Philosophy 148 — Announcements & Such

• Overall, people did very well on the mid-term ($\mu = 90, \sigma = 16$).

• HW #2 graded will be posted very soon. Raul won’t be able to give back the HW’s until next week, and I will collect HW #3’s for him today.

• Recall: we’re using a straight grading scale for this course. Here it is:
  
  – A+ > 97, A (94,97], A- (90,94], B+ (87,90], B (84,87], B- (80,84],
  
  C+ (77,80], C (74,77], C- (70,74], D [50,70], F < 50.

• I’ve posted solutions for HW #1 (HW #2 solutions are coming soon).

• HW #3 is due today. HW #4 has been posted. We will have a discussion devoted to it in two weeks time. Next week, I’m in Europe — so no office hours for me. We’ll have guest lecturers — so the class will meet!

• Today: More on “Logical” Probability and Inductive Logic
  
  – Review of Carnapian theory, and some new stuff on Carnap.
  
  – Another way of thinking about Inductive Logic.
Review of Carnapian “Logical” Probability 1

- Carnap endorses the **logicality desideratum** \((D_2)\) for \(c\), and since Carnap defines \(c\) in terms of \(Pr\), Carnap concludes this mandates a “logical” kind (theory/interpretation) of probability — the search for “logical \(Pr\)” is on.

- Carnap assumes that \(c(C, P) = Pr(C \mid P)\). The principle of indifference (PI) follows from this assumption, since then “\(K\) does not favor any \(s_i\) over any \(s_j\)” \(\Rightarrow \) “\(K\) confirms each \(s_i\) to the same degree as \(K\) confirms each \(s_j\)” \(\Rightarrow c(s_i, K) = c(s_j, K) \Rightarrow Pr(s_i \mid K) = Pr(s_j \mid K)\), for all \(i\) and \(j\), *ergo* (PI).

- This leads Carnap, initially, to endorse the Wittgensteinian function \(m^\dagger\), which assigns equal probability to each state description \(s_i\) of \(L\).

- But, \(m^\dagger\) implies that there can be no correlations between logically independent propositions. Carnap thinks this renders \(m^\dagger\) *inapplicable*.

- Here, Carnap is presupposing what I will call an **applicability desideratum**: \((D_3)\) Inductive logic (\(i.e.,\) the theory of confirmation) should be *applicable* (presumably, to *epistemology*) — in some substantive way.
Review of Carnapian “Logical” Probability 2

• *Exactly how* \( c \) (i.e., \( \Pr_T \)) should be applicable to epistemology is not completely clear. But, Carnap says things which *constrain* applicability.
  
  – It should be possible for logically independent claims to be *correlated* with each other. Presumably, because it is possible for logically independent claims to *evidentially support* each other.
  
  – Specifically, Carnap thinks \( Ga \) and \( Gb \) should be correlated “*a priori*”:

\[
\Pr(Gb \mid Ga) > \Pr(Gb) \quad \text{[this is called “instantial relevance”]}
\]

  – This *rules-out* \( m^\dagger \). And, this leads Carnap to adopt \( m^* \) instead. But, \( m^* \) has “inapplicability problems” of its own. Carnap later came to think we should also be able to have the following chain of inequalities:

\[
\Pr(Gb \mid Ga) > \Pr(Gb \mid Ga \& Fa \& \sim Fb) > \Pr(Gb)
\]

  – But (HW #3!), *neither* \( m^\dagger \) *nor* \( m^* \) are compatible with this. This lead Carnap to continue his search. He moved to a more complex \( m^{\lambda,\gamma} \).
Carnap’s Final Theory of “Logical” Probability ($m, n = 2, \lambda = \frac{1}{2}$)

<table>
<thead>
<tr>
<th>$F_a$</th>
<th>$G_a$</th>
<th>$F_b$</th>
<th>$G_b$</th>
<th>Carnap’s $m^{\frac{1}{2}, \gamma(s_i)}$ [where $\gamma_F, \gamma_G \in (0, 1)$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>$\frac{1}{9} \gamma_F \gamma_G (\gamma_G + \gamma_F (5 \gamma_G + 1) + 2)$</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>$-\frac{1}{9} \gamma_F (5 \gamma_F + 1) (\gamma_G - 1) \gamma_G$</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>$-\frac{1}{9} (\gamma_F - 1) \gamma_F \gamma_G (5 \gamma_G + 1)$</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>$\frac{5}{9} (\gamma_F - 1) \gamma_F (\gamma_G - 1) \gamma_G$</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>$-\frac{1}{9} \gamma_F (5 \gamma_F + 1) (\gamma_G - 1) \gamma_G$</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>$\frac{1}{9} \gamma_F (\gamma_G - 1) (\gamma_G + \gamma_F (5 \gamma_G - 6) - 3)$</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>$\frac{5}{9} (\gamma_F - 1) \gamma_F (\gamma_G - 1) \gamma_G$</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>$-\frac{1}{9} (\gamma_F - 1) \gamma_F (\gamma_G - 1) (5 \gamma_G - 6)$</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>$-\frac{1}{9} (\gamma_F - 1) \gamma_F \gamma_G (5 \gamma_G + 1)$</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>$\frac{5}{9} (\gamma_F - 1) \gamma_F (\gamma_G - 1) \gamma_G$</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>$\frac{1}{9} (\gamma_F - 1) \gamma_G (\gamma_F + (5 \gamma_F - 6) \gamma_G - 3)$</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>$-\frac{1}{9} (\gamma_F - 1) (5 \gamma_F - 6) (\gamma_G - 1) \gamma_G$</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>$\frac{5}{9} (\gamma_F - 1) \gamma_F (\gamma_G - 1) \gamma_G$</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>$-\frac{1}{9} (\gamma_F - 1) \gamma_F (\gamma_G - 1) (5 \gamma_G - 6)$</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>$-\frac{1}{9} (\gamma_F - 1) (5 \gamma_F - 6) (\gamma_G - 1) \gamma_G$</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>$\frac{1}{9} (\gamma_F - 1) (\gamma_G - 1) (-6 \gamma_G + \gamma_F (5 \gamma_G - 6) + 9)$</td>
</tr>
</tbody>
</table>
Carnapian Logical Probability: Analogy and Similarity 1

- More generally: if \( a \) has properties \( P_1 \ldots P_n \), and \( b \) has \( P_1 \ldots P_{n-2} \), but lacks \( P_{n-1} \), that should be of *some* relevance to \( b \)'s having \( P_n \); \( n = 3 \) case:

\[
\Pr(H_b \mid H_a) > \Pr(H_b \mid H_a \land F_a \land G_a \land F_b \land \neg G_b)
\]

\[
> \Pr(H_b \mid H_a \land F_a \land G_a \land \neg F_b \land \neg G_b) > \Pr(H_b)
\]

- *I.e.*, Differing on 2 properties should be worse than 1, but neither should completely undermine instantaial relevance. \( \Pr^\dagger \) and \( \Pr^* \) violate this.

- The “analogical” idea here involves certain judgments of “similarity” of objects, where “similarity” is measured by “counting shared predicates.”

- The slogan seems to be: The more properties \( a \) and \( b \) “share”, the more this *enhances* instansial relevance; and, the more properties they are known to *differ* on, the more this *undermines* instansial relevance.

- Unfortunately, any theory that satisfies this analogical principle will be *language-variant*, so long as the languages contain 3 or more predicates.
Carnapian Logical Probability: Analogy and Similarity 2

Let $x_1, \ldots, x_n$ be the objects that fall under the predicate $X$ [i.e., those in $\text{Ext}(X)$]. And, let $s(x_1, \ldots, x_n)$ be some measure of “the degree to which the objects falling under $X$ are similar to each other”.

Carnap doesn’t offer much in the way of a theory of $s(x_1, \ldots, x_n)$. But, his discussion suggests the following account of $s(x_1, \ldots, x_n)$:

Let $P(x)$ be the set of predicates that $x$ falls under. Then, define:

$$s(x_1, \ldots, x_n) = \left| \bigcap_i P(x_i) \right|$$

That is, $s(x_1, \ldots, x_n)$ is the size (cardinality) of the intersection of all the $P(x_i)$. This is “the size of the set of shared predicates of the $x_i$”.

There is a problem with this idea. Next, I will present an argument which shows that this measure of similarity is language variant.
Carnapian Logical Probability: Analogy and Similarity 3

- That is, “the degree of similarity of $a$ and $b$” depends sensitively on the syntax of the language one uses to describe $a$ and $b$. [Note: if $n = 2$, there is no language-variance — it requires 3 or more predicates.] Here’s why.

- 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 $Xx \equiv A x \equiv B x$, $Yx \equiv B x \equiv C x$, $Zx \equiv A x$, and $Ux \equiv D x$.

$ABCD$ and $XYZU$ are (extra-systematically) expressively equivalent. Anything that can be said in $ABCD$ can be said in $XYZU$, and conversely — intuitively, there is no semantic difference between the two languages.

- Now, consider two objects $a$ and $b$ such that:

  $Aa \& Ba \& Ca \& Da$

  $Ab \& \sim Bb \& Cb \& Db$

- Question: How similar are $a$ and $b$ in our “predicate-sharing” sense?
Carnapian Logical Probability: Analogy and Similarity 4

• Answer: That depends on which of our expressively equivalent languages we use to describe \( a \) and \( b \)! To see this, note that in \( XYZU \) we have:

\[
Xa \land Ya \land Za \land Ua \\
\sim Xb \land \sim Yb \land Zb \land Ub
\]

• Therefore, in \( ABCD \), \( a \) and \( b \) share three predicates. But, in \( XYZU \), \( a \) and \( b \) share only two predicates. Or, to use a modified notation, we have:

\[
s_{ABCD}(a, b) = 3 \neq 2 = s_{XYZU}(a, b)
\]

• On the other hand, probabilities should not be language-variant. It shouldn’t matter which language you use to describe the world — equivalent statements should be probabilistically indistinguishable.

• One consequence is that if \( p \equiv \equiv q, x \equiv \equiv y \) and \( z \equiv \equiv u \), then we shouldn’t have both \( \Pr(p \mid x) > \Pr(p \mid u) \) and \( \Pr(q \mid y) < \Pr(q \mid z) \). Carnapian principles of analogy and similarity contradict this requirement \((n \geq 3)\).
Carnapian Logical Probability: Analogy and Similarity

• Here’s a concise way of stating the general Carnapian analogical principle:

  (A) If \( n > m \), then \( \Pr(Xa \mid Xb \& s(a, b) = n) > \Pr(Xa \mid Xb \& s(a, b) = m) \)

• Applying this principle to our example yields both of the following:

  (1) \( \Pr(Da \mid Db \& Aa \& Ba \& Ca \& Ab \& \sim Bb \& Cb) > \Pr(Da \mid Db \& Aa \& Ba \& Ca \& Ab \& \sim Bb \& \sim Cb) \)
  (2) \( \Pr(Ua \mid Ub \& Xa \& Ya \& Za \& \sim Xb \& Yb \& Zb) > \Pr(Ua \mid Ub \& Xa \& Ya \& Za \& \sim Xb \& \sim Yb \& Zb) \)

• Now, let \( p \equiv Da, q \equiv Ua \), and

  \[
  \begin{array}{|c|c|}
  \hline
  x \equiv Db \& Aa \& Ba \& Ca \& Ab \& \sim Bb \& Cb & y \equiv Ub \& Xa \& Ya \& Za \& \sim Xb \& \sim Yb \& Zb \\
  z \equiv Ub \& Xa \& Ya \& Za \& \sim Xb \& Yb \& Zb & u \equiv Db \& Aa \& Ba \& Ca \& Ab \& \sim Bb \& \sim Cb \\
  \hline
  \end{array}
  \]

• Then, \( p \models q, x \models y \) and \( z \models u \), but the Carnapian principle (A) implies both (1) \( \Pr(p \mid x) > \Pr(p \mid u) \), and (2) \( \Pr(q \mid y) < \Pr(q \mid z) \). Bad.

• It seems that principle (A) must go. Otherwise, some restriction on the choice of language is required to block inferring both (1) and (2) from it.
A Closer Look at Inductive Logic and Applicability 1

- Carnap suggested various *bridge principles* for connecting inductive logic and inductive epistemology. The most well-known of these was:
  - **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 evidential support.

- A more precise way of putting this principle is:
  \[ (\text{RTE}) \quad E \text{ evidentially supports } H \text{ for an agent } S \text{ in an epistemic context } C \iff c(H, E | K) > r, \text{ where } K \text{ is } S\text{'s total evidence in } C. \]

- For Carnap, \( c(H, E | K) = \Pr_T(H | E & K) \), where \( \Pr_T \) is a suitable “logical” probability function. So, we can restate Carnap’s (RTE) as follows:
  \[ (\text{RTE}_C) \quad E \text{ evidentially supports } H \text{ for an agent } S \text{ in an epistemic context } C \iff \Pr_T(H | E & K) > r, \text{ where } K \text{ is } S\text{'s total evidence in } C. \]

- Carnap’s version of (RTE) faces a challenge (first articulated by Popper) involving *probabilistic relevance vs high conditional probability.*
A Closer Look at Inductive Logic and Applicability 2

- Popper discusses examples of the following kind (we’ve seen an example like this in the class before), which involve testing for a rare disease.
  - Let $E$ report a positive test result for a very rare disease (for someone named John), and let $H$ be the (null) hypothesis that John does not have the disease in question. We assume further that John knows (his $K$ entails) the test is highly reliable, and that the disease is very rare.

- In such an example, it is plausible (and Carnap should agree) that (to the extent that $\Pr_T$ is applicable to modeling the epistemic relations here):
  1. $\Pr_T(H \mid E & K)$ is very high.
  2. But, $\Pr_T(H \mid E & K) < \Pr_T(H \mid K)$.

- Because of (2), it would be odd to say that $E$ supports $H$ (for John) in this context. (2) suggests that $E$ is (intuitively) evidence against $H$ here.

- But, because of (1), Carnap’s $(\text{RTE}_C)$ implies that $E$ supports $H$ (for John) here. This looks like a counterexample to [the $\iff$ of] Carnap’s $(\text{RTE}_C)$.
This suggests the following refinement of Carnap’s (RTEC):

(RTE_C)  
$E$ evidentially supports $H$ for an agent $S$ in an epistemic context $C$

$\implies \Pr_T(H \mid E \& K) > \Pr_T(H \mid K)$, where $K$ is $S$’s total evidence in $C$.

In other words, (RTE_C) says that evidential support in (for $S$ in $C$) implies probabilistic relevance, conditional upon $K$ (for a suitable $\Pr_T$ function).

Note: this only states a necessary condition for evidential support.

While (RTE_C) avoids Popper’s objection, it faces serious challenges of its own (e.g., Goodman’s “Grue” example — more on that later). Here’s one:

Consider any context in which $S$ already knows (with certainty) that $E$ is true. That is, $S$’s total evidence in the context entails $E$ ($K \models E$).

In such a case, $\Pr(H \mid E \& K) = \Pr(H \mid K)$, for any probability function $\Pr$. Thus, (RTE_C) implies that $E$ cannot support anything (for $S$, in any such $C$). This shows that (RTE_C) isn’t a correct principle either. [“Old Evidence”]
A Closer Look at Inductive Logic and Applicability 4

• I think this whole way of approaching inductive logic is wrongheaded.

• First, why must Pr itself be logical, if c (which is defined in terms of Pr) is to be logical? Analogy: must the truth-value assignment function v itself be logical, if ⊨ (which is defined in terms of v) is to be logical?
  – But: there is a crucial disanalogy here, which I will discuss below.

• Second, Carnap’s proposal \( c(H, E \mid K) = Pr_T(H \mid E \& K) \) is suspect, because (as Popper pointed out) it is not sensitive to probabilistic relevance.
  – Note: this undermines Carnap’s argument for the “logicality” of (PI).

• Third, the applicability desideratum \( (D_3) \) may be fundamentally misguided. The search for logic/epistemology “bridge principles” is fraught with danger, even in the deductive case. And, since IL is supposed to generalize DL, it will also face these dangers and new ones (as above).
  – I think this is the true (but, surprisingly, un-appreciated) lesson of Goodman’s “grue” example. I will explain why in the confirmation unit.
An Alternative Conception of Inductive Logic 1

- In light of the above considerations, we might seek a measure $c$ satisfying the following (provided that $E$, $H$, and $K$ are logically contingent):

$$c(H, E \mid K) \text{ should be } \begin{cases} \text{Maximal (> 0, constant)} & \iff E \& K \models (\text{or } \models) H. \\ > 0 \text{ (confirmation rel. to } K) & \implies \Pr(H \mid E \& K) > \Pr(H \mid K). \\ = 0 \text{ (irrelevance rel. to } K) & \implies \Pr(H \mid E \& K) = \Pr(H \mid K). \\ < 0 \text{ (disconfirmation rel. to } K) & \implies \Pr(H \mid E \& K) < \Pr(H \mid K). \\ \text{Minimal (< 0, constant)} & \iff E \& K \models (\text{or } \models) \sim H. \end{cases}$$

- Carnap would add: “and $\Pr$ should be a ‘logical’ probability function $\Pr_T$”. But, I suggested that this was a mistake. OK, but then what do I say about the $\Pr$’s above? There is an implicit quantifier over the $\Pr$’s above…

  - $\exists$ is too weak a quantifier here, since there will always be some such $\Pr$.
  - $\forall$ is too strong a quantifier here, because that is demonstrably false!
  - What’s the alternative? The alternative is that $\Pr$ is a parameter in $c$ itself. That is, perhaps $\Pr$ is simply an argument of the function $c$. 
An Alternative Conception of Inductive Logic 2

- Here’s the idea. Confirmation is a four-place relation, between $E$, $H$, $K$, and a probability function $Pr$. The resulting relation is still logical in Carnap’s sense, since, given a choice of $Pr$, $c$ is logically (mathematically, if you prefer) determined, provided only that $c$ is defined in terms of $Pr$.
- So, on this conception, desiderata $(D_1)$ and $(D_2)$ are satisfied.
- As usual, the subtle questions involve the applicability desideratum $(D_3)$.
- What do we say about that? Well, I think any naive “bridge principle” like (RTE or RTE') is doomed to failure. But, perhaps there is some connection.
- Thinking back to the deductive case, there may be some connection between deductive logic and deductive inference. But, what is it?
- This is notoriously difficult to say. The best hope seems to be that there is some connection between knowledge and entailment. [Note: connecting or bridging justified belief and entailment seems much more difficult.]
Many people think there is some connection between knowledge and entailment. But, simple/naive version of bridge principles don’t work. Here is a progression of increasingly subtle “bridge principles”:

1. If $S$ knows $p$ and $p \models q$, then $S$ knows $q$.
   - What if $S$ doesn’t know that $p \models q$?

2. If $S$ knows $p$ and $S$ knows $p \models q$, then $S$ knows $q$.
   - What if $S$ doesn’t even believe $q$? [Discuss Dretske’s Zebra Example.]

3. If $S$ knows $p$ and $S$ knows $p \models q$ and $S$ believes $q$, then $S$ knows $q$.
   - What if $S$ believes $q$ for reasons other than $p$?

4. If $S$ knows $p$ and $S$ knows $p \models q$ and $S$ comes to believe $q$ because of their (initial) belief that $p$, then $S$ knows $q$.
   - What if $S$ no longer believes $p$, while/after they infer $q$?

5. If $S$ knows $p$ & $S$ knows $p \models q$ & $S$ competently deduces $q$ from $p$ (thus coming to believe $q$) while maintaining their belief in $p$, then $S$ knows $q$. 