A Probabilistic Theory of Coherence

BRANDEN FITELSON

1. The Coherence Measure $\mathcal{C}$

Let **E** be a set of *n* propositions $E_1, ..., E_n$. We seek a probabilistic measure $\mathcal{C}(\mathbf{E})$ of the 'degree of coherence' of **E**. Intuitively, we want $\mathcal{C}$ to be a quantitative, probabilistic generalization of the (deductive) *logical coherence* of **E**. So, in particular, we require $\mathcal{C}$ to satisfy the following intuitive desideratum.

$$(1) \quad \mathcal{C}(\mathbf{E}) \text{ is } \begin{cases} \text{Maximal (positive, constant)} & \text{if the } E_i \text{ are logically equivalent (and } \mathbf{E} \text{ is satisfiable)} \\ > 0 & \text{if } \mathbf{E} \text{ is positively dependent (see below for def.)} \\ 0 & \text{if } \mathbf{E} \text{ is independent (see below for def.)} \\ < 0 & \text{if } \mathbf{E} \text{ is negatively dependent (see below for def.)} \\ \text{Minimal (negative, constant)} & \text{if all subsets of } \mathbf{E} \text{ are unsatisfiable} \end{cases}$$

Desideratum (1) captures the qualitative features that a probabilistic generalization of logical coherence should satisfy — it requires $\mathcal{C}$ to respect the extreme deductive cases, and to be properly sensitive to probabilistic dependence (a general notion of probabilistic dependence will be defined precisely, and in a slightly non-standard way, below).

I propose a probabilistic measure of coherence $\mathcal{C}$ based on a slight modification of Kemeny and Oppenheim's (1952) measure of factual support *F*. The formulation of $\mathcal{C}$ is somewhat intricate. We begin with some preliminary definitions. First, we define the two-place function $F(X,Y)$. $F(X,Y)$ may be interpreted as the degree to which one proposition *Y* supports another proposition *X* (relative to a finitely additive, regular, Kolmogorov (1956) probability function Pr).[1]

$$F(X,Y) = \begin{cases} \dfrac{\Pr(Y \mid X) - \Pr(Y \mid \neg X)}{\Pr(Y \mid X) + \Pr(Y \mid \neg X)} & \text{if } X \text{ is contingent and } Y \text{ is not a necessary falsehood} \\ 1 & \text{if } X \text{ and } Y \text{ are necessary truths} \\ 0 & \text{if } X \text{ is a necessary truth and } Y \text{ is contingent} \\ -1 & \text{if } Y \text{ is a necessary falsehood} \end{cases}$$

---

[1]For simplicity, I am assuming that the probability function Pr is regular or *strictly coherent* in the sense of Shimony 1952. That is, I assume that Pr assigns probability 1 *only* to necessary truths, and probability 0 *only* to necessary falsehoods (and, therefore, that $\Pr(X \mid Y)$ is extreme only if *X* and *Y* are logically dependent). We could weaken this assumption, but it would complicate our definitions (as would a generalization to infinite sets of propositions, and perhaps to countably additive Pr). In the context of Bayesian epistemology, assigning extreme probability to contingent propositions is controversial (Jeffrey 1992). Our *F* is more complete than Kemeny and Oppenheim's (1952), which is defined only for contingent *X* and *Y*. We need this additional structure in *F* to ensure that $\mathcal{C}$ satisfies (1). Even our $F(X,Y)$ remains undefined in the case where *X* is a necessary falsehood but *Y* is not. We don't need to worry about this case here (however, note that this omission does make $\mathcal{C}$ a partial function). But, I think we can at least say that if *X* is contradictory, then $F(X,Y) \leq 0$ (*i.e.*, that nothing supports a contradiction).

Next, we use F to define the probabilistic independence and positive/negative dependence of a set of propositions **E**. Let $\mathbf{P}_i$ be the power set (*sans* null set) of the set $\mathbf{E}\backslash\{E_i\}$ (unless **E** is a singleton, in which case **P** = **E**). And, for each $x \in \mathbf{P}_i$, let X be the conjunction of the elements of x. Then, we define:

$$\mathbf{E} \text{ is } \begin{cases} \text{positively dependent} & \text{iff for all } E_i \in \mathbf{E} \text{ and for all } x \in \mathbf{P}_i,\ F(E_i, X) > 0 \\ \text{independent} & \text{iff for all } E_i \in \mathbf{E} \text{ and for all } x \in \mathbf{P}_i,\ F(E_i, X) = 0 \\ \text{negatively dependent} & \text{iff for all } E_i \in \mathbf{E} \text{ and for all } x \in \mathbf{P}_i,\ F(E_i, X) < 0 \end{cases}$$

Our definition of independence is (nearly) logically equivalent (except for certain extreme cases in which E_i and/or X are non-contingent[2]) to the standard definition of the independence of sets **E** seen in probability texts (Kolmogorov 1952: §I.5). Interestingly, the concepts of positive and negative dependence of a set **E** are *not* typically defined (at all) in standard probability textbooks (at least, I have not seen such definitions). Like the standard definition of independence, my definitions of positive/negative dependence require correlation (or anti-correlation) of *all subsets* of propositions in **E** (*i.e.*, pairwise, 3-wise, 4-wise, *etc*.). This means that sets with "mixed" correlations or anti-correlations (*e.g.*, pairwise but not 3-wise correlation or anti-correlation, etc.) will *not* count as *dependent* sets on my definition.

Now, we're ready to introduce the components of our quantitative measure of coherence $\mathcal{C}$. Let $\mathbf{S} = \bigcup\{\{F(E_i,X) \mid x \in \mathbf{P}_i\} \mid E_i \in \boldsymbol{E}\}$. In general, **S** will have $n \cdot (2^{n-1}-1)$ elements, where $n > 1$ is the number of elements of $\boldsymbol{E}$ (if $n = 1$, then **S** will have just one element: $F(E_1,E_1)$). For instance, if $n = 3$, then **S** will be $\{F(E_1,E_2), F(E_1,E_3), F(E_1,E_2\&E_3), F(E_2,E_1), F(E_2,E_3), F(E_2,E_1\&E_3), F(E_3,E_2), F(E_3,E_1), F(E_3,E_1\ \&E_2)\}$, which has $3\cdot(2^{3-1}-1) = 3\cdot3 = 9$ elements. Finally, we define $\mathcal{C}$ as follows.[3]

$$\mathcal{C}(\boldsymbol{E}) =_{df} \text{mean}(\mathbf{S})$$

That is, $\mathcal{C}$ is simply the *mean* value of **S**. It is easy to verify that $\mathcal{C}$ satisfies (1). This is because F is a proper probabilistic generalization of deductive *support* (Kemeny & Oppenheim 1952, Fitelson 2001: §3.2.3). In particular, we have:

[2]The standard (Kolmogorov 1952: §I.5) definition of independence says that E is independent *of itself* in cases where E is non-contingent. This is because in such cases $\Pr(E\ \&\ E) = \Pr(E) \cdot \Pr(E) = 1$ *or* 0, if $\Pr(E) = 1$ *or* 0, respectively. I think this is unintuitive. Intuitively, all consistent propositions are (maximally) positively correlated with (dependent on) themselves, and all contradictions are (maximally) negatively correlated with (dependent on) themselves. This is what my definitions of independence and dependence entail, owing to my definition of F on which $F(E,E) = 1$ for all consistent E, and $F(E,E) = -1$ for all self-contradictory E.

[3]Here, I take $\mathcal{C}$ to be the *straight* average of **S**. One could generalize our definition of $\mathcal{C}$ so as to assign different weights to different types of (in)dependence (*e.g.*, 2-wise (in)dependence *vs* 3-wise (in)dependence might be weighted differently in the average, or negative dependence among certain subsets of **E** might be weighed more heavily than positive dependence, *etc*.). One can think of our definition of $\mathcal{C}$ as assuming that all of the F-components of $\mathcal{C}$ (in **S**) have the same weight.

$$\mathcal{C}(\mathbf{E}) \text{ is } \begin{cases} 1 & \text{if the } E_i \text{ are logically equivalent (and } \mathbf{E} \text{ is satisfiable)} \\ > 0 & \text{if } \mathbf{E} \text{ is positively dependent} \\ 0 & \text{if } \mathbf{E} \text{ is independent} \\ < 0 & \text{if } \mathbf{E} \text{ is negatively dependent} \\ -1 & \text{if all subsets of } \mathbf{E} \text{ are unsatisfiable} \end{cases}$$

2. $\mathcal{C}$ *vs* Shogenji's Measure of Coherence

Interestingly, Shogenji's ratio measure of coherence (Shogenji 1999, Akiba 2000, Shogenji 2001) does *not* satisfy (1).[4] In the case where the E_i are logically equivalent (hence $\Pr(E_i) = p$, for all i), Shogenji's measure reduces to

$$\frac{\Pr(E_1 \,\&\cdots\&\, E_n)}{\Pr(E_1)\times\cdots\times\Pr(E_n)} = \frac{p}{p^n} = \frac{1}{p^{n-1}}$$

which is *not* a constant (nor is it maximal on the $[0,\infty)$ scale of Shogenji's measure), and still depends on the unconditional probabilities of the E_i. This is unintuitive, as this should be a case in which the degree of coherence is *maximal*, and does not depend on the priors of the E_i. Here, Shogenji's measure of coherence inherits an undesirable feature of the ratio measure of degree of *support* or *confirmation* (Pollard 1999, Fitelson 2001: §3.2.3).

It is also interesting to note that Shogenji's measure is based only on the n-wise (in)dependence of the set **E**. It is well known that a set **E** can be j-wise independent, but not i-wise independent, for any $i \neq j$ (indeed, we can have disagreement for any *combinations* of i- and j-wise independence as well – see (Pfeiffer 1994: §4.2) for several concrete examples). Since Shogenji's measure is based only on n-wise independence (dependence), in cases where a set is n-wise independent (dependent), but not j-wise (for some $j \neq n$) independent (dependent), Shogenji's measure does not take into account the 'mixed' nature of the coherence (incoherence) of **E** (and its subsets), and it judges **E** as having the same degree of coherence (incoherence) as a *fully independent* (or *fully dependent*) set. This seems incorrect to me. I think it's important for a measure of coherence to be sensitive to the (in)dependencies implicit in *all* subsets of **E**.

3. Akiba's Criticisms of Shogenji's Coherence Measure

Akiba's (1999) criticisms of Shogenji's coherence measure do not apply to our $\mathcal{C}$. For instance, Akiba complains that if E_1 entails E_2, then Shogenji's measure says that the degree of coherence of the set $\{E_1,E_2\}$ is $1/\Pr(E_2)$, which he finds unintuitive, since it only depends on the

[4]The fact that Shogenji's measure is always positive is a *merely conventional* (and therefore *insignificant*) violation of (1). This can be fixed simply by taking the logarithm of Shogenji's measure. The problems with Shogenji's measure discussed below are *not* merely conventional.

unconditional probability of E_2. I agree with Akiba's intuition here; and, so does $\mathcal{C}$. In such a case, we have (if E_1 and E_2 are contingent[5]):

$$\mathcal{C}(\{E_1, E_2\}) = \text{mean}(\mathbf{S}) = \frac{F(E_1, E_2) + F(E_2, E_1)}{2} = \frac{F(E_1, E_2) + 1}{2} = \frac{1}{1 + \Pr(E_2 \mid \neg E_1)}$$

So, $\mathcal{C}$ supports Akiba's intuition that the degree of coherence in this case should depend on the *precise relationship* between E_1 and E_2, and not merely on $\Pr(E_2)$ *alone* (similar things happen when $\mathcal{C}$ is applied to Akiba's other examples involving more than two events).

Akiba also discusses the problem of singleton sets of propositions. He points out that Shogenji's measure of coherence judges the self-coherence of *all* propositions to be the same. This is unintuitive, since (for instance) necessary truths should be viewed as more self-coherent than necessary falsehoods. Our measure $\mathcal{C}$ captures this intuition, since:

$$\mathcal{C}(\{E\}) = F(E,E) = \begin{cases} 1 & \text{if } E \text{ is a necessary truth} \\ -1 & \text{if } E \text{ is a necessary falsehood} \end{cases}$$

Nonetheless, it is still true that $\mathcal{C}$ judges the degree of coherence to be the same (+1) for all *satisfiable* singleton sets propositions. However, since there seem to be no clear countervailing intuitions about what a probabilistic theory of degree of coherence should say in the contingent, singleton case (Akiba 2000, Klein & Warfield 1994, Shogenji 1999), I do not view this as a shortcoming of my proposed probabilistic measure of coherence. Here, I am in agreement with Shogenji (1999) in understanding coherence as a *relation* between propositions. Intuitively, *all* propositions 'cohere with themselves' (maximally), except for necessary falsehoods.[6]

San José State University
San José, CA 95192–0096
branden@fitelson.org

References

Akiba, K. 2000. Shogenji's probabilistic measure of coherence is incoherent. *Analysis 60*: 356–59.

Fitelson, B. 2001. *Studies in Bayesian Confirmation Theory*. Ph. D. thesis, University of Wisconsin–Madison (Philosophy). The thesis can be downloaded in its entirety from the following URL `http://fitelson.org/thesis.pdf`.

Jeffrey, R. 1992. *Probability and the Art of Judgment*. Cambridge: Cambridge University Press.

Kemeny, J. and P. Oppenheim. 1952. Degrees of factual support. *Philosophy of Science 19*: 307–24.

Klein, P. and T. Warfield. 1994. What price coherence? *Analysis 54*: 129–32.

Kolmogorov, A. 1956. *Foundations of the Theory of Probability* (second english edition). New York: AMS Chelsea Publishing.

---

[5]If E_1 and E_2 are contingent and E_1 entails E_2, then the set $\mathbf{E} = \{E_1, E_2\}$ is positively dependent, and we have $F(E_2, E_1) = \Pr(E_1 \mid E_2) / \Pr(E_1 \mid E_2) = 1$, since $\Pr(E_1 \mid \neg E_2) = 0$.

[6] Thanks to Luc Bovens, Alan Hájek, Stephan Hartmann, Jim Hawthorne, and Tomoji Shogenji for useful discussions about coherence and probability. And, thanks to an anonymous referee for their comments and suggestions on previous drafts of this paper.

Pfeiffer, P. 1990. *Probability for Applications*. New York: Springer-Verlag.
Pollard, S. 1999. Milne's measure of confirmation. *Analysis 59*: 335–37.
Shimony, A. 1955. Coherence and the axioms of confirmation. *Journal of Symbolic Logic 20*: 1–28.
Shogenji, T. 1999. Is coherence truth-conducive? *Analysis 59*: 338–45.
Shogenji, T. 2001. On the probabilistic measure of coherence. Reply to 'Shogenji's probabilistic measure of coherence is incoherent' by K. Akiba. *Analysis 61*: 147–50.