Notes for Week 9 of *Confirmation*

10/31/07

Branden Fitelson

1 The Carnap–Popper Controversy

Almost all of the first-edition of Carnap's *LFP* is devoted to confirmation as *firmness* [which he explicates as $\Pr_{\top}(H \mid E) > r$, for a suitable "logical" conditional probability function $\Pr_{\top}(\cdot \mid \cdot)$ — see below for much more on the historical varieties of Carnapian "logical" probability functions]. But, in his discussion of Hempel, Carnap gives what he claims to be a counterexample to (SCC), which, as you'll recall from Week 7, is:

(SCC) If E confirms H relative to K and $H \vDash_K H'$, then E confirms H' relative to K.

This is rather puzzling, since (SCC) *holds* for confirmation as firmness (we'll return to this point when we discuss the "conjunction fallacy" in the final week of the seminar). What's happening here is that Carnap presupposes a different — probabilistic relevance — conception of confirmation in his discussion of Hempel, which he (there) calls "initial confrimation". This is very much like Maher's confirmation relation (from last week). It is something like "probability-raising, relative to empty background corpus". Popper wrote a (nasty) series of criticisms of Carnap, accusing him of "inconsistency" (or worse) on this score. While Popper's discussions were not always the most charitable (or helpful), he did make a compelling case that the probabilistic relevance conception is often the more important of the two. One example (that is similar to an example Popper discusses) involves a person ϕ who is wondering whether (H) they have a very rare disease. They just received a positive result (E) for the disease from highly reliable test. In such cases, it seems wrong to say that E "justifies" or "provides a reason to believe" or "supports" $\sim H$ (for ϕ). However, it is still reasonable — all things considered (where this *includes* E) — for ϕ to believe that they do not have the disease ($\sim H$). But, the right thing to say here is that it is *despite* having learned E that it is (still) reasonable for ϕ to believe $\sim H$. Intuitively, E *counter-supports* $\sim H$ (for ϕ) because E *supports* H (for ϕ). Thus, in this case, the firmness conception says that E confirms H, but the increase in firmness conception says that E disconfirms H. And, Popper argues that in such cases the increase in firmness conception is the more salient (or important). I think Popper is basically right here. And, Carnap did not really disagree. Largely in response to Popper, Carnap issued a second edition of *LFP*. But, the second edition does not make any changes to the main text. It just adds a brief preface, explaining how he should have been clearer about the two distinct conceptions of confirmation that were floating around in the first edition.[1] In any event, Popper should be given more credit for the shift that occurred in confirmation theory (that is, the shift to increase in firmness and away from firmness). But, largely because of his nastiness in this dispute with Carnap, his contribution is not properly acknowledged. See the Michalos book for lots more on this controversy.

2 Carnapian "Logical" Probability — A Trip Down Memory Lane

2.1 The Early Systems of *LFP*: $\mathfrak{c}^{\dagger}$ and $\mathfrak{c}^{*}$

In *LFP*, Carnap clearly seems to think that in order for confirmation relations that are defined in terms of conditional probabilities to count as "logical", the probability functions in question must *themselves* be "logical". This is somewhat odd, in light of what he says about "logicality" in deductive *vs* inductive logic:

> The principal common characteristic of the statements in both fields is their independence of the contingency of facts. This characteristic justifies the application of the common term 'logic' to both fields.
>
> That $\mathfrak{c}$ is an objective concept means this: if a certain $\mathfrak{c}$ value holds for a certain hypothesis with respect to a certain evidence, then this value is entirely independent of what anyone ... thinks about these sentences.
>
> They both involve the concept of range. 'E L-implies H' ['$E \vDash H$'] means the entire range of E is included in that of H, while ... '$\mathfrak{c}(H, E) = 3/4$' means three-fourths of the range of E is included in that of H.

[1] Carnap's brief remarks in the preface of LFP_2 already suggest that some radical changes would have been required in the rest of the text. As it turns out, the changes would have rippled through *LFP* in various crucial ways (I'll return to this later in the seminar).

The basic idea here already appears in the *Tractatus*. There, Wittgenstein talks about arguments that are "strong" but invalid — not *all* truth-value assignments on which E is true are truth-value assignments on which H is true, but "most" are. Before getting into Carnap's development of his theories of "logical" probability, I want to explain why *conditional probability* is a natural way to go here. It has to do with Carnap's remark on "range" above. A weaker requirement on "range" is the following *desideratum*:

($\mathcal{D}$) If $E \vDash H$, then $\mathfrak{c}(H, E)$ should take a *maximal* value, and if $E \vDash {\sim}H$, then $\mathfrak{c}(H, E)$ should take a *minimal* value.

This is a fundamental constraint on any "logical" confirmation function — it should *generalize* entailment and refutation. Interestingly, one of the most intuitive ways of "generalizing entailment" does *not* satisfy ($\mathcal{D}$). One classical definition of validity is as follows: the argument from E to H is valid iff $E \,\&\, {\sim}H$ is *impossible*. The most natural (and naïve) way of generalizing *this* definition would seem to be:

- The argument from E to H is "strong" iff $E \,\&\, {\sim}H$ is *improbable*. This leads to $\mathfrak{c}(H, E) = \Pr(E \supset H)$.

That is, one very natural way to "generalize entailment" is to take $\mathfrak{c}(H, E)$ to be the *probability of the material conditional* $E \supset H$. But, this is *not* the same as $\mathfrak{c}(H, E) = \Pr(H \mid E)$. It can be shown that $\Pr(H \mid E) \le \Pr(E \supset H)$. But, while $\Pr(H \mid E)$ satisfies ($\mathcal{D}$), $\Pr(E \supset H)$ does not. This is a good reason for Carnap to go for $\Pr(H \mid E)$ rather than $\Pr(E \supset H)$. But, Carnap goes much farther than this. In *LFP*, he seems to think that there is "a right way" to quantitatively generalize entailment — in a much finer-grained sense. Thus, his finer-grained notion of "logical range", implicit above. This brings us to Carnap's early theories of "logical" probability.

2.1.1 Carnap's First System: $\mathfrak{c}^\dagger$

Carnap's early systems operate over monadic predicate-logical languages $\mathcal{L}_Q^{m,n}$, which contain n monadic predicates and m individual constants (think: domains of size m). To fix ideas, consider the language $\mathcal{L}_Q^{2,2}$, which contains two monadic predicates F and G and two constants a and b. This yields 16 *state descriptions* of $\mathcal{L}_Q^{2,2}$. Table 1 depicts these 16 state descriptions of $\mathcal{L}_Q^{2,2}$ (ignore the last two columns for now):

Fa	Ga	Fb	Gb	State Descriptions (s_i)	$\mathfrak{m}^\dagger(s_i)$	$\mathfrak{m}^*(s_i)$
⊤	⊤	⊤	⊤	$Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$	1/16	1/10
⊤	⊤	⊤	⊥	$Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/20
⊤	⊤	⊥	⊤	$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/20
⊤	⊤	⊥	⊥	$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/20
⊤	⊥	⊤	⊤	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	1/16	1/20
⊤	⊥	⊤	⊥	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/10
⊤	⊥	⊥	⊤	$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/20
⊤	⊥	⊥	⊥	$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/20
⊥	⊤	⊤	⊤	${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$	1/16	1/20
⊥	⊤	⊤	⊥	${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/20
⊥	⊤	⊥	⊤	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/10
⊥	⊤	⊥	⊥	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/20
⊥	⊥	⊤	⊤	${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	1/16	1/20
⊥	⊥	⊤	⊥	${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/20
⊥	⊥	⊥	⊤	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/20
⊥	⊥	⊥	⊥	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/10

For any finite m and n, we can think of (the quantifier-free part of — I'll discuss quantified sentences at the end of these notes) $\mathcal{L}_Q^{m,n}$ as a propositional language with $m \times n$ atomic sentences. Now, a probability assignment over a language $\mathcal{L}$ is just an assignment of real numbers $\mathfrak{s}_i$ on $[0, 1]$ to each of its state descriptions s_i, such that $\sum_i \mathfrak{s}_i = 1$. This is analogous to a truth-value assignment, which assigns ⊤ (1) or ⊥ (0) to each atomic sentence of $\mathcal{L}$. Both truth-values and probability-values of sentences are "compositionally" determined. The truth-value of a (quantifier–free) sentence is determined by the truth-values of its atomic constituents. And, the probability of a (quantifier-free) sentence is determined by the probabilities of the state descriptions constituting what I will call its *state description equivalent* (SDE). Every (quantifier-free) sentence p of a language $\mathcal{L}_Q^{m,n}$ is equivalent to some disjunction of state descriptions $\bigvee s_i = p' \dashv\vDash p$ of $\mathcal{L}_Q^{m,n}$. And, the probability of p is simply the sum of the numbers $\mathfrak{s}_i$, corresponding to "basic probabilities" assigned to the state descriptions s_i constituting the disjunction p'. Of course, there are *many* such probability assignments. Carnap (and others like him) seem to think that *only some* of these probability assignments will count as "logical". Or, perhaps more accurately, Carnap *et al* believe that *only some* of these probabilty assignments are suitable for the construction of explicata for "confirmation". Which ones?

The first "logical" probability function $\mathfrak{c}^{\dagger}$ Carnap considers is one which is defined in terms of an underlying "logical measure function" $\mathfrak{m}^{\dagger}$, which assigns *equal probability to all state descriptions.* That is, for all i, $\mathfrak{s}_i = \mathfrak{m}^{\dagger}(s_i) = \frac{1}{2^{m \times n}}$, and then $\mathfrak{c}^{\dagger}(p \mid q) \overset{\text{def}}{=} \frac{\mathfrak{m}^{\dagger}(p \& q)}{\mathfrak{m}^{\dagger}(q)}$. Why $\mathfrak{m}^{\dagger}$? Carnap (and others) offer a rationale in terms of the "Principle of Indifference" (PI). But, for Carnap, the (PI) is supposed to be purely logical. As he explains:

> ...the statement of equiprobability to which the principle of indifference leads is, like all other statements of inductive probability, not a factual but a logical statement. If the knowledge of the observer does not favor any of the possible events, then with respect to this knowledge as evidence they are equiprobable.

This is confusing, since we're talking about a logical concept here. What does an "observer"'s "knowledge" have to do with anything here? Carnap's idea here seems to be that the (PI) is really a *conceptual/logical truth.* Here's the "proof" of (PI) — relative to "an observer's knowledge" K — Carnap seems to have in mind:

1. K does not favor any state description s_i over any state description s_j.
2. If K does not favor p over q or q over p, then $\Pr(p \mid K) = \Pr(q \mid K)$.
3. Therefore, for all i, j: $\Pr(s_i \mid K) = \Pr(s_j \mid K)$. This is just (PI) on the s_i's of $\mathcal{L}$, relative to K. *QED.*

This is cute, but not terribly helpful. First, "favoring" seems like an epistemic, not a logical concept. So, why should facts about "favoring" place constraints on logical probabilities? Perhaps we're already thinking of the probabilities here as an explicatum for "degree of credence" or "degree of belief", and this constraint comes from the requirement of similarity between explicandum and explicatum (more on this below)? Second, it is both question-begging and implausible to assume that the best explication of "no favoring" implies "same conditional probability". Popper gives examples that call this assumption into question. Take a case in which K is irrelevant to (independent of) both p and q, but $\Pr(p \mid K) > \Pr(q \mid K)$. It seems odd to say that K favors p over q (or q over p) in such a case. In an epistemic context, I'd say that p was more probable "a priori" (by which I just mean prior to the agent's having learned K, since I'm not sure what "a priori" probabilities are) than q, but this has nothing to do with K. Here, again, we have an ambiguity in the "favoring" explicandum (inherited from Popper's critique of Carnap's *LFP*). Third, notice how the case we need to establish here is not the case involving "an observer's knowledge" (think back to Maher's "unicorn example" which involved a similar slip), but rather the case involving "empty corpus" or "a priori knowledge" or something like this. And, it's not at all clear whether the above argument is compelling (or even makes much sense) for this case (recall I had a similar worry about Maher's "unicorn counterexample" to his $\text{NC}_{\top}$):

1. $K_{\top}$ does not favor any state description s_i over any state description s_j.
2. If $K_{\top}$ does not favor p over q or q over p, then $\Pr(p \mid \top) = \Pr(q \mid \top)$.
3. Therefore, for all i, j: $\Pr(s_i \mid \top) = \Pr(s_j \mid \top)$. This is just (PI) on the s_i's of $\mathcal{L}$, relative to $\top$ ($\mathfrak{m}^{\dagger}$). *QED.*

Again, to my ear, the argument for the case of "a priori" probabilities just sounds strange. Does it even make sense to talk about a tautology (or an *a priori* truth) "*favoring*" one hypothesis over another? The natural way of understanding "favoring" (as Popper does) is as a kind of *difference-making.* But, how *could* $K_{\top}$ (or $\top$) make any difference to the probabilities of p or q? And, *with respect to what background* would that "difference" be assessed? I will bracket these worries for now, because Carnap abandons $\mathfrak{c}^{\dagger}$ for other reasons anyway. Indeed, the reasons he offers for rejecting $\mathfrak{c}^{\dagger}$ make the considerations here *even stranger.*

2.1.2 Carnap's Second System: $\mathfrak{c}^{*}$

Carnap abandons $\mathfrak{c}^{\dagger}$ because — and these are his words — it doesn't allow for "learning from experience". Again, it seems that Carnap already has an epistemic explicandum in mind here, and the complaint seems to be about *dissimilarity* between explicatum and explicandum. Specifically, if (a) one thinks of $\mathfrak{c}^{\dagger}(p \mid q)$ as an explication of "rational degree of credence in p, conditional on q", and (b) one explicates *learning* in terms of *conditionalizing* one's credence function on what one has learned, then choosing $\mathfrak{c}^{\dagger}$ as our explicatum will have the consequence that (*e.g.*) learning Fa can never raise one's credence in Fb. Indeed, no conjunction of F's will raise the probability of another object's being F. Inspecting the $\mathfrak{m}^{\dagger}$ column of Table 1 yields:

$$\mathfrak{c}^\dagger(Fb \mid Fa) = \frac{\mathfrak{m}^\dagger(Fb \,\&\, Fa)}{\mathfrak{m}^\dagger(Fa)} = \frac{4 \cdot \frac{1}{16}}{8 \cdot \frac{1}{16}} = \frac{1}{2} = 8 \cdot \frac{1}{16} = \mathfrak{c}^\dagger(Fb \mid \top)$$

Carnap thought that Fb should count as "evidence in favor of" Fa (relative to "no background evidence"), and so he thinks this indicates a (fatal?) dissimilarity between explicandum and explicatum, if the explicatum is defined in terms of $\mathfrak{m}^\dagger$.[2] There is another (perhaps better?) way to view what Carnap is doing here. Perhaps Carnap thinks he's explicating a "logical" probability concept here alright, but he also thinks that there is an intimate connection between inductive logic and inductive epistemology. After all, he defends the RTE:

> **The Requirement of Total Evidence**. In the application of inductive logic to a given knowledge situation, the total evidence available must be taken as a basis for determining the degree of confirmation.

This imposes exactly the sort of constraint that would make this fact about $\mathfrak{c}^\dagger$ relevant to a "no learning from experience" complaint. Whichever way we understand what Carnap is up to here, one thing is certain — he moves from $\mathfrak{m}^\dagger$ to $\mathfrak{m}^*$ (also depicted in Table 1) because of this "no learning from experience problem". While $\mathfrak{m}^\dagger$ assigns equal probability to state descriptions, $\mathfrak{m}^*$ assigns equal probability to *structure descriptions*. A structure description is like a state description, but with "indistinguishable" individual constants. That is, if two state descriptions have the same structure — up to renaming of individual constants — then they are lumped together as falling under the same structure description. While there are 16 state descriptions of $\mathcal{L}_Q^{2,2}$, there are only 10 structure descriptions. If a structure description has multiple state descriptions falling under it, then each of those is assigned an equal share of the probability assigned to that structure description. This yields the $\mathfrak{m}^*$ function depicted in Table 1.[3] The resulting $\mathfrak{c}^*$ avoids the above problem:

$$\mathfrak{c}^*(Fb \mid Fa) = \frac{\mathfrak{m}^*(Fb \,\&\, Fa)}{\mathfrak{m}^*(Fa)} = \frac{2 \cdot \frac{1}{10} + 2 \cdot \frac{1}{20}}{2 \cdot \frac{1}{10} + 6 \cdot \frac{1}{20}} = \frac{3}{5} > \frac{1}{2} = 2 \cdot \frac{1}{10} + 6 \cdot \frac{1}{20} = \mathfrak{c}^*(Fb \mid \top)$$

So, $\mathfrak{c}^*$ allows for "learning from experience" (*i.e.*, it verifies "the principle of instantial relevance" or "basic singular predictive induction" — the relevance of Fb to Fa). But, it also assigns different probabilities to some pairs of state descriptions. Translating back into explicandum terms, does this mean that some state descriptions are "favored" over others by *a priori* evidence ($K_\top$)? If so, why would *that* be? And, how could such "favoring" jibe with Carnap's line on the "Principle of Indifference", above? In any event, Carnap soon moves beyond $\mathfrak{c}^*$ as well. A couple of years after *LFP*, he published a more general family of systems: $\mathfrak{c}_\lambda$.

2.2 The Johnson–Carnap λ–Continuum

Carnap published *The Continuum of Inductive Methods* in 1952 (two years after *LFP*). In this monograph, Carnap re-invents a family of "logical" probability functions that had already been invented in the teens by Keynes's teacher W.E. Johnson.[4] This family can be seen as a generalization of $\mathfrak{c}^\dagger$ and $\mathfrak{c}^*$. The details won't be important for us. But, the basic idea is that we now have a family of "logical measure functions" $\mathfrak{m}_\lambda$, where $\lambda \in [0, \infty]$. In this family of systems, $\mathfrak{m}_\infty = \mathfrak{m}^\dagger$, and $\mathfrak{m}_2 = \mathfrak{m}^*$. The parameter λ can be thought of as an "inverse speed of learning from experience" parameter. When $\lambda = \infty$, we get no learning from experience ($\mathfrak{m}^\dagger$), and when $\lambda = 2$, we get $\mathfrak{m}^*$. The move to a continuum makes Carnap's approach much more flexible. It can now emulate many more probability assignments. However, there are still many probability assignments that cannot be emulated in the λ-continuum, and some of these Carnap later came to view as indispensable for applications of inductive logic. This led to the addition of yet another parameter to Carnap's systems.

2.3 The Later Carnapian λ/γ–Continua

In his later work, Carnap became convinced that sensitivity to what he called "analogical" effects should be a desideratum for "logical" $\Pr_\top$. For instance, he thought the following "analogical" constraint should hold:

[2] Note the emphasis on *singular predictive induction*. Early Carnapian theory is *Hempelian* on $\forall$-induction! ($\text{NC}_\top$) is a *theorem* of Carnap's early systems (on *finite* domains—see §2.4): $\Pr[(Fa \supset Ga) \,\&\, (Fb \supset Gb) \mid Fa \,\&\, Ga] > \Pr[(Fa \supset Ga) \,\&\, (Fb \supset Gb) \mid \top]$. In fact, Carnap's early (λ-family) $\mathfrak{C}$-theories agree with Hempelian theory on "confirmatory instances" of $\forall$-claims. *E.g.*, $Fa \,\&\, Ga$, $\sim Fa \,\&\, \sim Ga$, $\sim Fa \,\&\, Ga$, $\sim Fa$, and Ga *all* confirm $(\forall x)(Fx \supset Gx)$. Check this for $\mathfrak{c}^\dagger$ and $\mathfrak{c}^*$, over $\mathcal{L}_Q^{2,2}$ — see my *Mathematica* (PrSAT) notebook.

[3] This distinction between treating individual constants as distinguishable ($\mathfrak{m}^\dagger$) or indistinguishable ($\mathfrak{m}^*$) is analogous to the difference between Maxwell-Boltzmann and bose-einstein/fermi-dirac statistics in classical *vs* quantum statistical mechanics.

[4] Carnap doesn't cite Johnson, but he had read Johnson's work in 1921. I discovered this in the Carnap Archive at Pitt a few years ago. According to Dick Jeffrey and Sandy Zabell, the most likely explanation is that Carnap *forgot* he had read Johnson in 1921.

$$\Pr_{\top}(Gb \mid Ga) > \Pr_{\top}(Gb \mid Ga \,\&\, Fa \,\&\, {\sim}Fb) > \Pr_{\top}(Gb \mid \top)$$

This implies ("a priori") instantial relevance of Ga to Gb, plus the additional "analogical" requirement that if a and b are also known to differ on some other property F, then this should undermine — but not completely eliminate — the degree of ("a priori") instantial relevance of Ga to Gb. As it turns out, no assignment in the λ-continuum can satisfy this additional "analogical" constraint. As an exercise, you should convince yourself (using the values in Table 1) that neither $\mathfrak{c}^{\dagger}$ nor $\mathfrak{c}^{*}$ satisfies the above "analogical" constraint.

Because of this insensitivity to "analogical effects", Carnap later added yet another parameter γ to his systems. The details aren't important here, but if you look at Maher's paper "Probability Captures the Logic of Scientific Confirmation", then you'll see a Carnapian two-parameter system for the class of languages $\mathcal{L}_Q^{n,2}$ (with two predicates), which satisfies these additional later Carnapian constraints (and others).[5]

Unfortunately, there is another problem that arises, when we start thinking about "analogical effects" in the way that Carnap did. As soon as we go beyond languages with two predicates, we run headlong into a thorny *language-relativity problem*. In his later work, Carnap generalizes the "analogical" requirement. He defines a syntactical notion of "similarity" of two objects a and b: $\mathbb{S}(a,b)$, which is defined as "the number of predicates that both a and b fall under". Unfortunately, as soon as our languages contain more than two predicates, $\mathbb{S}(a,b)$ becomes language-relative in a pernicious way. Consider two $\mathcal{L}_Q^{2,4}$ languages: the $ABCD$ language and the $XYZU$ language. The $ABCD$ language consists of four predicates A, B, C, and D. And, the $XYZU$ language also has four predicates X, Y, Z, and U such that Xx is (extra-systematically) equivalent to $Ax \equiv Bx$, Yx is equivalent to $Bx \equiv Cx$, Zx is equivalent to Ax, and Ux is equivalent to Dx.[6] Thus, $ABCD$ and $XYZU$ are *expressively equivalent* languages. Now, consider two objects a and b such that:

$Aa \,\&\, Ba \,\&\, Ca \,\&\, Da$		$Xa \,\&\, Ya \,\&\, Za \,\&\, Ua$
and	or, expressed in $XYZU$ terms:	and
$Ab \& {\sim}Bb \& Cb \& Db$		${\sim}Xb \& {\sim}Yb \& Zb \& Ub$

How similar are a and b in a "predicate-sharing" sense? This depends on which language we use to describe a and b. In $ABCD$, a and b share three predicates. But, in $XYZU$, a and b share only two predicates, *i.e.*:

$$\mathbb{S}_{ABCD}(a,b) = 3 \neq 2 = \mathbb{S}_{XYZU}(a,b)$$

Using $\mathbb{S}(a,b)$, we can state Carnap's generalized "analogical" constraint, as follows:

($\mathscr{A}$) If $n > m$, then $\Pr_{\top}(Xa \mid Xb \,\&\, \mathbb{S}(a,b) = n) > \Pr_{\top}(Xa \mid Xb \,\&\, \mathbb{S}(a,b) = m)$.[7]

If we apply this to a and b in the $ABCD$ and $XYZU$ languages, respectively, we get (for instance):

(1) $\Pr_{\top}[Da \mid Db \,\&\, (Aa \,\&\, Ba \,\&\, Ca \,\&\, Ab \,\&\, {\sim}Bb \,\&\, Cb)] > \Pr_{\top}[Da \mid Db \,\&\, (Aa \,\&\, Ba \,\&\, Ca \,\&\, Ab \,\&\, {\sim}Bb \,\&\, {\sim}Cb)]$

(2) $\Pr_{\top}[Ua \mid Ub \,\&\, (Xa \,\&\, Ya \,\&\, Za \,\&\, {\sim}Xb \,\&\, Yb \,\&\, Zb)] > \Pr_{\top}[Ua \mid Ub \,\&\, (Xa \,\&\, Ya \,\&\, Za \,\&\, {\sim}Xb \,\&\, {\sim}Yb \,\&\, Zb)]$

But, (1) and (2) violate the following *language-invariance* desideratum ($\Leftrightarrow$ is *extrasystematic equivalence*):

($\mathcal{I}$) If $p \Leftrightarrow r$ and $q \Leftrightarrow s$, then $\Pr_{\top}(p \mid q) = \Pr_{\top}(r \mid s)$.

Let $p \overset{\text{def}}{=} Da$, $r \overset{\text{def}}{=} Ua$, $q \overset{\text{def}}{=} Db\&Aa\&Ba\&Ca\&Ab\&{\sim}Bb\&Cb$, and $s \overset{\text{def}}{=} Ub\&Xa\&Ya\&Za\&{\sim}Xb\&{\sim}Yb\&Zb$. Then, ($\mathcal{I}$) implies that the term on the left-hand side of (1) must be equal to the term of the right-hand side of (2). Similarly, if we let $p \overset{\text{def}}{=} Ua$, $r \overset{\text{def}}{=} Da$, $q \overset{\text{def}}{=} Ub\&Xa\&Ya\&Za\&{\sim}Xb\&Yb\&Zb$, and $s \overset{\text{def}}{=} Da\&Db\&Aa\&Ba\&Ca\&Ab \,\&\, {\sim}Bb \,\&\, {\sim}Cb$, then ($\mathcal{I}$) implies that the term on the left-hand side of (2) must be equal to the term of the right-hand side of (1). But, then, (1) and (2) contradict the law of trichotomy for real numbers. Thus, Carnap's desideratum ($\mathscr{A}$) introduces a strong kind of language-relativity into his explicata. Is this a serious problem? It's not immediately clear. Just as Tarski's explicatum "truth-for-a-Tarski-interpretation-of-$\mathcal{L}$" is language-relative, so is Carnap's explicatum "inductive-conditional-probability-for-a-Carnap-intepretation-of-$\mathcal{L}$". Of course, one might still wonder whether Carnap's *explicanda* are (intuitively) language-relative. But, one could have the same worry about Tarskian "truth", no? Be that as it may, Tarski needs to specify an *assignment*

[5] Unlike any functions in the λ-continuum, Maher's two-parameter Carnapian system contains $\mathfrak{c}$-functions that force *violations* of ($\mathrm{NC}_{\top}$) on finite domains. Maher noticed this in his (2004). I now suspect this is a *mere artifact* of "instantial relevance" + "analogy".

[6] This example is a modification of an example David Miller used (in 1973) to show that Popper's qualitative theory of verisimilitude is language-relative. As far as I know, I'm the first to apply Miller's argument to Carnapian probabilities (for 3+ predicates).

[7] ⌜$\mathbb{S}(a,b) = n$⌝ is shorthand for ⌜a sentence α of $\mathcal{L}$ which gives a maximally specific description of both a and b in terms of the predicates of $\mathcal{L}$ (other than 'X') such that the number of predicates of $\mathcal{L}$ that apply to both a and b, according to α, is equal to n⌝.

function for $\mathcal{L}$, in order to generate truth-values-for-a-Tarski-interpretation-of-$\mathcal{L}$. And, Carnap needs to specify a *logical measure function* $\mathfrak{m}$ for $\mathcal{L}$, in order to generate probabilities-for-a-Carnap-interpretation-of-$\mathcal{L}$. At this level of abstraction, the two explications sound rather similar. How do they differ?

Well, for one thing, we can *use* Tarski's truth-for-a-Tarski-interpretation-of-$\mathcal{L}$ to explicate *logical consequence* (in $\mathcal{L}$). And, logical consequence is *not* language-relative (at least, not in the strong sense we saw above with Carnapian probabilities). On the other hand, if we *use* Carnap's probability-for-a-Carnap-interpretation-of-$\mathcal{L}$ to explicate *confirmation* (as increase in firmness, let's say, like Maher does), then it will inherit the language-relativity problem (except in the deductive cases themselves, of course). This is an important disanalogy, since one of the main desiderata here is that $\mathfrak{c}$ be a *generalization* of entailment. But, if $\mathfrak{c}$ is only language-invariant in the deductive limiting-cases, then this is *not much of a generalization*.[8]

There are further problems with Carnap's project, once we move from two to three+ predicates. Maher (in his paper "Probabilities for Multiple Properties", now on the website) gives *independent* reasons to worry about all of the existing Carnapian systems in the case of three or more predicates. Indeed, Maher does not know of any adequate Carnapian systems of this kind — and for reasons that have nothing to do with language-relativity. So, it seems that Carnap's project has failed — even by the lights of staunch Carnapians.

2.4 Quantified Sentences in Carnapian Systems

So far, we've been (implicitly) assuming that we're talking about *finite* domains of discourse. There are several reasons for this. First, only finite models are needed for the model theory of monadic predicate logic (if all we care about are *narrowly logical* properties). Second, because if we allow infinite domains of discourse, then we run into problems involving universally quantified sentences. Carnap (and others in this tradition) handles a $\forall$-quantified sentence $(\forall x)\phi x$ by looking at *the conjunction of its instances* $\bigwedge \phi$. Of course, this requires taking an "extra-systematic" perspective that "knows" there are exactly m objects in the domain of discourse, which are named by the m individual constants in $\mathcal{L}_Q^{m,n}$. After all, $\bigwedge \phi \nvDash_{\mathcal{L}} (\forall x)\phi x$. Let's bracket that issue. A "systematic" problem for Carnap's early systems was that *all universal claims had probability zero on all infinite domains*. There are technical fixes for this problem, which were implemented by Hintikka, Scott and Krauss, and others. Carnap later adopted some of these amendments. Since I have bigger (or at least different and smellier to my nose) fish to fry here, I won't discuss this issue further.

2.5 Some Closing Reflections on Carnap's Project

In the beginning, Carnap thought there was a *unique* "logical" probability function for any given language. Later, he moved to a family of probability functions, parameterized by a real-valued λ. Later still, he moved to a 2-parameter λ/γ-family. This seems odd, since, on its face, it introduces *logical indeterminacies* into a concept that is supposed to be logically determinate. Indeed, most commentators see this as Carnap *abandoning* his original "logical" project (since he talks about "pragmatic ways" of fixing the parameter values, *etc.*). Moreover, it seems that even the λ/γ systems are over-constrained in ways that *even Carnapians* think are too restrictive (esp. in the case of 3+ predicates). This seems to indicate the need for *still further parameters*. I think this is a degenerating research programme. Basically, I think the whole thing is based on a few mistakes (*e.g.*, that there are *a priori* probabilities, and/or that *pure* inductive logic owes us a "logical interpretation" of Pr). I recommend that we think of confirmation (in its logical sense) as a relation between E, H, K *and a probability model* $\mathfrak{M}$, which includes (at least) E, H, K *and a basic probability asignment* Pr over the state descriptions determined by E, H, and K (and whatever other statements might occur in $\mathfrak{M}$). In this way, the inductive logician is free to use *any* probability assignment (just as the deductive logician is free to use *any* assignment function). Two key questions arise now: "What remains of inductive *logic*?" and "What is the relationship between inductive logic and inductive epistemology?". I'll say more about these questions later. But, on the first question, there remains the task of characterizing the *comparative* (and perhaps *quantitative*) *structure* of the confirmation-as-increase-in-firnmess relation. This is non-trivial — *even when relativized to a probability model* $\mathfrak{M}$. On the second question, the problem of "bridge principles" between logic and epistemology must first be addressed in the deductive limting-cases (which is already hard!). As it turns out, addressing *that* problem also helps with the first problem. More on that later.

[8] See Fine's "Logical (Conditional) Probability" (now on website) for a different way of seeing where the language-relativity comes into the Carnapian systems. Fine gives an axiomatic development involving various symmetry and permutation invariance principles.