Four Decades of Scientific Explanation, Cont'd

- Administrative:
  - Let's talk about papers this week ...
  - I will be away next week ...
  - Guest Lecture on the 20th (Strevens)
- Probabilistic Accounts of Explanation
  - More on the S–R Account
    - "Homogeneity" and "Spurious Correlations"
    - The problem of the reference class
  - Back to the I–S Account
    - Further Discussion of maximal specificity
    - The problem of the reference class
- Time Permitting: Interpretations of Probability

Statistical–Relevance Explanation: Review

- A *partition* $\{A \,\&\, C_i\}$ of a class A is a collection of mutually exclusive and exhaustive subsets of A. Each subclass $A \,\&\, C_i$ in the partition $\{A \,\&\, C_i\}$ is called a *cell* of the partition.

- A partition $\{A \,\&\, C_i\}$ of A is *relevant* with respect to B if the probability of B in each cell $A \,\&\, C_i$ of the partition is different from each other cell. That is, if $\Pr(B \mid A \,\&\, C_i) \neq \Pr(B \mid A \,\&\, C_j)$, for all $i \neq j$.

- A partition $\{A \,\&\, C_i\}$ of A is *homogeneous* with respect to B if $\{C_i\}$ includes *all and only* the (causally?) relevant factors regarding B (in A).

  Objectively homogeneous = all and only relevant factors

  Epistemically homogeneous = all & only *known* relevant factors

- The gender-partition of A ($\{A \,\&\, C_i\}$) is *relevant* to acceptance (B). So is the department-partition ($\{A \,\&\, D_i\}$). The D_i *screen-off* the C_i from B (in A), but *not conversely*: $\Pr(B \mid A \,\&\, C_1 \,\&\, D_i) = \Pr(B \mid A \,\&\, C_2 \,\&\, D_i)$, but $\Pr(B \mid A \,\&\, D_i \,\&\, C_k) \neq \Pr(B \mid A \,\&\, D_j \,\&\, C_k)$. So, the department renders the gender irrelevant to acceptance. The D_i and not the C_i should be included.

Screening–Off and the Principle of the Common Cause

- Z *screens-off* X from Y (in reference class A) iff

$$\Pr(X \mid Y \,\&\, Z \,\&\, A) = \Pr(X \mid {\sim}Y \,\&\, Z \,\&\, A)$$

and

$$\Pr(X \mid Y \,\&\, {\sim}Z \,\&\, A) = \Pr(X \mid {\sim}Y \,\&\, {\sim}Z \,\&\, A)$$

- Also said: Y and Z are *conditionally independent*, *given* X (and ${\sim}X$) in A. We write this formally, as follows: $X \perp Y \mid Z$.
- Screening-off is a crucial tool for detecting and understanding causal structure. Reichenbach (1956) postulated the ***principle of the common cause***: if X and Y are (unconditionally) correlated in A, then either (1) X causes Y (in A), or (2) Y causes X (in A), or (3) there is a *common cause* Z of X and Y which *screens-off* X from Y.
- This kind of principle is central to most probabilistic causal modeling methods (including Bayesian network models of causation).

Screening–Off, (PCC), and S–R Explanation

- Reichenbach's (PCC) is true in deterministic worlds. This is because one can always invoke the entire past state of the universe Z to screen-off any two correlated events X and Y from each other.
- But in non-deterministic worlds, (PCC) may not be true generally. The most well-know (apparent) counterexamples to (PCC) are the correlations observed in certain quantum-mechanical experiments.
- For excellent discussions of (PCC), see Van Fraassen's "The Scientific Image" and Sober's "The Principle of the Common Cause."
- Salmon's basic suggestion seems to be that we should look at all subsets F of the total set of factors $\mathcal{F}$ which may be relevant to B (in reference class A). We seek a subset F of $\mathcal{F}$ which screens-off all other subsets F' of $\mathcal{F}$ from B (in A), such that F is *not* screened-off from B (in A) by any other subset F'' of F. This F will be (at least, *epistemically*) homogeneous and relevant with respect to B in A.

S–R and I–S Summarized and Compared

- In a nutshell, the two accounts can be expressed as follows:
  - I–S: An explanation of a particular fact f an *inductive argument* which confers a *high inductive probability* on the explanandum f.
  - S–R: An explanation of a particular fact f an assemblage of facts which are *statistically relevant* to the explanandum f *regardless of the degree of probability (of f) that results*.
- Both accounts conform to the *covering law* model. I–S explicitly requires that a statistical law occur in the explanans. S–R includes reference to the "prior probability" $\Pr(B \mid A)$, which can be thought of as a statistical law in the same sense (more on this point, below).
- Both accounts face the *problem of the reference class*. S–R faces *additional* problems arising from the possibility of spurious correlations and/or "hidden" relevant factors. The problem of the reference class is *the* problem facing *any* probabilistic account (of *anything*!) . . .

I–S and the Problem of the Reference Class

- Hempel added the *requirement of maximal specificity* (RMS) to attempt to control the non-monotonicity of inductive inference. It *seemed* that his problem was with "inductive inconsistency":

$$\frac{\begin{array}{l}\Pr(Gx \mid Fx \,\&\, Hx) = r \\ Fb \,\&\, Hb\end{array}}{Gb}\ [r] \qquad \frac{\begin{array}{l}\Pr({\sim}Gx \mid Fx \,\&\, Hx \,\&\, Jx) = r \\ Fb \,\&\, Hb \,\&\, Jb\end{array}}{{\sim}Gb}\ [r]$$

- But, as Coffa points out, if *this* was all he was worried about, there is a *much easier way* to "handle" this. We *know* which of Gb and ${\sim}Gb$ is true! So, the fact that there can be strong arguments for both Gb and ${\sim}Gb$ is not a serious problem for *emplaning* Gb (or ${\sim}Gb$, atcmb).
- What Hempel is really struggling with here is the *problem of the reference class*. This is a *fundamental* problem ...
- Questions: What is the probability that John Doe will still be alive 15 years from now? Is John Doe of average height? See the problem?

I–S and the Problem of the Reference Class II

- The problem here is that John is a member of *many* collections of people (or, for that matter, *organisms* or *physical objects*!). And, the 15-year survival probabilities may vary greatly from one class to another. The same can be said, of course, for the average height!
- So, which reference class do we choose to determine the probability? Salmon suggests that we should choose the *broadest possible homogeneous reference class*. That is, we should include all factors that we know to be relevant, and none that we know to be irrelevant.
- Problem #1: which factors are "relevant" will depend on what the total constellation of factors is taken to be. We need to choose a "ground reference class" to determine the "prior probabilities" in the first place.
- Problem #2: in cases of explanation, we *already know* that the object in question has the property in question. So, we *cannot* use the broadest homogeneous class (since it *contains* the explanandum itself!).

I–S and the Problem of the Reference Class III

- Since the explanandum is already part of our total evidence, we cannot us our *total* evidence to determine the "explanatory probabilities" (otherwise, all explanations are *trivial*!). This makes confirmation and explanation much different, from a probabilistic point of view.
- So, which parts of our total evidence are we to allow in our inductive explanations? Hempel's (RMS) is intended to answer this question …
- Hempel's (RMS) correctly *excludes* the explanandum-fact that b is a G, which is part of our total evidence, but would trivialize the explanation of Gb, since $\Pr(Gb \mid Fb \,\&\, Gb) = 1$, for *all* F's. But, (RMS) is *too weak*.
- Let Rb be the fact that John's recovery from strep (Gb) was reported by all the major news agencies (who corroborated their stories through reliable sources, etc.). Here, Fb is that John was treated with Penicillin. Presumably, $\Pr(Gb \mid Fb \,\&\, Rb) \approx 1$. But, it seems that Rb is *not explanatorily* relevant (although it is *evidentially* relevant to Gb).

I–S (S–R) and the Problem of the Reference Class IV

- Hempel's (RMS) does not exclude Rb from the (explanatory) reference class, even though it seems explanatorily irrelevant.
- Salmon proposes adding the following additional "unless clause" to Hempel's (RMS) "or unless b's membership in F_1 cannot be known before its membership in G has been ascertained."
- This is a common assumption added to probabilistic accounts of causation and explanation: that *causes (explanans) precede their effects (explananda)*. Salmon adds this to his own S–R account as well (as he *must*, on pain of falling prey to similar "counterexamples"!).
- So, Salmon suggests that the key here is to add (1) objective homogeneity of the reference class, and (2) temporal priority of the explanans. This goes for both I–S and S–R accounts.
- But, how do we determine if a partition is homogeneous in the appropriate way? Salmon suggests a method . . .

I–S (S–R) and the Problem of the Reference Class V

- Salmon suggests constructing an "associated sequence of events". Example: let F be a sequence of tosses of a coin, and let G be the coin landing heads up. We wonder whether the outcome of the previous toss is relevant to the outcome of the present toss. We form an associated sequence of tosses of the same coin A in which each toss x_i is associated with its predecessor x_{i-1} (B is the attribute of landing heads up). The selection S then picks out the tosses immediately following a heads up toss. For a standard coin, we'll have: $\Pr(G \mid F) = \Pr(G \mid F \,\&\, S)$.
- So, invariance across *all* (salient) associated sequences is the hallmark of homogeneity. Problem: which attributes B should be allowed for constructing "associated sequences?" Salmon requires that the B's be "objectively codefined" with the explanandum class (that is, that their values be determinable *before* the determination of the explanandum attribute is made — temporal precedence again).

Coffa's "Dispositional" I–S Model

- Coffa, like Salmon, was looking for a probabilistic model of explanation that was *not* inherently *epistemically relativized.*
- Coffa talks about the "objective disposition or propensity of a X to bring about Y". He says that, if nature is indeterministic, then there must be objective, *probabilistic* propensities $\mathcal{P}(Y \mid X)$.
- Coffa suggests that probabilistic explanations should be of the form:

$$\begin{array}{l} \mathcal{P}(Ga \mid Fa) = r \\ Fa \\ EC(a, G, F) \\ \hline Ga \end{array}$$

- Here, $EC(a, G, F)$ is an extremal clause which states that nothing in this explanatory situation is nomically relevant to a's having G except for a's having F. Nomic relevance is different from statistical relevance.

Coffa's "Dispositional Propensity" I–S Model II

- Coffa sticks with the basic I–S idea (*i.e.*, *not probabilistic relevance*). But, he moves away from the "epistemic" approach of Hempel.
- Does the "propensity interpretation of probability" avoid the problem of the reference class? Are propensities *probabilities* at all? It seems that certain propensities are "nonsensical" (*e.g.*, the propensity of future events to bring about past events). Humphreys' paradox.
- What does "nomically relevant" mean? This is not the same thing, presumably, as "probabilistically relevant under a propensity interpretation of probability," or is it? Salmon's account tries to explain "nomic relevance" in terms of an objectively homogeneous partition of factors. [Are propensities like Dormative Virtues?]
- For an excellent discussion of propensity interpretations of probability, see Gillies' recent *BJPS* paper "Varieties of Propensity".

Determinism, Indeterminism, and Explanation

- According to Salmon, Coffa, and many others, if the universe were deterministic, then all explanations would be deductive. That is, there would be no *irreducibly* probabilistic explanations in a deterministic world. Any probabilities that remained would be "purely epistemic". Such explanations would merely be *incomplete*.
- In Salmon's S–R theory, this means that in deterministic worlds, further partitions will always be possible, until a partition which *determines* the outcome is finally reached.
- In Coffa's theory, all "propensities" would be *perfect* in deterministic worlds, and so we would always have $\mathcal{P}(Ga \mid Fa) = 1$.
- In Hempel's theory, an omniscient modeler would always end-up with extreme epistemic probabilities in their I–S explanations.
- Are all incomplete explanations bad explanations? Isn't explanation also *pragmatic* in nature (meeting a *demand* for explanation)?

Interpretations of Probability I

- There are many different interpretations of probability.
  1. Subjective Interpretations
     (a) Probabilities as *justified* degrees of
     (b) Probabilities as *rational* degrees of belief
     (c) (Logical probability is often put here . . .)
  2. Objective Interpretations
     (a) Classical accounts (could also go in 1(a), above)
     (b) Actual frequency accounts
     (c) Hypothetical frequency accounts
     (d) Propensity accounts
     (e) (Logical probability is sometimes put here . . .)
- I would say that there are in fact many different *theories* of probability, each of which is more or less appropriate for certain applications. See Gillies' good recent book *Philosophical Theories of Probability*.

Interpretations of Probability II

- *Justified* belief and *Rational* belief are not (necessarily) the same thing. In fact, they are distinct. Consider the following counterexamples.
- Justified $\not\Rightarrow$ Rational: I may have overwhelming evidence that my wife has big feet. But, I may also know that I'm the kind of person who can't keep his mouth shut! So, it may *not* be rational for me to believe this — even though it would be *justified*, given my evidence.
- Rational $\not\Rightarrow$ Justified: Someone has offered me 1 million dollars to believe that there are 52,678 marbles in a jar in my attic (and I *love* money!). My evidence suggests there are 52,679 marbles in the jar (I counted them several times, and had others count them, etc.). So?
- The same point holds for degrees of belief. Rational degrees of belief are those which would be most well advised for me to have, all things considered in my complete life, and justified degrees of belief are those that are the most accurate, etc., given my total evidence.

Interpretations of Probability III

- Arguments are made which aim to show that rational degrees of belief and justified degrees of belief should satisfy the probability axioms.
- The "rationality" approaches usually involve either (1) "Dutch Books" which are supposed to show that if you're degrees of belief are not probabilistic, then you are susceptible to a "sure loss" (and, hence, that you are irrational), or (2) "representation theorems" which aim to show that if you're preferences satisfy certain "rationality constraints" (*e.g.*, intransitivity, etc.), then your degrees of belief will be probabilistic.
- The "justification" approaches usually involve some measure of "accuracy" of degrees of belief, and that the "most accurate" degrees of belief will (in general) be those which satisfy the probability axioms.
- All of these arguments have interesting variations and potential problems. See Rosenkrantz's book *Foundations and Applications of Inductive Probability* for a nice discussion of both types of accounts.

Interpretations of Probability IV

- Most subjectivists do not require anything more of an agent's degrees of belief than mere satisfaction of the probability axioms. So, for them, the distinction between rational and justified is of no consequence.
- Most people agree that it is the *justified* degrees of belief that are relevant to *epistemic* applications like confirmation are explanation (if you think those things are purely subjective and epistemic).
- What are *rational* degrees of belief good for? It doesn't seem that rational degrees of belief are relevant to decision making either (shouldn't you act on the basis of your *most accurate* assessment of probability — or on your "happiest" assessment of probability, etc.?).
- I do *not* think that *logical* interpretations of probability belong here. Logical probabilities are quantitative generalizations of notions of deductive logical consequence. See Roeper and Leblanc's book *Probability Theory and Probability Logic* for how this can be done.