No method of ampliative inference respects conditionalization*

Daniel Osherson

August 15, 2001

Let two events A, B be given. We consider probability distributions over the partition $\boldsymbol{P} = \{A \cap B, A \cap \bar{B}, \bar{A} \cap B, \bar{A} \cap \bar{B}\}$.

By a "constraint" is meant a probabilistically coherent set of statements each of the form $\mathsf{Prob}(E) = x$, where E is a subset of $\boldsymbol{P}$. Let $\mathcal{C}$ be the class of constraints. By a "method of ampliative inference" is meant any total function M from $\mathcal{C}$ to the class of probability distributions over $\boldsymbol{P}$, provided that M has the following property.

> For all $c \in \mathcal{C}$, $M(c)$ satisfies every member of c.

We say informally that M *must respect its input constraints*. For example, if $c = \{\mathsf{Prob}(A) = .9, \mathsf{Prob}(A \cap B) = .5\}$, then any method M of ampliative inference must be such that $M(c)(A) = .9$ and $M(c)(A \cap B) = .5$. The MAXENT recommendation is a method of ampliative inference in this sense (see [1, §11.1]).

Given a constraint c and an event X, we call X *coherent for* c just in case $c \cup \{\mathsf{Prob}(X) = 1\}$ is coherent. Now let M be a method of ampliative inference. We say that M *is compatible with conditioning* just in case for every constraint c and event X coherent for c, if $M(c)(X) > 0$ then $M(c)$ conditioned on X is the same distribution as $M(c; \mathsf{Prob}(X) = 1)$. The latter distribution is the result of applying M to the constraint consisting of c with $\mathsf{Prob}(X) = 1$ added in.

Now we want to show that there are no methods of ampliative inference compatible with conditioning. For a contradiction, suppose that M were such a method. Consider the following three constraints:

*Electronic mail: osherson@princeton.edu.

(1) $\{\mathsf{Prob}(A) = .3,\ \mathsf{Prob}(B) = .8\}$

(2) $\{\mathsf{Prob}(A \cup B) = 1.0\}$

(3) $\{\mathsf{Prob}(\bar{A} \cup B) = 1.0\}$

It is easy to verify that each of $A \cup B$ and $\bar{A} \cup B$ is coherent for (1). Suppose we apply M to (1). Then, since M respects its input constraints:

(4) $M(1)(A) = .3$ and $M(1)(B) = .8$, where $M(1)$ is the distribution that results from applying M to constraint (1) (and similarly below).

Suppose that after applying M to (1), we learn that $A \cup B$ occurs, which figures in constraint (2). Then, to respect conditionalization, it is required that:

(5) $M(1,2)$ is the conditionalization of $M(1)$ on $A \cup B$. Also, $M(1,2)(A) = .3$ and $M(1,2)(B) = .8$ [since M respects its input constraints, including (1)].

Suppose now that after applying M to (1), we learn $\bar{A} \cup B$ instead of $A \cup B$. Then, to respect conditionalization, it is required that:

(6) $M(1,3)$ is the conditionalization of $M(1)$ on $\bar{A} \cup B$. Also, $M(1,3)(A) = .3$ and $M(1,3) = .8$.

Now let us show that Properties (4) - (6) are not jointly satisfiable by any such function M. For notational ease, let:

$$M(1)(A \cap B) = ab,\ M(1)(A \cap \bar{B}) = a\bar{b},\ M(1)(\bar{A} \cap B) = \bar{a}b,\ M(1)(\bar{A} \cap \bar{B}) = \bar{a}\bar{b}.$$

By Property (4):

(7) $ab + a\bar{b} = .3$

$ab + \bar{a}b = .8$

By Property (5), $M(1,2)$ is the conditionalization of $M(1)$ on $A \cup B$, and this conditionalization must respect the information in constraint (1). Hence [because $M(1)(A \cap (A \cup B)) = ab + a\bar{b}$, $M(1)(B \cap (A \cup B)) = ab + \bar{a}b$, and $M(1)(A \cup B) = ab + a\bar{b} + \bar{a}b$]:

$$(8) \qquad \frac{ab + a\bar{b}}{ab + a\bar{b} + \bar{a}b} = M(1)(A|A \cup B) = M(1,2)(A) = .3$$
$$\frac{ab + \bar{a}b}{ab + a\bar{b} + \bar{a}b} = M(1)(B|A \cup B) = M(1,2)(B) = .8$$

A little manipulation shows that (7) and (8) imply:

(9) $ab = .1 \quad \bar{a}b = .7 \quad a\bar{b} = .2 \quad \bar{a}\bar{b} = 0$

By Property (6), $M(1,3)$ is the conditionalization of $M(1)$ on $\bar{A} \cup B$, and this conditionalization must also respect the information in constraint (1). Hence [because $M(1)(A \cap (\bar{A} \cup B)) = ab$, $M(1)(B \cap (\bar{A} \cup B)) = ab + \bar{a}b$, and $M(1)(\bar{A} \cup B) = ab + \bar{a}b + \bar{a}\bar{b}$]:

$$(10) \qquad \frac{ab}{ab + \bar{a}b + \bar{a}\bar{b}} = M(1)(A|\bar{A} \cup B) = M(1,3)(A) = .3$$
$$\frac{ab + \bar{a}b}{ab + \bar{a}b + \bar{a}\bar{b}} = M(1)(B|\bar{A} \cup B) = M(1,3)(B) = .8$$

But both equations in (10) are contradicted by (9).

Teddy Seidenfeld (personal communication) characterizes the foregoing fact in the following way. “It is impossible to have a general method that always treats new (consistent) evidence as probabilistically irrelevant to the old constraints.”

References

[1] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley Interscience, New York NY, 1991.