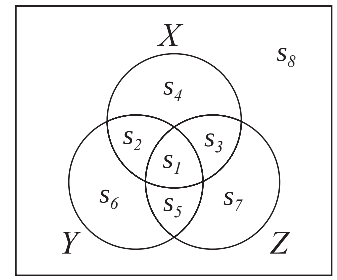
Conditionals Seminar: Day 3

- Administrative:
  - I've added a link to Hájek's SEP entry (good background on probability)
  - The lecture notes for my PHIL 148 class may also be useful on this score
  - **Oct. 5**: Alan Hájek will present most of the Chapter 5-related material
  - I'll present this week and next week (ch. 4 and intro to Ch. 5). Then, we'll move on to student presentations. Volunteers for Chapter 6/7 (in 2 weeks)?
- More Background: A Brief (Classical) Probability Calculus Primer
- Back to Bennett: Chapter 2 Wrap-Up, and then Chapters 3 & 4
  - Ramsey and Grice (or, more accurately, Ramsey and *Bennett's* Grice)
  - Jackson on Conventional Implicature, the Ramsey Test, $\rightarrow$, and $\supset$
  - The or-to-if inference [one more time]
  - Next: Chapter 4 – Lead-up to "The Equation", and Lewisian attacks on it

A Brief (Classical) Probability Calculus Primer 1: Small Boolean Algebras

- For present purposes, it suffices to think of the probability calculus as a (very simple) quantiative augmentation of classical, Boolean logic (or algebra).
- In fact, we won't even need to worry about algebras with more than three atomic propositions! Two representations of a 3-atom Boolean algebra:

X	Y	Z	States
T	T	T	s_1
T	T	F	s_2
T	F	T	s_3
T	F	F	s_4
F	T	T	s_5
F	T	F	s_6
F	F	T	s_7
F	F	F	s_8

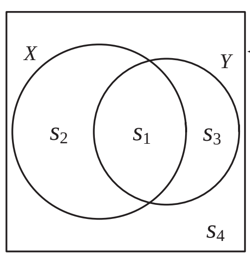


- In a 3-atom algebra, there are $2^3 = 8$ *states*. Each proposition in an algebra can be expressed as a *disjunction of states*. From here, probability is *easy*.

A Brief (Classical) Probability Calculus Primer 2: Probability Models 1

- A *probability function* $\Pr(\cdot)$ on a Boolean algebra $\mathcal{B}$ is just a function that maps each *state* s_i of $\mathcal{B}$ to a real number on the unit interval $[0, 1]$ such that the sum of all the $\Pr(s_i)$ is equal to one. Example of a 2-atom $\mathcal{B}$ plus a Pr:

X	Y	States	$\Pr(s_i)$
T	T	s_1	$\frac{1}{6}$
T	F	s_2	$\frac{1}{4}$
F	T	s_3	$\frac{1}{8}$
F	F	s_4	$\frac{11}{24}$



The area of the box is 1, since Pr(**T**) = 1.

- On the left, we have a "stochastic truth-table" representation of $\langle \mathcal{B}, \Pr \rangle$, and on the right we have a Venn Diagram representation in which the area of the regions is proportional to the probability of the corresponding propositions.
- A pair $\langle \mathcal{B}, \Pr \rangle$ consisting of a Boolean algebra of propositions $\mathcal{B}$ and a probability function Pr over $\mathcal{B}$ is called a (classical) *probability model*.

A Brief (Classical) Probability Calculus Primer 3: Probability Models 2

- Once you have a stochastic truth-table, which gives the probabilities of the states, calculating probabilities of arbitrary propositions in $\mathcal{B}$ is easy.
- Since each proposition x in $\mathcal{B}$ is equivalent to a disjunction of states, and since all pairs of states are mutually exclusive, the probability of any x is just the sum of the probabilities of the states that are these "disjuncts of x."
- In our 2-atom example, $\Pr(X \,\&\, Y) = s_1 = \frac{1}{6}$, $\Pr(X \vee Y) = s_1 + s_2 + s_3 = \frac{5}{6}$, $\Pr(X \supset Y) = \Pr(\neg X \vee Y) = s_1 + s_3 + s_4 = \frac{3}{4}$, *etc*., for all $x \in \mathcal{B}$.
- Traditionally, *conditional* $\Pr(\cdot \mid \cdot)$ is *defined in terms of un*conditional $\Pr(\cdot)$, as the *ratio*: $\Pr(x \mid y) = \frac{\Pr(x \& y)}{\Pr(y)}$, *if* $\Pr(y) > 0$; otherwise, $\Pr(x \mid y)$ is *undefined*.
- Kolmogorov (1933) gave the following three axiom characterization of $\Pr(\cdot)$:
  1. For all $x \in \mathcal{B}$, $\Pr(x) \geqslant 0$.
  2. For all $x \in \mathcal{B}$, if x is equivalent to $\top$, then $\Pr(x) = 1$.
  3. For all $x, y \in \mathcal{B}$, if $x \,\&\, y$ is equivalent to $\bot$, then $\Pr(x \vee y) = \Pr(x) + \Pr(y)$.

A Brief (Classical) Probability Calculus Primer 3: Probability Models 3

- It is easy to see that Kolmogorov's axiomatic characterization and my intuitive ones (in terms of stochastic truth-tables/Venn Diagrams) are equivalent.
- I prefer to use my algebraic approach, since it tends to be much easier to prove things algebraically than axiomatically (and I will do so whenever possible).
- [Note: $\Pr(\cdot \mid y)$ is itself a *probability* function (if defined), for all $y \in \mathcal{B}$, since (1) $\Pr(x \mid y) \geqslant 0$, for all $x, y \in \mathcal{B}$, (2) $\Pr(\top \mid y) = 1$, for all $y \in \mathcal{B}$, and (3) if x and z are mut. exclusive, then $\Pr(x \vee z \mid y) = \Pr(x \mid y) + \Pr(z \mid y)$, for all $y \in \mathcal{B}$.]
- This is, basically, all the background we'll need. Four key (easy) theorems:

  T1. For all $x, y \in \mathcal{B}$, $\Pr(x) = \Pr(x \,\&\, y) + \Pr(x \,\&\, \neg y)$.
  T2. $\Pr(x) = \Pr(x \mid y) \cdot \Pr(y) + \Pr(x \mid \neg y) \cdot \Pr(\neg y)$, *iff* $\Pr(y) \in (0, 1)$.
  T3. $\Pr(x \supset y) \geqslant \Pr(y \mid x)$, *iff* $\Pr(x) > 0$.
  T4. $\Pr(y \mid x) = \Pr(\neg x \mid \neg y) \Leftrightarrow \Pr(x \,\&\, y) = \Pr(\neg x \,\&\, \neg y)$, unless either $\Pr(x) = 0$ and $\Pr(y) = 1$ [$\not\Leftarrow$], or $\Pr(y \mid x) = 1 > \Pr(y)$ [$\not\Rightarrow$]. And, *radical* differences between $\Pr(y \mid x)$ and $\Pr(\neg x \mid \neg y)$ are possible.

§13: Ramsey and Grice — Some Points of Agreement & Conflict

- Sometimes, the Ramsey test (applied to "degree of assertibility") and Gricean prescriptions about assertability are in agreement, and sometimes they aren't.
- **Point of Agreement**:
  - '$A \to C$' does not get a high "assertability value" for me *merely because* I am highly confident in $\neg A$ (or merely because I am highly confident in C).
  - Grice: because I should (at least, in *some* cases) rather say something stronger and simpler if I am confident in such a thing.
  - Ramsey: because the criterion is a high probability in C, *given that* A. And, neither '$\Pr(A)$ is low' nor '$\Pr(C)$ is high' entails '$\Pr(C \mid A)$ is high'.
- **Points of Conflict** (five of them – NOTE: some of these seem unfair to Grice):
  1. Grice (sometimes) condemns asserting '$A \to C$' when $\neg A$ is *known*. Bennett suggests this generalizes ("in degree"?) to (some) cases where A is highly improbable (on my P). [I guess Grice would only frown here,

rather than condemning?] But, even if $\Pr(A)$ is very low, $\Pr(C \mid A)$ can (nonetheless) be very high. Plausibly, then, in some such cases, the theories are bound to disagree. Bennett gives no concrete cases to judge here, but some such cases seem inevitable, since Grice's criterion is clearly not equivalent to high conditional speaker-probability. Note: $\Pr(C \mid A)$ is *undefined* when $\Pr(A) = 0$! In such cases, '$A \to C$' fails the Ramsey test. So, the kind of disagreement Bennett has in mind here (which does seem to be somewhat overstated in the text) does not generalize from $\Pr(A)$ being low to $\Pr(A)$ being *zero*. Bennett discusses this later under the heading 'the zero intolerance of $\to$'. Bennett thinks this is an essential feature of the Ramsey Test, and that it is an essential feature of $\to$, but it seems to me that this is an artifact of the ratio analysis of conditional probability.

2. Grice (sometimes) condemns asserting '$A \to C$' when C is *known*. Bennett suggests this generalizes ("in degree"?) to (some) cases where C is highly probable (on my P). But, even if $\Pr(C)$ is very high, $\Pr(C \mid A)$ can be very high. So, while Grice will (sometimes) frown (at least, a little?), Ramsey might (in the same cases) smile. Note: $\Pr(C \mid A) = 1$ if

$\Pr(C) = 1$! So, Ramsey *must endorse* '$A \to C$' as highly assertible when C is *known*! This is a strike *against* Ramsey as *sufficient* for assertibility.

3. Grice (according to Bennett's Jackson) would say that logical truths are unassertible (since they are utterly informative). This is icorrect. In fact, Grice recognizes that sometimes logical truths are assertible sometimes (as he explicitly says in "Logic and Conversation"). Be that as it may, there is still a point of (true) disagreement lurking in the neighborhood here. All conditionals whose consequents are necessary truths will pass (all versions of) the Ramsey Test, since $\Pr(\top \mid A) = 1$, for all A. Grice, I take it, would say that *some* of these '$A \to \top$''s are unassertible (sometimes). So, this is a source of disagreement for Grice and Ramsey, and Bennett is aware that this is a tricky issue. We'll return to "necessary consequents" later on.
4. Some conditionals that are *truth-functionally equivalent* differ greatly in terms of their *assertibility* (and/or *believability*). For instance:

   (*i*) '(Even) if the bible is divinely inspired, it is not literally true.'
   (*ii*) 'If the Bible is literally true, it is not divinely inspired.'

If the horseshoe analysis is correct, then (*i*) and (*ii*) are *logically equivalent*, because they are *contrapositives*. Bennett says that (therefore) there is no way for Grice to account for the fact that (*i*) can be perfectly assertible/believable, while (*ii*) can't be. Why? He seems to think that the general conversational principles in Grice's theory are incapable of distinguishing truth-functionally equivalent claims in terms of their assertibility. I don't see why. Paper topic: Could there be a Gricean story about (*i*) and (*ii*)? Similar topics for all of these "divergences" between Ramsey and Grice. Bennett thinks this indicates that we need a "thicker" account of the meaning of →. In any event, *Ramsey* surely has no problem here, since '$\Pr(\neg T \mid D)$ is high' ⇏ '$\Pr(\neg D \mid T)$ is high'. [Thm. **T4** above]

5. Jackson's "Argument From Classroom Experience": Jackson's students are happy with Grice's stories about 'and' and 'or', but they strenuously resist his story about 'if'. He and Bennett think this suggests Grice's story is incomplete, and that the question 'Why does → seem not to be truth-functional?' cannot be adequately answered without some appeal to the *meaning* of '→' ("thick"). I guess Jackson's students like Ramsey?

Chapter 3: Jackson on Grice and Ramsey — §14: Setting the Stage

- In §14, some important background is given, and some subtleties clarified:
  - The horseshoe analysis is that → and ⊃ have the same *truth conditions*.
  - Grice and Jackson both are horseshoe analysts in this sense. The difference between them is the way they explain away the apparent counterexamples.
  - Grice uses considerations of *conversational* implicature to do this, and Jackson uses Ramsey Test + considerations of *conventional* implicature.
  - **Jargon**. C is *robust* with respect to A (for an agent) iff $\Pr(C \mid A)$ is high. For Jackson, this is equivalent to '$A \to C$' *passing the RT* (for the agent).
  - Jackson thinks robustness of C wrt A is *necessary but not sufficient* for assertibility/believability of '$A \to C$'. We have already seen our own potential failures of sufficiency (*e.g.*, '$A \to \top$'). Grice is *still relevant* here!
  - Jackson: robustness is *necessary* (*not* sufficient) for '$A \to C$' to be useful in *modus ponens*. He's thinking of "ticket entailment". Such MP "tickets" don't get you very far if '$A \to C$' is *not* robust (sometimes, even if it *is*!).

§15: Jackson on Conventional Implicature and →

- Like Grice, Jackson distinguishes what is *asserted* or *said* when one utters '$A \to C$', and what is *implicated* by such an utterance. And, like Grice, Jackson accepts the horseshoe analysis of what is said or asserted: $A \supset C$.
- Unlike Grice, Jackson does not talk about *conversational* implicature, which is determined by the rules of conversation that are presupposed to be in force in the context of utterance. Rather, Jackson talks about *conventional* implicature.
- Conventional implicature has to do with the "conventional meaning" of the indicative conditional → itself. And, for Jackson, this conventional meaning is encoded by the (passing of the) Ramsey Test (*i.e.*, the robustness of C wrt A).
- The basic idea is that when someone utters '$A \to C$', they *assert* $A \supset C$, but they *conventionally implicate* that C is robust with respect to A (*i.e.*, relative to *the speaker's* degree of belief/probability function Pr, $\Pr(C \mid A)$ is high).
- This isn't *quite* Strawsonian; it doesn't *quite* say that a "connection" between A and C is part of the meaning of →. But, it's close. It's "thick*er*", to be sure.

- Here, Jackson relies on an analogy with certain other connectives, like 'but':
  (1) Noam Chomsky would be a good Commencement speaker, and he is the country's most famous radical left-winger.
  (2) Noam Chomsky would be a good Commencement speaker, but he is the country's most famous radical left-winger.
- According to Jackson, (1) and (2) have the same *truth conditions* (they assert the same proposition). The difference between them is that (2) *conventionally implicates* some kind of *contrast* between the conjuncts of the assertion.
- This conventional implicature is generated not by conversational rules that are supposed by the hearer of the utterance to be in force, but by the *conventional meaning* of the connective 'but' itself. Jackson wants to tell a similar story for:
  (3) Either I didn't eat an egg for breakfast or you ate a million eggs.
  (4) If I ate an egg for breakfast, then you ate a million eggs.
- Jackson: (3) and (4) have the same truth conditions, but (4) conventionally implicates that its consequent is robust wrt its antecedent (*for its speaker*).

- In order for this to be plausible, the analogy between 'but' and 'if' has to hold water. Bennett thinks the analogy is leaky, and so is Jackson's story about 'if'.
- Before we discuss the details of the breakdown of '→'::'⊃' as 'but'::'and', two bits of background on Jacksonian conventional implicature are useful.
- **First**: the primary aim in discourse is to communicate beliefs. The asserted content of an utterance fixes the belief the speaker aims to communicate. The conventional implicatures can help/hurt by removing/adding obstacles.
- This explains why our response to 'Susan was poor but honest' is not to say that it is *false* (assuming both conjuncts are true!). Rather, we might say something like 'You may be right about Susan, but I wouldn't put it that way.'
- But, if the hearer is under the impression, say, that there is a strong correlation between being poor and being dishonest, then this use of 'but' may be to good communicative effect, since it may *remove* an obstacle in the hearer's mind.
- **Second**: (typically) what a speaker *implicates* can be *true* (apt) *even if* what she *asserts* is *improbable*. *E.g.*, 'Wittgenstein wasn't a deep thinker, but he influenced many deep thinkers' (implicature true, assertion improbable).

§16: The Case Against Jackson's Theory of →

- I'll only discuss the first 3 of 5 objections Bennett raises to Jackson's account.
  0. Recall that, typically, (*e.g.*, with 'but') what a speaker *implicates* can be *true* (apt) *even if* what she *asserts* is *improbable*. But, with indicative conditionals, this is *not* the case! If I utter '$A \to C$', then (Jackson says) I *assert* $A \supset C$, and I *implicate* that C is robust with respect to A. But, if $\Pr(A \supset C) < \epsilon$, then $\Pr(C \mid A) < \epsilon$. This follows from our **T3** (above).
  1. Bennett: how does the *implicature* of '$A \to C$' do any Jacksonian work?

     When I tell you 'If (A) Nixon's top aides were not blackmailing him into defending them, then (C) he gave them more loyalty than they deserved', I signal to you that my $\Pr(C \mid A)$ is high. I signal this to you . . . because this may help me to get across the belief I am primarily trying to communicate, namely that *either Nixon's top aides were blackmailing him into defending them or he gave them more loyalty than they deserved*. The conventionally implied robustness will improve your chances of acquiring precisely this belief rather than being distracted by irrelevant side-issues.

Bennett asks rhetorically: "*What* irrelevant side-issues? What are the threatening distractions, and how does the implication of robustness remove them?" There seem to be no answers forthcoming from Jackson. [Lewis provides a different motivation for asserting '$A \to C$', in cases where the speaker disbelieves A and believes C: to get the hearer to believe C even if he – unlike the speaker – believes A. This doesn't lend aid to Jackson here – it is more Adamsian. We'll discuss this later.]

2. There is also a *formal* disanalogy between 'but', 'nevertheless', 'yet', 'anyhow', and 'however' on the one hand, and 'if' on the other:

   When W links two sentences, it can be replaced by 'and' without affecting the truth conditions of what is asserted; when used as an operator on one sentence, it can be deleted without affecting said truth conditions.

   This property is shared by all the 'but'-like operators W (for which Jackson has compelling examples of the role of conventional implicature). But, 'if' does not have this property. This is further evidence of disanalogy.

- Bennett: Jackson's problem is that he (like Grice) doesn't look at a sufficient amount of 'if'-data, and he relies too much on an unmotivated 'but' analogy.

§18: The "Or-to-If" Inference [Take 2]

- I'm skipping §17, but one comment on it is in order. Bennett thinks that Jackson didn't go *far enough* toward a Stawsonian ("thick") account of →. He thinks the Ramsey Test contributes to the "*core* of asserted content". [Note: be careful not to read this as *truth conditions*! More on this later.]
- Bennett: there are two problems with the *abductive* "Or-to-If" Inference:
  - There are other (good) explanations of the facts about assertibility (and acceptability) in the story, which do not presuppose the ⊃-analysis, and
  - The ⊃-analysis *does not even explain* said facts in the first place.
- On the first point, one can use Grice + Ramsey (*without* ⊃) to explain why it was "all right for you to say" '$\neg P \to Q$' after Vladimir told you '$P \vee Q$'.
- On the second point, the ⊃-analysis *doesn't even explain* why it was "all right for you to *say*" '$\neg P \to Q$' after Vladimir told you '$P \vee Q$'. [Presumably, it *would* explain why it was "all right for you to *accept*" $\neg P \to Q$.] Bennett:

- Vladimir was behaving badly unless he was more confident of the disjunction than of either disjunct [$\therefore \Pr(P) \neq 1$]; and Grice's theory about conversational implicature explains why. If he was not misbehaving, therefore, he accepted the disjunction independently of whether one disjunct (either one) turned out to be false; so for him Q is robust with respect to $\neg P$ [$\because \Pr(P \vee Q \mid \neg P) = \Pr(Q \mid \neg P)$, if $\Pr(P) \neq 1$]. That would make it all right by the Ramsey test for him to assert $\neg P \to Q$; and your trust in him makes it all right for you to assert this also.
- Anyway, whether or not some rival to the horseshoe analysis does explain the acceptability of the or-to-if inference, the analysis itself does not. It contributes only the thesis that $P \vee Q$ entails $\neg P \to Q$, and thus that the truth of what Vladimir told you guarantees the truth of what you told Natalya. But it is a famous fact that a true material conditional may be an absurd thing to say; so this entailment thesis does not imply or explain the fact that if you accepted what Vladimir told you, then it was all right for you to say what you did to Natalya.

- This reads the or-to-if inference as an *abduction*, where the *explanandum* is that it was "all right for you to say" '$\neg P \to Q$' to Natalya after Vladimir told you '$P \vee Q$', and the *explanans* is that the $\supset$-analysis is correct. Is IBE *valid*?

Chapter 4: The Equation — §19 (Other Approaches)

- Strawson's (*non*-$\supset$) idea was that the *meaning* of $\to$ involves a *metaphysical* "connection" between antecedent and consequent. Counterexamples:
  'If she apologized to him, then he lied to me.'
  '(Even) if he apologizes, I shall (still) be angry.'
- Davis and Stalnaker give (*non*-$\supset$) "possible world" semantics for $\to$ (later on).
- Lycan has yet another (*non*-$\supset$) approach, which we'll also discuss later on.
- Bennett thinks the Ramsey Test is key to a correct (*and non*-$\supset$!) analysis of $\to$. "The *acceptability* of $A \to C$ for a person at a time is governed by the probability the person then assigns to C on the supposition of A [$\Pr(C \mid A)$]."
- Bennett makes some strong assumptions here about "acceptability" and Pr:

  > ...acceptability depends upon probability and nothing else: you ought not to accept what you do not find probable, and there is no obstacle to your accepting what you do find probable. So we can move to the thesis that the probability for you of $A \to C$ is proportional to your probability for C on the supposition of A.

Digression: Why Acceptability May Depend On Something More Than Probability

- Let $R(x)$ mean "it would not be irrational for you to accept x". Let $\Pr(x) \geqslant r$ say that x has "high probability," according to your (rational) Pr. Consider the following claim about high Pr and non-irrational acceptance (NIA).

(1) $R(x) \iff \Pr(x) \geqslant r$. [high (rational) Pr is necessary and sufficient for NIA]

- (1) faces challenges in both directions. First, the ($\Rightarrow$) direction:
- Let B_1 and B_2 be two propositions such that (*i*) $R(B_1)$ and $R(B_2)$, and (*ii*) B_1 and B_2 are *independent*, according to your (rational) credence function Pr.
- Now, assume (*iii*) $\Pr(B_1) = \Pr(B_2) = r < 1$. Then, B_1 & B_2 will *not* have high Pr: $\Pr(B_1 \,\&\, B_2) = r^2 < r$. Can't this be consistent with $R(B_1 \,\&\, B_2)$?
- Example: B_1: Bonds will hit another home run this season; B_2: it will rain in Berkeley sometime next January. *Mightn't* it be OK to accept B_1 & B_2?
- Generally, if you accept two independent propositions non-irrationally, and you (rationally) assign each of them Pr = r (the "threshold value"), then, if ($\Rightarrow$) is correct, you *cannot* rationally accept their conjunction. Too strong?

- Intuitively, high Pr may also not be *sufficient* for "what you'd not be irrational to accept". Let L_i state that ticket i will not be the winning ticket in a lottery that you *know* to be very large (n) and to have exactly one winning ticket.
- Then, rational degrees of belief may be such that $\Pr(L_i)$ is (very) high, for each i. But, $\Pr(L_1 \,\&\, \cdots \,\&\, L_n) = 0$. Assuming ($\Leftarrow$), you'd not be rational to believe *each* L_i. But, you *would* be irrational to believe their conjunction!
- So, ($\Leftarrow$) implies that $\mathcal{B}$ = "the set of propositions you'd not be irrational to accept" isn't closed under *known* logical consequence. *The Lottery Paradox*. If you're inclined to think that $\mathcal{B}$ *is* closed under known-$\vDash$, you'll reject ($\Leftarrow$).
- Perhaps a weaker claim can be salvaged here? What if we retreat to:

(2) $R(x) \Rightarrow \Pr(x)$ is *not low*. [*low* $\Pr(x)$ is sufficient for $\neg R(x)$]

- This principle may seem plausible. But, again, if you think $\mathcal{B} = \{x \mid R(x)\}$ is closed under known-$\vDash$, then you'll reject *even* (2). Extend the first case above to n independent propositions, each NIA, and each with a Pr = r (high). Their conjunction will have low Pr, provided n is *large enough*.

§20: Kinds of Probability

- We have seen the formal theory of probability (axiomatically and algebraically presented). There are various *interpretations* (models) of this formal theory:
  - Physical Probabilities (*a.k.a.*, objective chances of events): If the universe is indeterministic, then even knowing all the laws and all the history of the universe will yield only knowledge of *probabilities of* future events.
    - Propensities of some token events to bring about others (in set-ups).
    - Hypothetical limiting frequencies of events in chance set-ups.
  - Statistical Frequencies in Classes of Events. Even in deterministic universes, systems will exhibit statistical regularities in the (actual) frequencies of events (relative to fixed reference classes of events). *E.g.*, Models of games of chance, and classical systems in statistical physics.
  - **Personalistic Probabilities (Degrees of belief of "rational" agents).**
    - ***Pragmatically* rational degrees of belief (expected utility theory, etc.)**
    - ***Epistemically* rational degrees of belief (probabilistic epistemology)**
- Hájek's SEP entry (online) is *the* thing to read on "kinds of probability".

§21: Elements of "Probability Logic"

- Here, Bennett misuses the term "probability logic" — a term of art which has to do with using the probability calculus to evaluate argument strength (or the transmission of probabilities by inferences). He means "probability calculus".
- We've already seen the probability calculus (NOTE: our presentation is more rigorous and illuminating, and this will be important for giving a more accurate and comprehensible story about the issues raised in chapter 5).
- A few more theorems Bennett discusses here (obvious, *algebraically*):
  - If q entails r, then $\Pr(q) \leqslant \Pr(r)$.
  - Therefore, if q is logically equivalent to r, then $\Pr(q) = \Pr(r)$.
  - $\Pr(\neg p) = 1 - \Pr(p)$
- **Definition**. p and q are *independent* under Pr iff $\Pr(p \,\&\, q) = \Pr(p) \cdot \Pr(q)$.
- **Definition**. p and q are *correlated* under Pr iff $\Pr(p \,\&\, q) > \Pr(p) \cdot \Pr(q)$.
- **Definition**. p and q are *anti-correlated* under Pr iff $\Pr(p \,\&\, q) < \Pr(p) \cdot \Pr(q)$

§22: The Ratio "Formula"

- The conditional probability of p, *given* q is written $\Pr(p \mid q)$, and it is supposed to represent (for our purposes) the degree of belief an agent [whose credence function is $\Pr(\cdot)$] assigns to *p on the supposition that q is true*.
- We've seen that $\Pr(p \mid q)$ is, traditionally, *defined* in terms of the unconditional probability function $\Pr(\cdot)$, as the *ratio* of $\Pr(p \,\&\, q)$ to $\Pr(q)$.
- A ratio *definition* implies that $\Pr(p \mid q)$ is *undefined* when $\Pr(q) = 0$. Bennett does *not* want to take the ratio as a *definition* or an *analysis* of $\Pr(\cdot \mid \cdot)$. He merely wants wants to think of the "ratio *formula*" as a *truth* about $\Pr(\cdot \mid \cdot)$.
- He thinks the *Ramsey Test* is "definitive" of conditional probability, and (I guess) that the ratio formula is (somehow) implied by the Ramsey Test.
- He concedes that there are cases in which the ratio can't be determined (or has no value), despite the fact that $\Pr(\cdot \mid \cdot)$ can be (or does). He avoids those cases (mainly, Hájek's $\vDash$ examples) by eschewing "independent conditionals".

- There are other cases in which $\Pr(p \,\&\, q)$ is not known, but $\Pr(p \mid q)$ is. For this problem, Bennett introduces another formula (he says is "arithmetically equivalent" to RF), which can be used (sometimes) to determine $\Pr(p \,\&\, q)$:
  $\Pr(p \,\&\, q) = \Pr(p \mid q) \cdot \Pr(q)$
- He ignores continuous magnitudes (points with "measure zero"). He says that "no such theoretical possibility could have any bearing on our lives". This seems pretty reasonable to me, since we're talking about *credences*.
- Bennett *needs* the ratio formula to be true (how else will he prove that conditional probabilities are *probabilities*, using only facts about *unconditional* probability?), and he provides an argument for its plausibility.
- This argument examines the dilemma $\Pr(p \,\&\, q) < \Pr(q)$ or $\Pr(p \,\&\, q) = \Pr(q)$. [Probability calculus prevents $\Pr(p \,\&\, q) > \Pr(q)$.] He considers these cases, in turn, and says that the ratio gives the intuitive answer.
- It is, mainly, the cases where $\Pr(A) = 0$ that give the ratio the most trouble, from an intuitive point of view. And, these are the only cases (in which all unconditional probabilities are known) where Hájek's arguments have force.

§23: Indicative Conditionals are Zero-Intolerant

- Bennett spends a fair amount of time here trying to argue that the ratio formula is not really missing anything (concerning →) by "going undefined" when $\Pr(q) = 0$. He thinks that indicatives are "zero intolerant" anyway.
- He says "nobody has any use for $A \to C$ when for him $\Pr(A) = 0$." *I.e.*,

  > . . . combine the Ramsey test thesis with the Ratio Formula . . . the acceptability to a person of $A \to C$ is given by his probability for $A \,\&\, C$ divided by his probability for A: 'Believe $A \to C$ to the extent that you think $A \,\&\, C$ is nearly as likely as A'. You can do nothing with this in the case where your $\Pr(A) = 0$.

- He acknowledges three counterexamples, but then claims they can be ignored:
  1. **Conversational Stretches**. We assert $A \to C$ when *for us* $\Pr(A) = 0$, because the hearer is *not* such that *for them* $\Pr(A) = 0$. But, this sounds like *assertibility* again, and not *acceptability* of the conditional.
  2. **Inferred →s**. "The thesis that indicative conditionals are zero-intolerant should be confined to ones that stand on their own feet, and not applied to

inferred ones." $A \to C$ may be accepted, *despite* $\Pr(A) = 0$, since it's a *known consequence of* $A' \to C'$; $R(A' \to C')$, and $\Pr(A') \neq 0$. *Lotteries?*

3. Independence, Non-Interference, and something nearby.
   (a) With independent conditionals, we can simply move straight to certainty in the consequent without worrying about how to "accommodate" $\Pr(A) = 1$ – we can just "dump it in there and go straight to $\Pr(C) = 1$."
   (b) Non-interference conditionals are accepted on the ground that the person holds C to be true and thinks that A's being true would not interfere with that (no need for Ramsey here either, according to B).
   (c) Other examples don't fit exactly into either of these classes:

   > 'If Hobbits exist, then Tolkien's novel refers to real beings as its heroes.'

   With such $A \to C$s, the person accepts $A \to B$ as a non-interference conditional, where B asserts the existence of some item having a certain meaning or content, and C assigns to B a relational property ('is true', 'does not refer to anything', *etc*.) which it must have if A is true.

- Bennett thinks some purported counterexamples in which $A \to C$ is thought to be true, despite A's having probability zero, are really *subjunctives* in disguise.

§24: The Equation

- **The Equation** asserts that the acceptability (for S) of the indicative conditional is the conditional probability of C, given A (that S assigns):

  $\Pr(A \to C) = \Pr(C \mid A) = \frac{\Pr(A \& C)}{\Pr(A)}$

- Something like this thesis has been defended by Stalnaker, Adams, and Jeffrey (among others). I add the ratio at the end, since *Bennett* accepts it as true.
- Bennett suggests here that '$A \to C$' *does not express a proposition*. He says, for this reason, that there is no account of the *truth-conditions* of '$A \to C$'. But, he still thinks this consistent with the "Ramseyan spirit" of The Equation.
- This will take a lot of explaining on his part! If '$A \to C$' does not express a proposition, then the probability calculus cannot be used to discuss it either!
- In Chapter 5, various "triviality" arguments against The Equation are discussed. Bennett thinks these, too, point to his "no-proposition" thesis.

- I think this is misguided and unnecessary as a response. I'll explain why next.
- I'll let Alan summarize the most compelling arguments he sees for "triviality". But, meanwhile, I will give my own take on the "triviality proofs".
- I will give the simplest "proof" I know (here, I follow Milne's recent paper).
- I will do so in a way that makes strong contact with the horseshoe analysis.
- I will identify the minimal set of premises one needs for such "proofs".
- I will explain why Bennett's response is unnecessary and ill-motivated.
- I will argue that what's really at issue here does not involve conditional probability at all, but certain kinds of (non-obvious, and perhaps *false*) assumptions about *unconditional* probabilities of *iterated* conditionals.
- Bennett will still have what he calls "philosophical" arguments for the conclusion he wants (that indicatives do not express propositions).
- But, I think the considerations I raise about Lewisian arguments against The Equation will also raise doubts about Bennett's "philosophical" motivations.

Chapter 5: Lewisian "Triviality Proofs" — A Brief, Opinionated Overview

- A.H. will present this stuff in more detail (and Lewis-style) on Oct. 5. Meanwhile, here's my take on it. I begin with my favorite "proof" (Milne):

  Premise 1. S's (*propositional*!) degree of belief structure $\langle \mathcal{B}, \Pr \rangle$ is a *probability model* (all we *need* is the assumption that S's degrees of belief Pr over *propositions* [$\mathcal{B}$] satisfy the two *expansion laws* **T1** and **T2**, above).

  Premise 2. S's algebra of propositions (propositional belief structure) $\mathcal{B}$ is *closed* under indicative implication. That is, for all $a, c \in \mathcal{B}$, $a \to c \in \mathcal{B}$.

  Premise 3. S's Pr satisfies The Equation: $\forall a, c \in \mathcal{B}$, $\Pr(a \to c) = \Pr(c \mid a)$.

  Premise 4. S's Pr satisfies the following constraint (Lewis): $\forall a, b, c \in \mathcal{B}$,
  $$\Pr(a \to c \mid b) = \Pr(c \mid a \,\&\, b)$$

  Conclusion. S's Pr is such that, for all *contingent* [$\Pr \in (0,1)$] $a, c \in \mathcal{B}$,
  $$\Pr(a \to c) = \Pr(a \supset c)$$

- This argument is valid (Milne). Of course, Bennett denies its conclusion! So, he must deny at least one of its four premises. In fact, he denies Premise 2.

- If I were an Equation Guy, my reaction to this "proof" would not be to deny that $\mathcal{B}$ is closed under $\to$. To be honest, I'm not even sure what that would mean (more later). I think the problem with this argument is Premise 4.
- Lewis calls Premise 4 "closure of the class of probability functions under conditionalization", and Bennet reports a "proof" of Premise 4 due to Alan H.
- I think this "proof" is fallacious (more when Alan is here!). Moreover, Lewis' constraint has nothing in particular to do with conditionalization. In fact, it is equivalent to a constraint on *unconditional* probabilities of *iterated* $\to$s!
- From Premises 1–3 (*i.e.*, The Equation), we can derive the following pair:
  - For all $a, b, c \in B$, $\Pr(b \to (a \to c)) = \Pr(a \to c \mid b)$.
  - For all $a, b, c \in B$, $\Pr(c \mid a \,\&\, b) = \Pr((a \,\&\, b) \to c)$.
- Therefore, given Premises 1–3, Lewis' constraint (Premise 4) is *equivalent to*:

  Premise 4*. S's Pr satisfies the following constraint: $\forall a, b, c \in \mathcal{B}$,
  $$\Pr(b \to (a \to c)) = \Pr((a \,\&\, b) \to c)$$
- This is a (probabilistic) *importation/exportation law* for the indicative $\to$.

- The structure of the "triviality proof" is a *reductio*. Bennett wants to "aim" the *reductio* at Premise 2. But, one can just as easily "aim" it at Premise 4.
- And, there are independent reasons for *Bennett* to reject Premise 4*.
- Unless $\to$ satisfies the import-export *law* (IEL), it's hard to see why we should accept Premise 4*. Gibbard argues that if $\to$ satisfies IEL, then $\to = \supset$!
- Here's Gibbard's argument. Consider the following indicative conditional:

  (1) If $A \supset C$, then $A \to C$.

  Assuming that $\to$ satisfies the IEL, (1) is logically equivalent to:

  (2) If $((A \supset C) \,\&\, A)$, then C.

  The antecedent of (2) entails its consequent. So (2) is a logical truth. So by IEL, (1) is a logical truth. All agree that $A \to C$ entails $A \supset C$. So (1) entails

  (3) $(A \supset C) \supset (A \to C)$

  So (3) is a logical truth. Thus, $(A \supset C)$ entails $A \to C$, and $\to = \supset$.
- Thus, IEL entails the $\supset$-analysis of $\to$. Since Bennet rejects the $\supset$-analysis, he has reason to reject the IEL, and its probabilistic analogue Premise 4*.