Philosophy 148 — Announcements & Such

- Administrative Stuff
  - I'll be 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.
  - People did very well on the quiz ($\mu = 93$). HW #1 assigned (due 2/28).
  - Today's Agenda
    - Some "real world" probability examples (and problems with them)
    - Then, starting over from scratch — with a guiding analogy:

$$\frac{\text{truth-on-}\mathcal{I}}{\text{truth}} :: \frac{\text{probability-in-}\mathcal{M}}{\text{probability}}$$

  - That is, we'll start again (from scratch) by comparing the informal notions of "truth" and "probability", and their analogue *formal* or *analytic* notions "truth-on-$\mathcal{I}$" and "probability-in-$\mathcal{M}$".
  - This will give us a "bottom-up" approach for the rest of the course.

Inverse Probability and Bayes's Theorem II

- Here's a famous example, illustrating the subtlety of Bayes's Theorem:

  The (unconditional) probability of breast cancer is 1% for a woman at age forty who participates in routine screening. The probability of such a woman having a positive mammogram, given that she has breast cancer, is 80%. The probability of such a woman having a positive mammogram, given that she does not have breast cancer, is 10%. What is the probability that such a woman has breast cancer, given that she has had a positive mammogram in routine screening?

- We can formalize this, as follows. Let H = such a woman (age 40 who participates in routine screening) has breast cancer, and E = such a woman has had a positive mammogram in routine screening. Then:

$$\Pr(E \mid H) = 0.8, \Pr(E \mid {\sim}H) = 0.1, \text{ and } \Pr(H) = 0.01.$$

- Question (like Hacking's O.Q. #5): What is $\Pr(H \mid E)$? What would you guess? Most experts guess a pretty high number (near 0.8, usually).

- If we apply Bayes's Theorem, we get the following answer:

$$\Pr(H \mid E) = \frac{\Pr(E \mid H) \cdot \Pr(H)}{\Pr(E \mid H) \cdot \Pr(H) + \Pr(E \mid {\sim}H) \cdot \Pr({\sim}H)}$$
$$= \frac{0.8 \cdot 0.01}{0.8 \cdot 0.01 + 0.1 \cdot 0.99} \approx 0.075$$

- We can also use our algebraic technique to compute an answer.

E	H	$\Pr(s_i)$
T	T	$a_1 = 0.008$
T	F	$a_2 = 0.099$
F	T	$a_3 = 0.002$
F	F	$a_4 = 0.891$

$$\Pr(E \mid H) = \frac{\Pr(E \,\&\, H)}{\Pr(H)} = \frac{a_1}{a_1 + a_3} = 0.8$$
$$\Pr(E \mid {\sim}H) = \frac{\Pr(E \,\&\, {\sim}H)}{\Pr({\sim}H)} = \frac{a_2}{1 - (a_1 + a_3)} = 0.1$$
$$\Pr(H) = a_1 + a_3 = 0.01$$

- Note: The posterior is about eight times the prior in this case, but since the prior is *so* low to begin with, the posterior is still pretty low.

- This mistake is usually called the *base rate fallacy*. I will return to this example later in the course, and ask whether it really is a mistake to report a large number in this example. Perhaps it is not a mistake.

Inverse Probability and Bayes's Theorem III

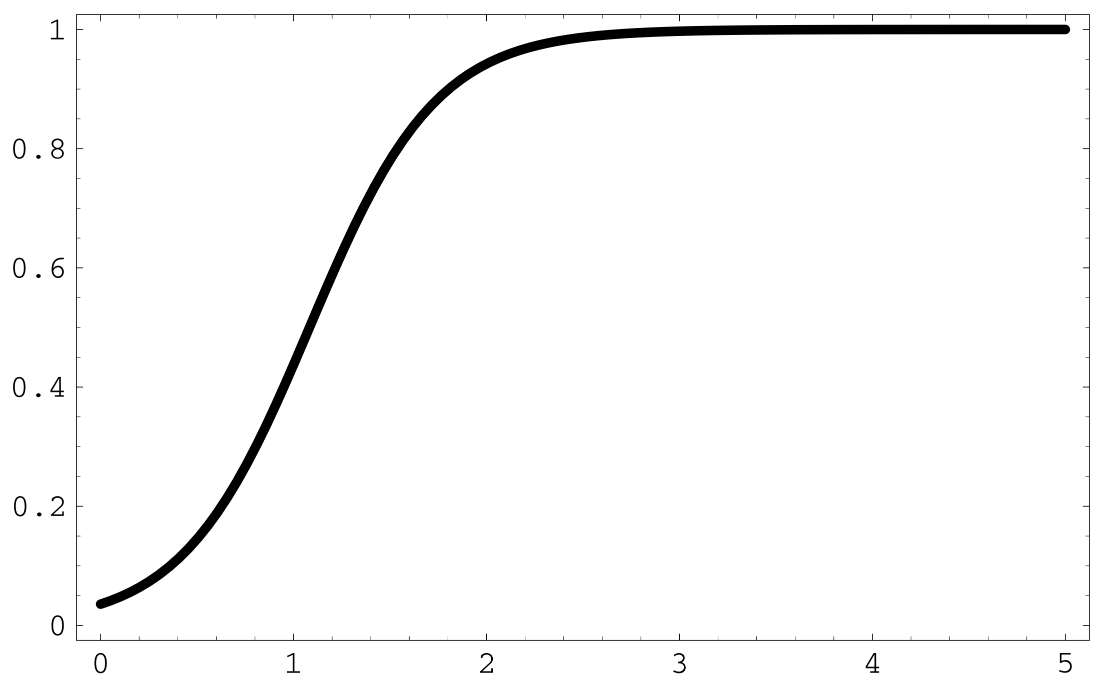
- Hacking's O.Q. #6: You are a physician. You think it is quite probable (say 90% probable) that one of your patients has strep throat (S). You take some swabs from the throat and send them to the lab for testing. The test is imperfect, with the following likelihoods (Y is + result, N is −):
  - $\Pr(Y \mid S) = 0.7$, $\Pr(Y \mid \sim S) = 0.1$
- You send five successive swabs to the lab, from the same patient. You get the following results, in order: Y, N, Y, N, Y. What is $\Pr(S \mid YNYNY)$?
- Hacking: Assume that the 5 test results are *conditionally independent*, *given* both S and $\sim S$, *i.e.*, that S *screens-off* the 5 tests results. So:
  - $\Pr(YNYNY \mid S) = 0.7 \cdot 0.3 \cdot 0.7 \cdot 0.3 \cdot 0.7 \approx 0.03087$
  - $\Pr(YNYNY \mid \sim S) = 0.1 \cdot 0.9 \cdot 0.1 \cdot 0.9 \cdot 0.1 \approx 0.00081$

$$\Pr(S \mid YNYNY) = \frac{\Pr(YNYNY \mid S) \cdot \Pr(S)}{\Pr(YNYNY \mid S) \cdot \Pr(S) + \Pr(YNYNY \mid \sim S) \cdot \Pr(\sim S)}$$
$$= \frac{0.03087 \cdot 0.9}{0.03087 \cdot 0.9 + 0.00081 \cdot 0.1} \approx 0.997$$

General Analysis of Hacking's "Odd Question #6"

- If n is the number of Y results, then $(5 - n)$ is the number of N results (out of 5 results). Bayes's theorem allows us to calculate $\Pr(S \mid E_n)$, where E_n is evidence consisting of n Y results and $(5 - n)$ N results (any order):

$$\frac{\Pr(E_n \mid S) \cdot \Pr(S)}{\Pr(E_n \mid S) \cdot \Pr(S) + \Pr(E_n \mid {\sim}S) \cdot \Pr({\sim}S)} = \frac{0.7^n \cdot 0.3^{5-n} \cdot 0.9}{0.7^n \cdot 0.3^{5-n} \cdot 0.9 + 0.1^n \cdot 0.9^{5-n} \cdot 0.1}$$



An Anecdotal Prelude to "Interpretations" of Probability

- After the O.J. trial, Alan Dershowitz remarked that "fewer than 1 in 1,000 women who are abused by their mates go on to be killed by them".
- He said "the *probability*" that Nicole Brown Simpson (N.B.S.) was killed by her mate (O.J.) — *given that he abused her* — was less than 1 in 1,000.
- Presumably, this was supposed to have some consequences for people's *degrees of confidence* (*degrees of belief*) in the hypothesis of O.J.'s guilt.
- The debate that ensued provides a nice segué from our discussion of the formal theory of probability calculus to its "interpretation(s)".
- Let A be the proposition that N.B.S. is abused by her mate (O.J.), let K be the proposition that N.B.S. is killed by her mate (O.J.), and let $\Pr(\cdot)$ be whatever probability function Dershowitz has in mind here, over the salient algebra of propositions. Dershowitz is saying the following:

(1) $$\Pr(K \mid A) < \frac{1}{1000}$$

- Shortly after Dershowitz's remark, the statistician I.J. Good wrote a brief response in *Nature*. Good pointed out that, while Dershowitz's claim may be true, it is not salient to the case at hand, since it *ignores evidence*.
- Good argues that what's relevant here is is the probability that she was killed by O.J., given that she was abused by O.J. *and that she was killed*.
- After all, we do know that Nicole was killed, and (plausibly) this information should be taken into account in our probabilistic musings.
- To wit: let K' be the proposition that N.B.S was killed (by *someone*). Using Dershowitz's (1) as a starting point, Good does some *ex cathedra* "back-of-the-envelope calculations," and he comes up with the following:

(2) $$\Pr(K \mid A \,\&\, K') \approx \tfrac{1}{2} \gg \tfrac{1}{1000}$$

- This would seem to make it far more probable that O.J. is the killer than Dershowitz's claim would have us believe. Using statistical data about murders committed in 1992, Merz & Caulkins "estimated" that:

(3) $$\Pr(K \mid A \,\&\, K') \approx \tfrac{4}{5}$$

- This would seem to provide us with an *even greater* “estimate” of “the probability” that N.B.S. was killed by O.J. Dershowitz replied to analyses like those of Good and Merz & Caulkins with the following rejoinder:

> . . . whenever a woman is murdered, it is highly likely that her husband or her boyfriend is the murderer without regard to whether battery preceded the murder. The key question is how salient a characteristic is the battery as compared with the relationship itself. Without that information, the 80 percent figure [as in Merz & Caulkins’ estimation] is meaningless. I would expect that a couple of statisticians would have spotted this fallacy.

- Dershowitz’s rejoinder seems to trade on something like the following:

(4) $\Pr(K \mid K') \approx \Pr(K \mid A \,\&\, K')$ [*i.e.*, K', not A, is doing the real work here]

- Not to be outdone, Merz & Caulkins give the following “estimate” of the salient probabilities (again, this is based on statistics for 1992):

(5) $$\Pr(K \mid K') \approx 0.29 \ll \Pr(K \mid A \,\&\, K') \approx 0.8$$

- We could continue this dialectic *ad nauseam*. I’ll stop here. This anecdote raises several key issues about “interpretations” and “applications” of Pr.

- Our discussants seem to be talking about some kind of "objective" probabilities involving *N.B.S.*'s murder (and murderer) *in particular*.
  * But, the "estimates" Merz & Caulkins appeal to involve *statistical frequencies* of murders in some *population*.
- First, *are there such things as* objective probabilities *at all*? If so, what are they (are there different kinds?) and what determines them?
- More specifically, are there objective probabilities of *token events*, or *only* frequencies (in populations)? If there are such probabilities of token events, then how (if at all) do they relate to frequencies?
  * Specifically, which population is "the right one" in which to include the token event (this is known as the *reference class problem*)?
- Finally, how are "objective" probabilities related to *degrees of belief*.
  * Generally, how are "objective" and "subjective" probabilities related?
- We'll be thinking more about some of these questions in the next unit.
- But, first, we're going to back-up and start "from scratch" ...

T-on-$\mathcal{I}$: Truth : : Pr-on-$\mathcal{M}$: Probability (I)

- In logic (and formal semantics), we have a *formal* notion called *truth-on-an-interpretation* (or *truth-on-$\mathcal{I}$*). This is *not truth* (simpliciter).
- It's useful to think about examples now. Here's a very simple example.
- Consider a 2-atom sentential language $\mathcal{L}$, where the atoms are *extra-systematically* understood as having the following content:
  - $X \stackrel{\text{def}}{=}$ John is unmarried.
  - $Y \stackrel{\text{def}}{=}$ John is a bachelor.
- As usual, we can picture all four interpretations of $\mathcal{L}$, as follows:

X	Y	Interpretations
T	T	$\mathcal{I}_1$
T	F	$\mathcal{I}_2$
F	T	$\mathcal{I}_3$
F	F	$\mathcal{I}_4$

☞ Facts about *truth-on-$\mathcal{I}_i$* do *not* depend on (*extra-systematic*) *content*.

T-on-$\mathcal{I}$: Truth : : Pr-on-$\mathcal{M}$: Probability (II)

X	Y	Interpretations
T	T	$\mathcal{I}_1$
T	F	$\mathcal{I}_2$
F	T	$\mathcal{I}_3$
F	F	$\mathcal{I}_4$

- Specifically, we have the following facts about *truth-on-*$\mathcal{I}_i$:
  - X is T-on-$\mathcal{I}_1$ and T-on-$\mathcal{I}_2$, but X is F-on-$\mathcal{I}_3$ and F-on-$\mathcal{I}_4$.
  - Y is T-on-$\mathcal{I}_1$ and T-on-$\mathcal{I}_3$, but X is F-on-$\mathcal{I}_2$ and F-on-$\mathcal{I}_4$.
- Indeed, *all* facts about *truth-on-*$\mathcal{I}_i$ are determined for *all* sentences p of $\mathcal{L}$ — *just by our conventions about truth-tables for truth-functional logic*.
- In this sense, *truth-on-*$\mathcal{I}_i$ does not depend on the extra-sysetmatic content of the sentences of $\mathcal{L}$. But, the *truth* (simpliciter) of sentences *does*.

☞ In this sense, *truth* is *external* to logic (and to formal semantics).

- OK, so then what *is truth*, and how is it related to *truth-on-*$\mathcal{I}$?

T-on-$\mathcal{I}$: Truth :: Pr-on-$\mathcal{M}$: Probability (III)

- For each interpretation $\mathcal{I}_i$ of $\mathcal{L}$, there is a corresponding state-description s_i of $\mathcal{L}$. As a result, ⌜p is T-on-$\mathcal{I}_i$⌝ is synonymous with ⌜$s_i \vDash p$⌝.
- What this reveals is that *truth-on-$\mathcal{I}$* is a *systematic logical concept*.
- On the other hand, *truth* is an *extra-systematic* concept.
- In our example, Y *extra-systematially entails* (or *conceptually necessitates*) X [$Y \Vdash X$], since it is a conceptual truth that all bachelors are unmarried.
- This allows us to *extra-systematically rule-out* the truth of the third state description s_3 of $\mathcal{L}$. That is, s_3 cannot be true, despite the fact that $Y \nvDash X$.
- This is a *very strong* sense of "*ruling-out* an interpretation". There are also two weaker senses of "ruling out" that can obtain:
  - Although $Y \nVdash X$, Y *conceptually probabilifies* X.
  - Y *epistemically* (but *not conceptually*) *probabilifies* X.
- Next: examples of each of these other two "grades of ruling-out".

T-on-$\mathcal{I}$: Truth :: Pr-on-$\mathcal{M}$: Probability (IV)

- Consider the following example, again involving two sentences X and Y.
  - $X \stackrel{\text{def}}{=}$ The coin will land heads when it is tossed.
  - $Y \stackrel{\text{def}}{=}$ The coin is heavily biased in favor of heads.
- Here, we have *neither* $Y \vDash X$ *nor* $Y \Vdash X$. But, $\mathcal{I}_3$ still seems somehow "inappropriate". If the coin was 2-headed, then we would have $Y \Vdash X$.
- There is some sort of *extra-systematic conceptual probability relation* between Y and X. But, $\sim X$ and Y are *not* conceptually *inconsistent* here.
- The natural thing to do here is to try to represent this as some sort of *probabilistic extra-systematic constraint*. But, which constraint is it?
  1. $Y \Vdash \Pr(X) \approx 1$. [$Y$ *e.-s.*-entails that X is highly probable. Meaningful?]
  2. $\Pr(Y \to X) \approx 1$. [The conditional $Y \to X$ is highly probable.]
  3. $\Pr(X \mid Y) \approx 1$. [The conditional probability of X, given Y, is high.]
- None of these *rules-out* the *truth* of s_3. But, they all place *e.-s.*-constraints on how *probable* s_3 is. For instance, (2) forces $\Pr(s_3)$ to be low (why?).

T-on-$\mathcal{I}$: Truth : : Pr-on-$\mathcal{M}$: Probability (V)

- Initially, we have only systematic constraints. Specifically, we have no systematic logical relations between atomic sentences, and the only systematic probabilistic constraints are $a_i \in [0, 1]$ and $\sum_i a_i = 1$. *E.g.*:

X	Y	Interpretations/S.D.'s	Models ($\mathcal{M}$)
T	T	$\mathcal{I}_1$ / s_1	$a_1 \in [0, 1]$
T	F	$\mathcal{I}_2$ / s_2	$a_2 \in [0, 1]$
F	T	$\mathcal{I}_3$ / s_3	$a_3 \in [0, 1]$
F	F	$\mathcal{I}_4$ / s_4	$a_4 = 1 - (a_1 + a_2 + a_3)$

- Then, we associate extra-systematic contents with the atoms, *e.g.*:
  - $X \stackrel{\text{def}}{=}$ John is unmarried.
  - $Y \stackrel{\text{def}}{=}$ John is a bachelor.
- In this case, we can *conceptually rule-out* interpretation $\mathcal{I}_3$ on extra-sysetmatic grounds. In other words, s_3 is (necessarily) false.

- That leads to the following extra-systematic revision of our initial STT:

X	Y	Interpretations/S.D.'s	Models ($\mathcal{M}$)
T	T	$\mathcal{I}_1$ / s_1	$a_1 \in [0,1]$
T	F	$\mathcal{I}_2$ / s_2	$a_2 \in [0,1]$
F	T	$\mathcal{I}_3$ / s_3	0
F	F	$\mathcal{I}_4$ / s_4	$a_4 = 1 - (a_1 + a_2)$

- In other cases, we will *not* be able to *rule-out* any interpretations. But, we will be able to *rule-out* certain *probability assignments/models*. *E.g.*:
  - $X \overset{\text{def}}{=}$ The coin will land heads when it is tossed.
  - $Y \overset{\text{def}}{=}$ The coin is heavily biased in favor of heads.
- In this case, let's assume the right constraint is $\Pr(X \mid Y) \approx 1$. Then, this will impose the following extra-systematic constraint on our initial STT:

$$\frac{a_1}{a_1 + a_3} \approx 1$$

- This doesn't *rule-out* any *interpretations*, but it does *rule-out* some *probability models*. Finally, there is a third "grade of ruling-out"...

T-on-$\mathcal{I}$: Truth :: Pr-on-$\mathcal{M}$: Probability (VI)

- Here is another example of a pair of sentences:
  - $X \stackrel{\text{def}}{=}$ The ball is black.
  - $Y \stackrel{\text{def}}{=}$ The ball is either black or white.
- Some philosophers claim that there is *some* sense in which we should have $\Pr(X \mid Y) = \frac{1}{2}$ here — as an extra-systematic constraint, of course.
- But, intuitively, it's a different sort of constraint than the one in our last example. In our last example "biased" was *itself* a *probabilistic* concept.
- Here, there is no *probabilistic extra-systematic content* involved.
- As such, if some extra-systematic probabilistic constraint is called for here, it's not for purely conceptual reasons. I will call this an *epistemic* extra-systematic constraint (an instance of the "Principle of Indifference").
- This can be motivated by unpacking $\Pr(X \mid Y)$ as (something like) "the degree of confidence one should have in X — *if Y were all one knew*."
- We'll come back to this *epistemic* understanding of probabilities shortly.

T-on-$\mathcal{I}$: Truth : : Pr-on-$\mathcal{M}$: Probability (VII)

- We'll come back to the probabilistic issues soon enough. Let's back up first, and think more about (extra-systematic) *truth* (simpliciter).
- There are various "Theories" or "Philosophical Explications" of *truth*. I have posted a nice overview by Haack (and the SEP entry by Glanzberg).
- I will separate the philosophical theories of truth into two categories:
  - Objective Theories of Truth.
    - Correspondence theories.
  - Subjective Theories of Truth.
    - Epistemic theories.
      - Coherence theories.
    - Pragmatic theories.
- There are also theories that are neutral on the subjective/objective question. For instance, deflationary theories (like the redundancy theory).

T-on-$\mathcal{I}$: Truth :: Pr-on-$\mathcal{M}$: Probability (VIII)

- According to correspondence theories of truth, p is true if p *corresponds to* some *truthmaker* $\mathfrak{t}_p$ (that is, if there exists a truthmaker $\mathfrak{t}_p$ for p).
- There are different views on the bearers of truth-values (sentences, propositions, beliefs) and truthmakers (facts, states of affairs).
- Moreover, there are different views about whether truthmakers must exist in some “mind-independent realm”. *Realists* will require that there is a mind-independent realm of truthmakers. Anti-realists will not.

Sentence (s): "John loves Mary."

↓ expresses

Proposition (p): that John loves Mary.

↓ corresponds to

Truthmaker ($\mathfrak{t}_p$): John's (actual) loving of Mary.

- If p is *false*, there is no corresponding $\mathfrak{t}_p$ at the bottom of the diagram.

T-on-$\mathcal{I}$: Truth : : Pr-on-$\mathcal{M}$: Probability (IX)

- Subjective theories of truth do not involve any sort of correspondence between sentences/propositions/beliefs and some realm of truthmakers.
- The *epistemic* theory of truth, for instance, holds that (Alston):

  > The truth of a truth bearer consists not in its relation to some "transcendent" state of affairs, but in the epistemic virtues the former displays within our thought, experience, and discourse. Truth value is a matter of whether, or the extent to which, a belief is justified, warranted, rational, well grounded, or the like.

- The coherence theory of truth is a instance of the epistemic theory (where coherence with one's other beliefs is the salient "epistemic virtue").
- The pragmatic theory of holds that "truth is satisfactory to believe". Basically, a belief is true if believing it "works" for its believer.

☞ We will adopt an objective/realist stance toward truth in this course. I find it hard to *understand* the other conceptions of truth. Explain.

T-on-$\mathcal{I}$: Truth :: Pr-on-$\mathcal{M}$: Probability (X)

- Just as we can talk about p being *true-on-*$\mathcal{I}_i$, which is synonymous with $s_i \vDash p$, we can also talk about p having *probability-*r*-on-*$\mathcal{M}$.
- And, like *truth-on-*$\mathcal{I}_i$, *probability-on-*$\mathcal{M}$ is a *logical/formal* concept.
- That is, once we have *specified* a probability model $\mathcal{M}$, this *logically determines* the *probability-on-*$\mathcal{M}$ values of all sentences in $\mathcal{L}$.
- Moreover, just as the *truth-on-*$\mathcal{I}_i$ of sentence p does not imply anything about p's *truth* (*simpliciter*), neither does the *probability-on-*$\mathcal{M}$ of p imply anything about p's *probability* (*simpliciter*) — if there be such a thing.
- Finally, just as we have different philosophical "theories" of truth, we will also have different philosophical "theories" of probability.
- And, as in the case of truth, there will be objective theories and subjective theories of probability. However, there will be more compelling reasons for "going subjective" in the probability case than in the truth case.
- Ultimately, we will be most interested in the *assessment of arguments*.