Independence Assumptions and Bayesian Updating

Clark Glymour

Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.

Recommended by E. Pednault

Relying on a result they attribute to Hussian [1], Pednault, Zucker, and Muresan [2] claim to have proved that the independence assumptions used in the PROSPECTOR [3] program do not permit the probabilities of hypotheses to be changed by any evidence. The purpose of this note is to show that the theorem stated by Pednault et al. is false, as is the result claimed by Hussian.

Let P be a probability function and let $H_1, \ldots, H_n$ be a set of jointly exhaustive and mutually exclusively hypotheses in the sense that

$$\sum_i P(H_i) = 1 \tag{1}$$

and

$$P(H_i \ \& \ H_j) = 0 \quad \text{if} \ i \neq j . \tag{2}$$

Let $E_1, \ldots, E_m$ be a sequence of evidence sentences. We say that *no updating occurs* iff for every subsequence $E_i, \ldots, E_k$ of $E_1, \ldots, E_m$ and every H_i

$$\frac{P(H_i/E_i, \ldots, E_k)}{P(\bar{H}_i/E_i, \ldots, E_k)} = \frac{P(H_i)}{P(\bar{H}_i)} .$$

The independence assumptions used in the PROSPECTOR program are

$$P(E_1, \ldots, E_m/H_i) = \prod_{j=1}^{m} P(E_j/H_i) , \tag{3}$$

$$P(E_1, \ldots, E_m/\bar{H}_i) = \prod_{j=1}^{m} P/(E_j/\bar{H}_i) . \tag{4}$$

Pednault et al. are surely correct in claiming that requirements (3) and (4) are excessively strong restrictions on the probability function P. Indeed, Pearl [4] and Kim [5] have argued that keeping requirement (3) alone is empirically more reasonable and yields an efficient updating scheme. If, however, Hussian's result were correct, (3) and the prior independence of the evidence would be inconsistent with any updating, and Pednault et al. further claim that:

"**Proposition.** *If the hypotheses* H_1, H_2, ..., H_N *with* $N > 2$ *are complete and mutually exclusive, i.e., if* $\Sigma_{i=1}^{N} P(H_i) = 1$, *and if the assumptions* (3) *and* (4) *are satisfied, then ... no updating takes place.*"

That this proposition is false may be shown by producing a system of five or more events H_1, H_2, H_3, E_1, E_2, and a probability function P, such that equations (1), (2), (3), (4) hold but updating *does* take place.

Consider the following case:

We specify that

   (i) $P(H_i) = \frac{1}{3}$ for all i,
  (ii) $P(H_i \,\&\, H_j) = 0$ for $i \neq j$,
 (iii) $P(E_1) = \frac{1}{2}$,
 (iv) $P(E_1 \,\&\, H_i) = \frac{1}{6}$ for all i,
  (v) $P(E_2) = \frac{1}{3}$,
 (vi) $P(E_2 \,\&\, H_1) = \frac{1}{3}$,
(vii) $P(E_2 \,\&\, H_2) = P(E_2 \,\&\, H_3) = 0$,
(viii) $P(E_1 \,\&\, E_2) = \frac{1}{6}$.

These conditions suffice to determine the value of P for every Boolean combination of E_1, E_2, H_1, H_2, H_3. In particular, they entail that

$$P(E_1/H_i) = P(E_1 \,\&\, H_i)/P(H_i) = \tfrac{1}{2},$$

$$P(E_2/H_1) = 1,$$

$$P(E_2/H_2) = P(E_2/H_3) = 0,$$

$$P(E_1 \,\&\, E_2/H_1) = \tfrac{1}{2},$$

$$P(E_1 \,\&\, E_2/H_2) = P(E_1 \,\&\, E_2/H_3) = 0,$$

as can readily be verified from the representation of the probability distribution as a Venn diagram (see Fig. 1).

Observe that:

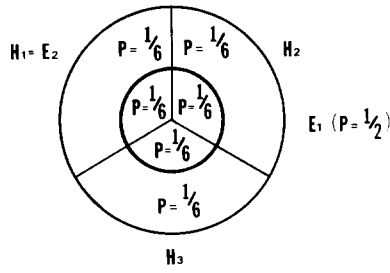$$\sum P(H_i) = 1 \quad \text{by (i)}. \tag{1}$$

FIG. 1.

$$P(H_i \;\&\; H_j) = 0 \quad \text{if } i \neq j \quad \text{by (ii)}. \tag{2}$$

$$P(E_1 \;\&\; E_2/H_i) = P(E_1/H_i) \cdot P(E_2/H_i), \tag{3}$$

because if $i = 1$,

$$P(E_1 \;\&\; E_2/H_i) = \tfrac{1}{2} = \tfrac{1}{2} \cdot 1 = P(E_1/H_i) \cdot P(E_2/H_i) \;;$$

and if $i = 2$ or 3.

$$P(E_1 \;\&\; E_2/H_i) = 0 = \tfrac{1}{6} \cdot 0 = P(E_1/H_i) \cdot P(E_2/H_i).$$

And

$$P(E_1 \;\&\; E_2/\bar{H}_i) = P(E_1/\bar{H}_i) \cdot P(E_2/\bar{H}_i), \tag{4}$$

because if $i = 1$,

$$P(E_1 \;\&\; E_2/\bar{H}_i) = P(E_1 \;\&\; E_2/H_2 \vee H_3)$$

$$= \frac{P(E_1 \;\&\; E_2 \;\&\; H_2) + P(E_1 \;\&\; E_2 \;\&\; H_3)}{P(H_2) + P(H_3)} = 0$$

and

$$P(E_1/\bar{H}_i) = \frac{P(E_1 \;\&\; H_2) + P(E_1 \;\&\; H_3)}{P(H_2) + P(H_3)} = \frac{\tfrac{1}{3}}{\tfrac{2}{3}} = \tfrac{1}{2}$$

and $P(E_2/\bar{H}_i) = 0$. So

$$P(E_1/\bar{H}_i) \cdot P(E_2/\bar{H}_i) = 0 \;;$$

and if $i = 2$ or 3, $k \neq i$, $k \neq 1$,

$$P(E_1 \& E_2 / \bar{H}_i) = P(E_1 \& E_2 / H_1 \vee H_k)$$

$$= \frac{P(E_1 \& E_2 \& H_1) + P(E_1 \& E_2 \& H_k)}{P(H_1) + P(H_k)}$$

$$= \frac{\frac{1}{6} + 0}{\frac{2}{3}} = \frac{3}{12} = \frac{1}{4}$$

and

$$P(E_1 / \bar{H}_i) \cdot P(E_2 / \bar{H}_i) = P(E_1 / H_1 \vee H_k) \cdot P(E_2 / H_1 \vee H_k)$$

$$= \left[\frac{P(E_1 \& H_1) + P(E_1 \& H_k)}{P(H_1) \& P(H_k)} \right]$$

$$\times \left[\frac{P(E_2 \& H_1) + P(E_2 \& H_k)}{P(H_1) + P(H_k)} \right]$$

$$= \left[\frac{\frac{1}{6} + \frac{1}{6}}{\frac{2}{3}} \right] \cdot \left[\frac{\frac{1}{3} + 0}{\frac{2}{3}} \right] = \frac{1}{4}.$$

So $P(E_1 \& E_2 / \bar{H}_i) = P(E_1 / \bar{H}_i) \cdot P(E_2 / \bar{H}_i)$. Thus all four conditions of the hypothesis of the theorem are met. But

$$P(H_i) / P(\bar{H}_i) = \tfrac{1}{2} \quad \text{for all } i$$

and

$$\frac{P(H_2 / E_1 \& E_2)}{P(\bar{H}_2 / E_1 \& E_2)} = \frac{P(H_3 / E_1 \& E_2)}{P(\bar{H}_3 / E_1 \& E_2)} = 0$$

and

$$\frac{P(H_1 / E_1 \& E_2)}{P(\bar{H}_1 / E_1 \& E_2)}$$

is undefined. Hence the theorem is false.

The error in the proof given by Pednault et al. lies in their use of a result claimed by Hussian, namely,

"**Theorem.** *Let the set of hypotheses H_i, $i = 1, 2, \ldots, N$ be exhaustive and mutually exclusive; if*

$$P(E_1, \ldots, E_m) = \prod_{k=1}^{m} P(E_k),$$

$$P(E_1, \ldots, E_m/H_i) = \prod_{k-1}^{m} P(E_k/H_i), \quad \text{for all } i,$$

then, for all i and k

$$P(E_k/H_i) = P(E_k)."$$

Hussian's derivation contains an algebraic error. The counterexample just given to the theorem claimed by Pednault et al. also shows that Hussian's result is false, and other counterexamples to the latter's theorem are easily produced. I do not know whether the result claimed by Pednault et al. would be true if one required in addition to (1)–(4) that for all i, $P(H_i/E_1, \ldots, E_k) \neq 0$.

From the Bayesian point of view, eliminative induction is just a special case in which the evidence determines a posterior probability of unity for one hypothesis and of zero for all others. Thus the example given here shows that the very strong independence assumptions in PROSPECTOR are consistent with learning by eliminative induction. This conclusion does not address the empirical adequacy of PROSPECTOR's updating scheme, which also contains provisions for ad hoc updating rules [6]. It does, I hope, lay to rest the claim that the program's Bayesian component is impotent if strictly applied.

REFERENCES

1. Hussian, A., On the correctness of some sequential classification schemes in pattern recognition, *IEEE Trans. Comput.* **21** (1972) 318–320.
2. Pednault, E., Zucker, S. and Muresan, I., On the independence assumption underlying subjective Bayesian updating, *Artificial Intelligence* **16** (1981) 213–222.
3. Duda, R., Hart, P. and Nilsson, N., Subjective Bayesian methods for rule-based inference systems, Tech. Note 124, Artificial Intelligence Center, SRI, International, Menlo Park, CA.
4. Pearl, J., Reverend Bayes on inference engines: a distributed hierarchical approach, *Proc. National Conf. Artificial Intelligence*, Carnegie-Mellon University, Pittsburgh, PA (1982) 133–136.
5. Kim, J. and Pearl, J., A computational model for combined causal and diagnostic reasoning in inference systems, *Proc. 8th Internat. Joint Conf. Artificial Intelligence*, Karlsruhe, West Germany, 1983.
6. Duda, R., Gaschnig, J. and Hart, P., Model design in the PROSPECTOR system for mineral exploration, in: D. Michie (Ed.), *Expert Systems in the Microelectronic Age* (Edinburgh Press, Edinburgh, 1979) 153–167.