similarities for each pair of categories. Distance in the output reflects dissimilarity. (Included in the ratings was semantic similarity to the superordinate concept *bird*.)

[16]*Prob* perfectly "predicts" the probabilities assigned to $Q(o_1) \cdots Q(o_6)$ since the quadratic programming problem (20) takes the latter values as constraints on the solution.

[17]For their derivation, see Neapolitan (1990).

[18]A few times it was necessary to draw a new sample because one of the variables was not implied by any of the chosen complete conjunctions.

References

Aguilar, C. M., and D. L. Medin (1999): "Asymmetries of comparison," *Psychonomic Bulletin & Review*, 6, 328–337.

Ahn, W., C. W. Kalish, D. L. Medin, and S. A. Gelman (1995): "The role of covariation versus mechanism information in causal attribution," *Cognition*, 54, 299–352.

Barsalou, L. W., and B. H. Ross (1986): "The roles of automatic and strategic processing in sensitivity to superordinate and property frequency," *Journal of Experimental Psychology: Learning, Memory, & Cognition*, I, 116–134.

Batsell, R., L. Brenner, D. Osherson, M. Vardi, and S. Tsvachidis (2002): "Eliminating Incoherence from Subjective Estimates of Chance," *Proceedings of the Ninth International Workshop on Knowledge Representation.*

Blok, S. V. (2004): "Modeling induction as conditional probability judgment," *Dissertation Abstracts International*, 65(1), 457, Northwestern University.

Carey, S. (1985): *Conceptual change in childhood.* MIT Press, Cambridge MA.

Cohen, L. J. (1980): "Some Historical Remarks on the Baconian Conception of Probability," *Journal of the History of Ideas*, 61, 219–31.

Coleman, D. (2001): "Baconian Probability and Hume's Theory of Testimony," *Hume Studies*, 27(2), 195–226.

Coley, J. D., S. Atran, and D. L. Medin (1997): "Does rank have its privilege? Inductive inference within folk biological taxonomies," *Cognition*, 64, 73–112.

Dawes, R., H. L. Mirels, E. Gold, and E. Donahue (1993): "Equating inverse probabilities in implicit personality judgments," *Psychological Science*, 4(6), 396–400.

Eells, E., and B. Fitelson (2002): "Symmetries and asymmetries in evidential support," *Philosophical Studies*, 107.

Gentner, D., and A. Markman (1997): "Structure Mapping in Analogy and Similarity," *American Psychologist*, (1).

Gigerenzer, G. (1996): "Reply to Tversky and Kahneman," *Psychological Review*, 103(3), 592–3.

Goldstone, R. (1994a): "Similarity, Interactive Activation, and Mapping," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(1), 3–28.

Goldstone, R. L. (1994b): "The role of similarity in categorization: providing a groundwork," *Cognition*, 52, 125–157.

Goodman, N. (1955): *Fact, Fiction, and Forecast*. Bobbs-Merrill, Indianapolis IN.

Goodman, N. (1972): "Seven Strictures on Similarity," in *Problems and projects*. Bobbs-Merrill, New York NY.

Hampton, J., and I. Cannon (2004): "Category-based induction: An effect of conclusion typicality," *Memory & Cognition*, 32(2), 235–243.

Heit, E. (2000): "Properties of inductive reasoning," *Psychonomic Bulletin & Review*, 7, 569–592.

Heit, E., and U. Hahn (1999): "Diversity-Based Reasoning in Children Age 5 to 8," in *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society*, Mahawah NJ. Erlbaum.

Jeffrey, R. C. (1983): *The Logic of Decision (2nd Edition)*. The University of Chicago Press, Chicago IL.

Johnson-Laird, P. N. (1983): *Mental Models*. Harvard University Press, Cambridge MA.

Juslin, P., and M. Persson (2002): "PROBabilities from EXemplars (PROBEX): a 'lazy' algorithm for probabilistic inference from generic knowledge," *Cognitive Science*, 26, 563–607.

Kahneman, D., and A. Tversky (1996): "On the Reality of Cognitive Illusions," *Psychological Review*, 103(3), 582–591.

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

Lassaline, M. E. (1996): "Structural alignment in induction and similarity," *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 22, 754 – 770.

Lo, Y., A. Sides, J. Rozelle, and D. Osherson (2002): ""Evidential diversity and premise probability in young children's inductive judgment," *Cognitive Science*, 26, 181–206.

López, A., S. Atran, J. Coley, D. Medin, and E. Smith (1997): "The tree of life: Universal and cultural features of folkbiological taxonomies and inductions," *Cognitive Psychology*, 32, 251–295.

López, A., S. A. Gelman, G. Gutheil, and E. E. Smith (1992): "The Development of Category-based Induction," *Child Development*, 63, 1070–1090.

Luenberger, D. G. (1984): *Linear and Nonlinear Programming (2nd Edition)*. Addison-Wesley, Reading MA.

Medin, D. L., J. D. Coley, G. Storms, and B. K. Hayes (2003): "A relevance theory of induction," *Psychonomic Bulletin & Review*.

Medin, D. L., R. L. Goldstone, and D. Gentner (1993): "Respects for similarity," *Psychological Review*, 100, 254–278.

Medin, D. L., and M. M. Schaffer (1978): "Context model of classification learning," *Psychological Review*, 85, 207–238.

Neapolitan, R. (1990): *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*. John Wiley & Sons, New York NY.

Nilsson, N. (1986): "Probabilistic logic," *Artificial Intelligence*, 28(1), 71–87.

Ortiz, C. L. (1999): "A commonsense language for reasoning about causation and rational action," *Artificial Intelligence*, 111, 73–130.

Osherson, D., E. E. Smith, and E. Shafir (1986): "Some origins of belief," *Cognition*, 24, 197–224.

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

Pearl, J. (1988): *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufamann, San Mateo, CA.

Quine, W. V. O. (1960): *Word and Object*. MIT Press, Cambridge MA.

Rips, L. (1975): "Inductive judgments about natural categories," *Journal of Verbal Learning and Verbal Behavior*, 14, 665–681.

Rips, L. J., E. J. Shoben, and E. E. Smith (1973): "Semantic distance and the verification of semantic relations," *Journal of Verbal Learning and Verbal Behavior*, 12, 1–20.

Russell, S. J., and P. Norvig (2003): *Artificial Intelligence: A Modern Approach*. Prentice Hall, Upper Saddle River NJ, second edn.

Shafir, E., E. Smith, and D. Osherson (1990): "Typicality and reasoning fallacies," *Memory and Cognition*, 18(3), 229–239.

Skyrms, B. (2000): *Choice & Chance: An Introduction to Inductive Logic*. Wadsworth, Belmont CA.

Sloman, S. A. (1993): "Feature based induction," *Cognitive Psychology*, 25, 231–280.

Sloman, S. A., and D. Lagnado (2005): "The problem of induction," in *Cambridge Handbook of Thinking & Reasoning*, ed. by R. Morrison, and K. Holyoak.

Sloman, S. A., and L. J. Rips (1998): "Similarity as an explanatory construct," *Cognition*, 65, 87–101.

SMITH, E. E., E. SHAFIR, AND D. OSHERSON (1993): "Similarity, plausibility, and judgments of probability," *Cognition*, 49, 67–96.

TENTORI, K., N. BONINI, AND D. OSHERSON (2004): "The Conjunction Fallacy: A Misunderstanding about Conjunction?," *Cognitive Science*, 28(3), 467–477.

TENTORI, K., V. CRUPI, N. BONINI, AND D. OSHERSON (2005): "Comparison of Confirmation Measures," manuscript.

TURNER, H. (1999): "A logic of universal causation," *Artificial Intelligence*, 113, 87–123.

TVERSKY, A. (1977): "Features of Similarity," *Psychological Review*, 84, 327–352.

TVERSKY, A., AND I. GATI (1978): "Studies of similarity," in *Cognition and Categorization*, ed. by E. Rosch, and B. Lloyd. Erlbaum, Hillsdale NJ.

TVERSKY, A., AND D. KAHNEMAN (1983): "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment," *Psychological Review*, 90, 293–315.

TABLE 1

Distances (in millimeters) in Figure 1 of Rips, Shoben & Smith (1973)

	robin	sparrow	chicken	bluejay	pigeon
sparrow	4.0				
chicken	39.0	38.5			
bluejay	10.3	6.4	38.1		
pigeon	17.5	17.4	21.5	18.6	
parrot	15.5	16.0	23.5	17.9	2.3

	hawk	eagle	parakeet	cardinal	goose
eagle	4.8				
parakeet	41.0	43.0			
cardinal	36.2	36.9	13.4		
goose	40.0	44.8	34.1	42.6	
duck	40.5	44.9	31.5	40.5	3.2

Note. To illustrate, $\mathsf{sim}(\text{robin,parrot}) = 1.0 - (15.5/39.0)$.

Table 2

The 36 complex statements figuring in the experiment

Structure	#	Formulas
$p \wedge q$	7	$Q(o_1) \wedge Q(o_2)$ $Q(o_5) \wedge Q(o_6)$ $Q(o_1) \wedge Q(o_5)$ $Q(o_1) \wedge Q(o_6)$ $Q(o_2) \wedge Q(o_5)$ $Q(o_2) \wedge Q(o_6)$ $Q(o_3) \wedge Q(o_4)$
$p \wedge \neg q$	7	$Q(o_1) \wedge \neg Q(o_3)$ $Q(o_1) \wedge \neg Q(o_4)$ $Q(o_2) \wedge \neg Q(o_5)$ $Q(o_2) \wedge \neg Q(o_6)$ $Q(o_6) \wedge \neg Q(o_1)$ $Q(o_1) \wedge \neg Q(o_2)$ $Q(o_6) \wedge \neg Q(o_5)$
$\neg p \wedge \neg q$	7	$\neg Q(o_4) \wedge \neg Q(o_3)$ $\neg Q(o_6) \wedge \neg Q(o_3)$ $\neg Q(o_6) \wedge \neg Q(o_5)$ $\neg Q(o_2) \wedge \neg Q(o_1)$ $\neg Q(o_6) \wedge \neg Q(o_1)$ $\neg Q(o_2) \wedge \neg Q(o_5)$ $\neg Q(o_3) \wedge \neg Q(o_5)$
$p \vee q$	3	$Q(o_2) \vee Q(o_3)$ $Q(o_4) \vee Q(o_5)$ $Q(o_6) \vee Q(o_4)$
$p : q$	5	$Q(o_1) : Q(o_2)$ $Q(o_2) : Q(o_3)$ $Q(o_3) : Q(o_2)$ $Q(o_4) : Q(o_1)$ $Q(o_5) : Q(o_2)$
$p : \neg q$	4	$Q(o_1) : \neg Q(o_5)$ $Q(o_2) : \neg Q(o_6)$ $Q(o_3) : \neg Q(o_4)$ $Q(o_6) : \neg Q(o_5)$
$\neg p : q$	3	$\neg Q(o_4) : Q(o_3)$ $\neg Q(o_5) : Q(o_1)$ $\neg Q(o_6) : Q(o_3)$

Note. Indices are relative to the orderings in (21)a,b. The symbols $\wedge$ and $\vee$ denote conjunction and disjunction, respectively; $\neg$ denotes negation. Conditional probabilities are indicated with $|$.

Table 3

Average probabilities assigned to variables

#	(21)a,(22)a ($N = 10$)	(21)a,(22)b ($N = 10$)	(21)b,(22)a ($N = 9$)	(21)b,(22)b ($N = 8$)
1	.79	.75	.54	.35
2	.79	.74	.61	.43
3	.39	.36	.30	.53
4	.45	.42	.61	.40
5	.28	.59	.40	.44
6	.29	.47	.48	.36

Note. The numbers under # are relative to the orderings in (21)a,b.

Table 4

Correlations between predicted and observed probabilities

	(1) Median individual correlation	(2) Correlation for the average subject	(3) Correlations with random provisional chances	(4) Correlations based on 32 provisional chances
(21)a,(22)a	.811	.915	.757	.795
(21)a,(22)b	.763	.899	.539	.872
(21)b,(22)a	.609	.753	.531	.601
(21)b,(22)b	.595	.773	.238	.754

Note. Each correlation involves the set of 36 complex observations. All coefficients in columns (1), (2), and (4) are reliable ($p < .01$). Column (3) shows the correlation between predicted and observed probabilities when random numbers in the unit interval are substituted for the provisional chances provided by f in the quadratic program (20), using averaged data. In each case, the correlations in columns (2) and (3) differ reliably ($p < .05$). The correlations in the column (4) are the average of six computations with 32 randomly chosen complete conjunctions.

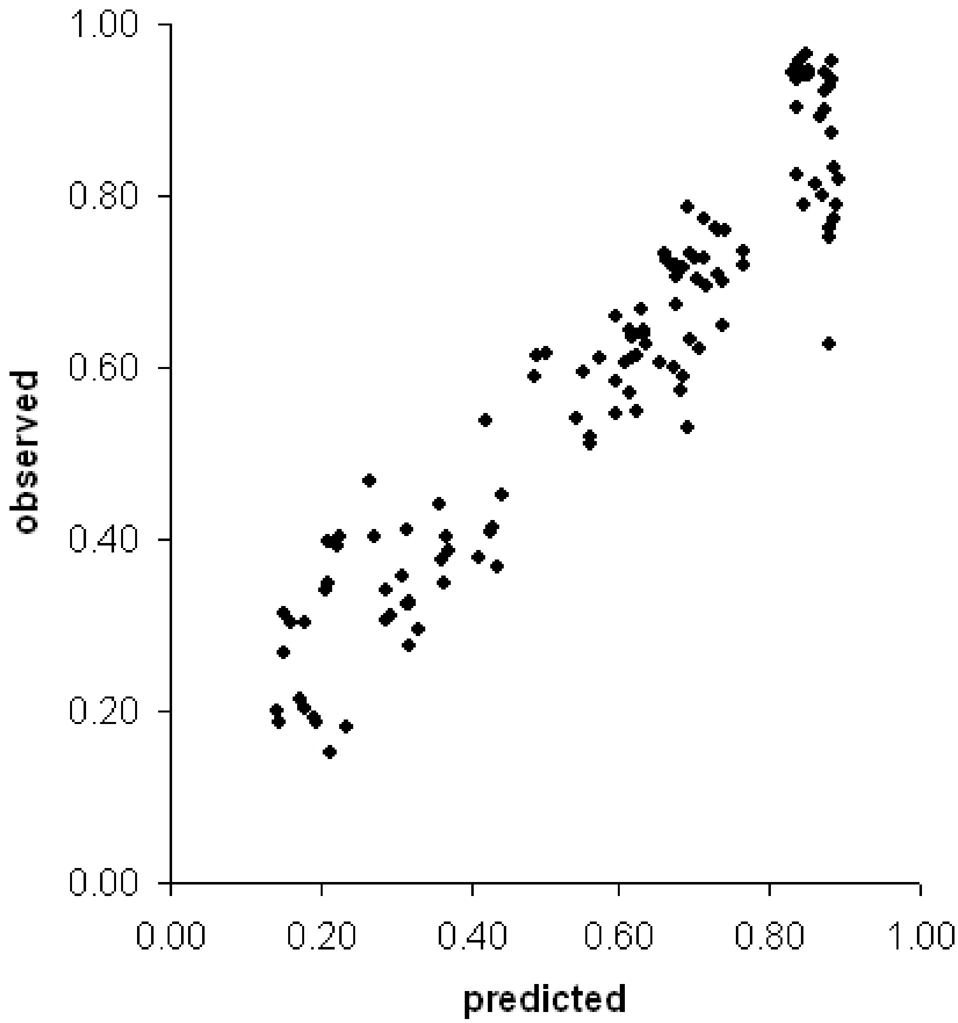

Figure 1: **Predicted versus observed probabilities for the 200 arguments.**

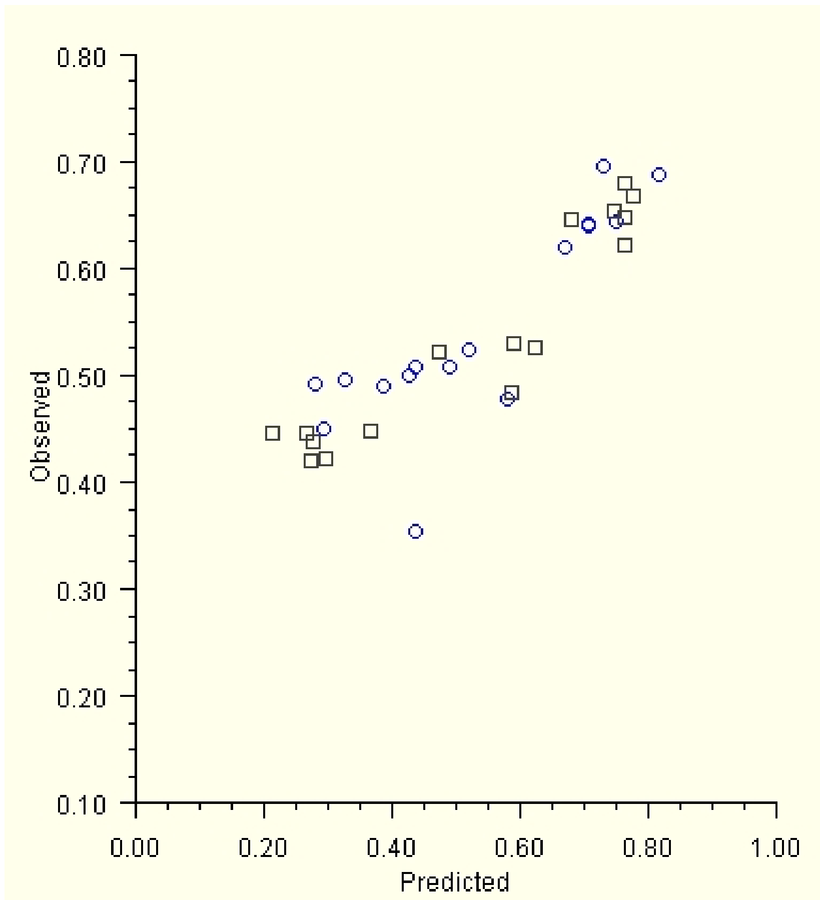

Figure 2: **Predicted versus observed probabilities for 32 arguments.** The circles represent the sixteen one-premise arguments in the follow-up experiment; the squares represent the sixteen two-premise arguments.

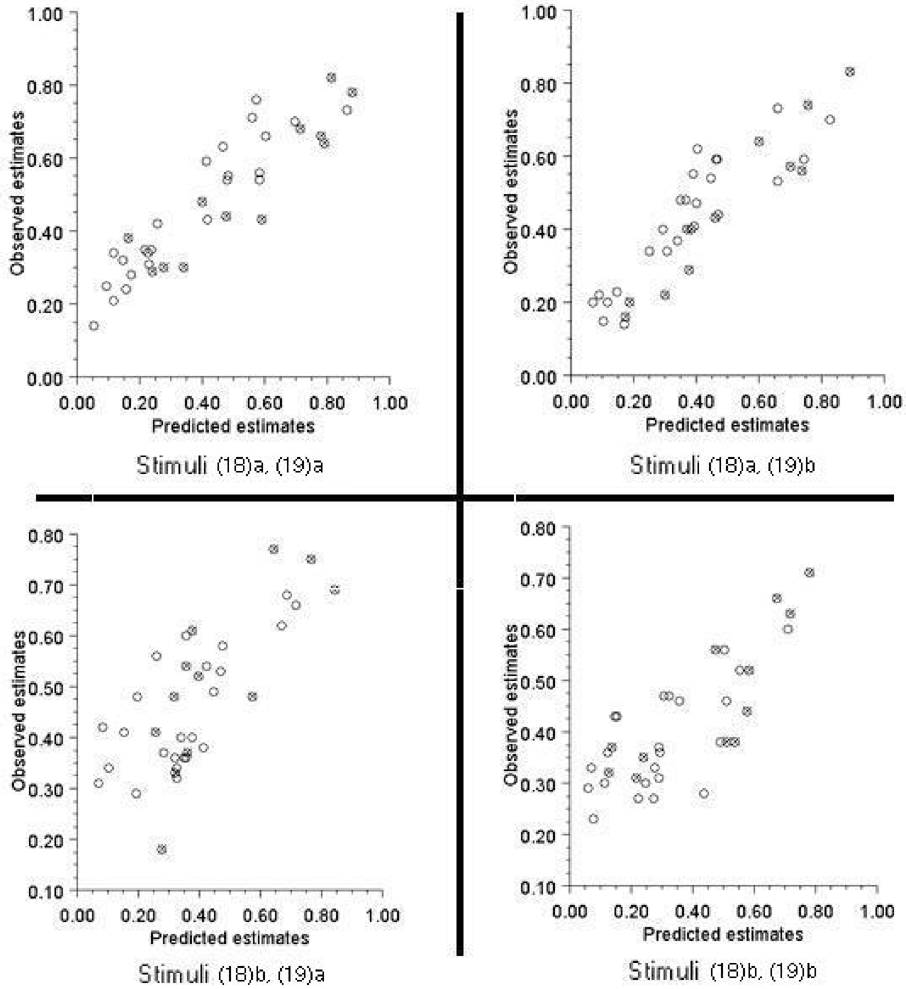

Figure 3: **Predicted versus observed probabilities for the four sets of stimuli.** Open circles correspond to the 24 absolute, complex statements. Crossed circles correspond to conditional statements.