Errata: “The relationship between knowledge, belief, and certainty”

Joseph Y. Halpern
Cornell University
halpern@cs.cornell.edu

There is an error in Theorem 7.1 in “The relationship between knowledge, belief, and certainty” (*Annals of Artificial Intelligence and Mathematics* **4**, 1991, pp. 301–322). The theorem says that F is a uniform frame iff every instance of Miller's principle is valid in F. It is true that every instance of Miller's principle is valid in a uniform frame. The converse does not hold in general.

The problem in the proof of Theorem 7.1 occurs in the proof of claim (6) on p. 320. It is claimed that if $a \neq b$, then we can find an interval $I = [d, e]$ such that $a \in I$, $b \notin I$, and $d > 0$. This is false if $a = 0$. The proof is correct as long as $a \neq 0$ and, in fact, this argument is basically the key to proving the corrected version of the theorem, as we now show.

The correct statement of the theorem is that F is an *almost-uniform* frame iff every instance of Miller's principle valid in F. An almost-uniform frame is one where uniformity holds almost everywhere in the following sense. Using the notation of the paper, given a frame $F = (S, PR)$, let $[s] = \{t \in S : PR(s) = PR(t)\}$. Note the sets $[s]$ form a partition of S. Let $G_F = \{s : PR(s)([s]) = 1\}$. Then F is almost-uniform if

AS1. $PR(s)(G_F) = 1$ for all $s \in S$

AS2. $PR(s)(t) = \sum_u PR(s)(u) \times PR(u)(t)$.

By way of contrast, F is uniform if $G_F = S$. Note that if $s \in G_F$, then AS2 holds at s, since if $PR(s)(u) > 0$, then $PR(u)(t) = PR(s)(t)$. Thus, if a frame is uniform then it is almost uniform; the converse does not hold in general.

It is also worth noting that in the presence of AS1, AS2 simplifies to

AS2$'$. $PR(s)(t) = PR(s)[t] \times PR(t)(t)$.

For if $u \notin G_F$, then by AS1, $PR(s)(u) = 0$; if $u \in G_F$ and $u \notin [t]$, then $PR(u)(t) = 0$; and if $u \in G_F$ and $u \in [t]$, then $PR(u)(t) = PR(t)(t)$. Given this, it is easy to show that almost-uniform frames are equivalent to Samet's *almost-partition models* [4].

Theorem 7.1: (corrected) A frame F is almost uniform iff every instance of Miller's principle is valid in F.

Proof: Suppose F is almost uniform and let $N = (S, \pi, PR)$ be a probability structure based on F. We want to show that Miller's principle holds for all $s \in S$. Given a formula φ and an interval $I = [a, b]$, notice that $S_{w(\varphi)\in I}$ can be partitioned into sets of the form $[t]$, since $PR(t')(S_\varphi)$ is the same for each element $t' \in [t]$. Standard calculations show

$$PR(s)(S_{\varphi\wedge w(\varphi)\in I}) = PR(s)(S_\varphi | S_{w(\varphi)\in I}) \times PR(s)(S_{w(\varphi)\in I}) \tag{1}$$

and

$$PR(s)(S_\varphi | S_{w(\varphi)\in I}) = \sum_{[t]\subseteq S_{w(\varphi)\in I}, PR(s)([t])>0} PR(s)(S_\varphi | [t]) \times PR(s)([t] | S_{w(\varphi)\in I}). \tag{2}$$

Using AS2$'$, if $PR(s)([t]) > 0$, we have

$$PR(s)(S_\varphi | [t]) = \sum_{t'\in S_\varphi\cap[t]} PR(s)(t')/PR(s)[t] = \sum_{t'\in S_\varphi\cap[t]} PR(t')(t') = \sum_{t'\in S_\varphi\cap[t]} PR(t)(t'). \tag{3}$$

By AS1, if $PR(s)([t]) > 0$, then $t' \in G_F$ for some, and hence all, $t' \in [t]$. Thus, $PR(t)([t]) = 1$, and

$$\sum_{t'\in S_\varphi\cap[t]} PR(t)(t') = PR(t)(S_\varphi). \tag{4}$$

Note that for $[t] \subseteq S_{w(\varphi)\in I}$, we have we have $a \leq PR(t)(S_\varphi) \leq b$. Using this observation and (2), (3), and (4), we get

$$PR(s)(S_\varphi | S_{w(\varphi)\in I}) \leq \sum_{[t]\subseteq S_{w(\varphi)\in I}, PR(s)([t])>0} bPR(s)([t] | S_{w(\varphi)\in I}) \leq b$$

and similarly

$$PR(s)(S_\varphi | S_{w(\varphi)\in I}) \geq a.$$

Putting this back into (1), we get that

$$aPR(s)(S_{w(\varphi)\in I}) \leq PR(s)(S_{\varphi\wedge w(\varphi)\in I}) \leq bPR(s)(S_{w(\varphi)\in I}).$$

Miller's principle now follows.

For the converse, suppose that Miller's principle is valid in $F = (S, PR)$. We want to show that F is almost uniform. We prove this using the following claim:

Claim: For all $x, y, z \in S$, if $PR(x)(y) > 0$ and $PR(y)(z) > 0$, then $PR(x)(z) > 0$ and $PR(y)(z) = PR(z)(z)$.

Proof: Suppose $PR(y)(z) = a$. By Miller's Principle, we have that

$$PR(x)(\{z\} \cap \{w : PR(w)(z) = a\}) = aPR(x)(\{w : PR(w)(z) = a\}) \geq aPR(x)(y) > 0.$$

(The first equality follows by considering the structure based on F where p is true only at the state z and applying Miller's Principle to p and arbitrarily small intervals containing a.) Thus, $z \in \{w : PR(w)(z) = a\}$, so $PR(z)(z) = PR(y)(z) = a$ and $PR(x)(z) > 0$, proving our result. ∎

Note that this claim tells us that in frames satisfying Miller's Principle, the support relation is transitive.

Returning to the proof, we first show that if $s \in S$, then $PR(s)(G_F) = 1$. To do this, it suffices to show that if $PR(s)(t) > 0$ and $PR(t)(u) > 0$, then $u \in [t]$, that is, $PR(t) = PR(u)$. To show this, it suffices to show that if $PR(u)(v) > 0$, then $PR(t)(v) = PR(u)(v)$.

By the claim, it immediately follows that if $PR(s)(t) > 0$, $PR(t)(u) > 0$, and $PR(u)(v) > 0$, then $PR(t)(v) > 0$, $PR(s)(u) > 0$, $PR(t)(v) = PR(v)(v)$, and $PR(u)(v) = PR(v)(v)$. Thus, $PR(t)(v) = PR(u)(v)$.

To show that AS2 holds, by our earlier discussion, it suffices to show that AS2$'$ holds. Consider any state t. If $PR(t)(t) = a$, then by Miller's Principle (applied to a formula p in a structure where p is true only at t), we have $PR(s)(\{t\} \cap \{w : PR(w)(t) = a\}) = aPR(s)(\{w : PR(w)(t) = a\})$. Clearly $\{t\} \cap \{w : PR(w)(t) = a\} = \{t\}$. Thus $PR(s)(t) = aPR(s)(\{w : PR(w)(t) = a\})$. If $a = 0$, then AS2$'$ is immediate. If $a > 0$, then by using the facts that (1) $PR(u)(t) = 0$ if $u \in G_F$ and $u \notin [t]$ and (2) $PR(s)(G_F) = 1$, we get that $PR(s)(t) = aPR(s)([t]) = PR(t)(t) \times PR(s)([t])$, again giving us AS2$'$. ∎

Samet [3, 4] proves a result even stronger than Theorem 7.1. He shows, using Markov chains in [4] (where he only considers finite frames) and ergodic theory in [3] (where he allows arbitrary frames), that AS2 (which he calls the *averaging* property) implies AS1 and that therefore Miller's Principle holds in a frame iff it satisfies AS2. (See Theorems 4 and 5 in [4] and Theorem 4 in [3].)

Gaifman [1] also proves a variant of Theorem 7.1 in his framework, with two agents (see his Theorem 2). As it says in the paper, he uses a version of Miller's Principle that is stronger than that considered here. He requires not only that that $w_1(\varphi | w_2(\varphi) \in I) \in I$ but that $w_1(\varphi | \psi \wedge w_2(\varphi) \in I) \in I$, where ψ is a conjunction of formulas of the form $w_2(\psi') \in I'$. Gaifman shows that this stronger version of Miller's Principle is equivalent to the requirement that analogues of AS1 and AS2 (appropriate for his framework, with two agents) hold.

The change in Theorem 7.1 affects other statements later in the paper. Theorem 7.3 and Corollary 7.4, which show that KD45^C provides a sound and complete axiomatization for certainty in uniform structures, are correct, but since Miller's principle is not characterized by uniformity, the sentence after Corollary 7.4, which says that KD45^C provides a sound and complete axiomatization for the class of structures characterized by Miller's principle, is not. Fortunately, a slight modification of KD45^C provides a characterization of almost-uniform structures. We simply replace the axiom **5** by

5′. $K(\neg K\varphi \Rightarrow K\neg K\varphi)$

This axiom can be shown to be characterized by a variant of the Euclidean condition. We say a relation $\mathcal{K}$ satisfies the Euclidean′ condition if $(s,t), (t,u), (t,v), \in \mathcal{K}$ implies that $(u,v) \in \mathcal{K}$. Using standard techniques of modal logic, we can show that **5′** characterized by the Euclidean′ condition. It is easy to see that **5′** (with K replaced by $Cert$) is sound in almost-uniform frames.

Let KD45′ consist of the axioms of KD45, with **5** replaced by **5′**. An argument similar to that of Theorem 7.2 shows that KD45′ characterizes almost-uniform frames. I omit further details here.

Finally, note there is also a persistent confusion of the role of the agent and the expert in the discussion of Gaifman's work in the paper. In particular:

- On page 315, line 2, and again on p. 317, line -9, w_2 should be taken to represent the expert and w_1 the agent about whom we are reasoning, not the other way around. Another interpretation of w_1 and w_2, taken by Samet [2] (where results essentially identical to Gaifman's are proved) is that w_1 is the agent's prior probability and w_2 his posterior, after learning some information. If the state defines the information, then we would expect w_2 to depend on the state.

- On p. 317, line -4, PR_1, the agent's probabiity assignment, should be taken to be independent of the state, not PR_2; similarly, on page 318, line 4, it should be PR_2, not PR_1.

- On p. 318, line 5, note that the condition $pr(S_{good}) = 1$ is the analogue of AS1 (since it says that $PR_1(s)(S_{good}) = 1$ for all s). The statement here is not correct since, as mentioned above, Gaifman also requires an analogue of AS2 as well as AS1 for his result.

I thank Dov Samet for alerting me to all these problems and for pointing out the need for AS2.

References

[1] H. Gaifman. A theory of higher order probabilities. In J. Y. Halpern, editor, *Theoretical Aspects of Reasoning about Knowledge: Proc. 1986 Conference*, pages 275–292. Morgan Kaufmann, San Francisco, Calif., 1986.

[2] D. Samet. Bayesianism without learning. unpublished manuscript, 1997.

[3] D. Samet. Quantified beliefs and believed quantities. In I. Gilboa, editor, *Theoretical Aspects of Rationality and Knowledge: Proc. Seventh Conference*, pages 263–272. Morgan Kaufmann, San Francisco, Calif., 1998.

[4] D. Samet. On the triviality of higher-order beliefs. *Journal of Philosophical Logic*, forthcoming.