Explanation and Unification

Erica Klempner
3/16/04

Michael Friedman, *Explanation and Scientific Understanding*
Philip Kitcher, *Explanatory Unification*

Friedman and Kitcher: same basic motivation, to cash out scientific explanation in terms of *unification*.

I. Friedman

- Explanation of *laws*. Attempts to answer Hempel and Oppenheim's "notorious footnote 33":

    "… Kepler's laws, K, may be conjoined with Boyle's law, B, to a stronger law K&B; but derivation of K from the latter would not be considered as an explanation of the regularities stated in Kepler's laws by themselves. The derivation of Kepler's laws from Newton's laws of motion and of gravitation, on the other hand, would be recognized as a genuine explanation in terms of more comprehensive regularities, or so-called higher-level laws. The problem therefore arises of setting up clear-cut criteria for the distinction of levels of explanation or for a comparison of generalized sentences as to their comprehensiveness. The establishment of adequate criteria for this purpose is as yet an open problem." (*Studies in the Logic of Explanation)*

- Why does heated water turn to steam? – Explanation in terms of molecular theory. One phenomenon "reduced" to another.

    "Thus, the central problem for the theory of scientific explanation comes down to this: what is the relation between phenomena in virtue of which one phenomenon can constitute an explanation of another, and what is it about this relation that gives understanding of the explained phenomenon?" (p6)

- Friedman's analysis of the previous literature's response to this central problem:

    Either: 1. tries to give a precise account of the explanation relation at the expense of saying how explanation gives understanding (e.g. Hempel); or
    2. emphasizes understanding at the expense of a precise account of the explanation relation (e.g. Scriven).

- **Objectivity.** Wants explanation to be *objective* (not pragmatic), and wants an objective notion of "understanding" to go with it. "Objective" for Friedman seems to mean non-relative. Says that Hempel disregarded the importance of understanding because he thought it was a hopelessly psychological or pragmatic notion, varying from person to person. But Friedman thinks we can have a psychological and non-relative notion of understanding – similar to the notion of "rational belief". This notion of understanding will presumably impose *normative* constraints on how explanation is *supposed* to affect a person's actual understanding.

- **Conditions on the explanation relation.**
    1. **Generality**: "most, if not all, scientific theories that we all consider to be explanatory should come out as such according to our theory." (p13) "Familiarity" theories fail.
    2. **Objectivity**: "what counts as an explanation should not depend on the idiosyncracies and changing tastes of scientists and historical periods." (p14) "Intellectual fashion" theories fail (as do familiarity theories, I guess).
    3. **Connects explanation and understanding**: shows "what kind of understanding scientific explanations provide and how they provide it." (p14) Everyone else fails, but particularly Hempel and the DN-guys.

    Given Friedman's logical exposition of the explanation relation (below), possible criticisms do not get a foothold in condition (2) and must focus on (1) and (3).

- **Unification is the answer!**
    E.g. the kinetic theory of gases explains, by *unifying*, three different gas-related phenomena:
    1. Boyle-Charles law
    2. Graham's law of diffusion

3. Specific heat capacities.

And also integrates these gassy phenomena with *other* phenomena like the motion of the planets because the kinetic theory of gases, and Kepler's laws, are themselves derivable from the more general laws of mechanics.

> "[T]his is the essence of scientific explanation – science increases our understanding of the world by reducing the total number of independent phenomena that we have to accept as ultimate or given." (p15)

II. Kitcher

- More general than Friedman: not confined to explanation of laws.

- Seems to take on board Friedman's 3 conditions on the explanation relation. Puts forward a further condition not explicitly mentioned by Friedman, but clearly related to his account:

    4. **Explanatory power**: "Embryonic theories are often defended by appeal to their explanatory power. A theory of explanation should enable us to judge the adequacy of the defense." (p508)

    I.e., I take it that Friedman could say that, if S is "more unifying" than T (i.e., in terms of his technical apparatus, if K-card(Y $\cup$ {S}) < K-card(Y $\cup$ {T}) – see below), then S has more explanatory power, given the Y-phenomena, than T.

- Against, e.g., van Fraassen: "explanatory power as a special virtue of theories is a myth." According to him, Kitcher says, theories are virtuous only insofar as they are *empirically adequate* and *simple*.

    How do *unifying* and *simple* differ?
    – Note that these "virtues" have been thought "superempirical" by realists, who seem therefore to have also thought that, given that they *are* (presumably) virtues, principles of unification/simplicity can lend support to scientific realism.
    – (Branden says that) when *predictive accuracy* is taken to be the prime virtue, where this is conceived of statistically as "goodness of fit" of theory to (hypothetical) data, the statistical results show that better predictive accuracy |– more unified/simpler – i.e., these virtues come along for free: they are empirical and not superempirical after all. (– Where does this leave realism?)

III. Friedman's technical details:

- K = {accepted law-like sentences}   (K is deductively closed: i.e., if K |– S, then S $\in$ K.)

- A *partition* of a sentence S is a set P of sentences s.t. P is logically equivalent to S and each S′ in P is *acceptable independently* of S. (Friedman leaves "independent acceptability" imprecise, but it is some kind of evidential notion: X and Y are independently acceptable iff there are "sufficient grounds for accepting one which are not also sufficient grounds for accepting the other." – p16.)

- S is *K-atomic* if it has no partition (in K).
- The *K-partition* of a set of sentences Y is a set Z of K-atomic sentences s.t. Y$\leftrightarrow$Z.
- The *K-cardinality* of Y is the number of members of the smallest K-partition of Y.
- S *reduces* Y if K-card(Y $\cup$ {S}) < K-card(Y).
- $con_k(S)$ = {independently acceptable consequences of S}

- (D1) S_1 ***explains*** S_2 iff $S_2 \in con_k(S_1)$ and S_1 reduces $con_k(S_1)$.

    E.g., Newton's laws explain Boyle's law because Newton's laws reduce their independently acceptable consequences: Boyle's law, Kepler's laws, etc. However, the conjunction of Boyle's law and Kepler's laws

doesn't explain Boyle's law because the conjunction doesn't reduce {Boyle's law, Kepler's laws}. ("Notorious footnote 33" apparently quieted.)

- "Weakens" D1 in order to safeguard the explanatory power of a good explanation conjoined to an irrelevant law (if S_1 explains S_2 and S_3 is some independent – and irrelevant – law, then S_1 & S_3 will not explain S_2, as S_1 & S_3 will not reduce $con_k(S_1$ & $S_2)$):

  (D1′) S_1 explains S_2 iff there exists a partition Γ of S_1 and an $S_i \in \Gamma$ such that $S_2 \in con_k(S_i)$ and S_i reduces $con_k(S_i)$.

  But this turns out to be different, rather than weaker, as (we'll get to this in a minute) Kitcher has shown that D1 entails that only K-atomic sentences explain, whereas D1′ entails that no K-atomic sentences explain. Salmon thinks D1′ should just be jettisoned.

- Either way, the explanation relation is not 2-place, between one phenomenon and another that explains it. It relies on more "global" aspects of the scientific enterprise, relating both phenomena to the "total set of accepted phenomena".

IV. Questions for Friedman:

1. **Explanation of particular events.** How does Friedman's account of the explanation of laws bear on explanations of particular events (if at all)? Is a completely separate account needed for events? Or can Friedman's framework do duty for events as well?

2. **Only K-atomic sentences explain.** Kitcher proves that this follows from D1 (see attached for the proof). Then he objects that often we use conjunctions of laws (not K-atomic) to explain complex phenomena (e.g. the adiabatic expansion of an ideal gas; why lightning is followed by thunder). So Friedman seems to violate his generality condition on the explanation relation.

3. **But can there be any K-atomic sentences?** – Salmon thinks maybe not: it seems that you can partition any law-like generalization.
- (1) $(x)[Fx \supset Gx]$
  $\equiv$ (2) $(x)[Fx$ & $(Hx \lor \sim Hx) \supset Gx]$
  $\equiv$ (3) $\{(x)[Fx$ & $Hx \supset Gx], (x)[Fx$ & $\sim Hx \supset Gx]\}$
  where the members of the partition set in (3) are – at least on the face of it – acceptable independently of (1).

  E.g., "All humans are mortal" gets partitioned as {"All men are mortal", "All women are mortal"}.

- Kitcher's defense of Friedman: such partitions will not, in general, be law-like – i.e., either H or ~H will not be *projectable* (Hempel and Oppenheim: *projectable* = "purely qualitative"; Kitcher (taking from Quine): *projectable* = "natural kind").

- Salmon's responses: (a) That's just not plausible ("women" is projectable); (b) In any case, we can base the partition of (1) on an exclusive and exhaustive disjunction $H_1, H_2, \dots H_n$ s.t. for each $i \leq n$, $(x)[Fx$ & $H_ix \supset Gx]$, and *now* surely each H_i is projectable. (Kitcher can of course just deny this.)

- Friedman's own response to Salmon:
  "The problem you raise … is just the *kind* of problem I had in mind in requiring that a partition of a sentence S consist of sentences that are acceptable independently of S. In your case, it would seem in general that our ground for accepting the two conjuncts is just that they follow from 'All F are G.' If, on the other hand, we have grounds for accepting the two conjuncts independently – by testing for the two conjuncts directly, say – then it would seem that 'All F are G' is in no way an explanation of the two, but just a summary of what we already know." (Salmon, *Four Decades of Scientific Explanation*, p97)

I.e., either (a) the two conjuncts aren't acceptable independently of 'All F are G' – which seems to be the case in Salmon's original formulation of the problem; or (b) they are.
If (a): 'All F are G' *is* K-atomic. (And Salmon's claim is false.)
If (b): 'All F are G' *is not* K-atomic and therefore (given Kitcher's result) can't explain. It is merely a "summary" of its consequences.

But this is extremely puzzling. Recalling D1, for S to explain T, T must be a member of the set of independently acceptable consequences of S. But given that only K-atomic sentences explain, Friedman's only hope for explanation lies in case (a). But, by hypothesis, in case (a) the conjuncts *aren't* independently acceptable. So Friedman has ruled out explanation altogether.

Is he saved if he just drops the requirement that $con_k(S)$ = {*independently acceptable* consequences of S}?

V. Kitcher's technical details

- Happy with explanations as *arguments*. But wants an objective characterization of what *features* of some arguments – the explanatory ones – enable them to provide explanations. (Fill in the sufficient conditions left vacant by Hempel.)

- **Basic idea** is that science provides a number of different *argument patterns* that can be drawn on in explanation (as in Kitcher's example of Galileo and the fusilier).

- K = {accepted sentences} (not necessarily law-like, as with Friedman)
- E(K) = the *explanatory store* over K = "the set of arguments acceptable as the basis for acts of explanation by those whose beliefs are exactly the members of K" (p512)

  What set of arguments is "acceptable"?
- *schematic sentence* = sentence with some nonlogical expressions replaced by dummy letters
    - e.g.:   "Organisms homozygous for the sickling allele develop sickle cell anemia."
    - can be: "Organisms homozygous for A develop P."
    - or: "For all X if X is O and A then X is P."

- *filling instructions* = directions for how to fill in the dummy letters
    - e.g.:   A = an allele, P = a phenotypic trait

- *schematic argument* = sequences of schematic sentences
- *classification* = a set of sentences describing the inferential characteristics of a schematic argument: which sentences of the argument are premises, which conclusions, what rules of inference are used
- *argument pattern* = <schematic argument, set of sets of filling instructions (one for each dummy letter), classification>

- Argument *stringency* determined by two constraints:
    1. conditions on substitution of dummy letters, imposed by:
        a. nonlogical expressions in the pattern
        b. filling instructions
    2. conditions on logical structure imposed by the classification.

- Kitcher's proposal is that E(K) = {arguments that maximally unify K}, where "unifying power is achieved by generating a large number of accepted sentences as the conclusions of acceptable arguments which instantiate a few, stringent patterns." (p520)

- This means that the unifying power of a set of argument patterns is proportional to:
    1. the total number of *consequences* of all arguments generated by the argument patterns;
    2. the *stringency* of the argument patterns;

and inversely proportional to:

      3.   the number of argument patterns in the set;

although Kitcher doesn't attempt to rank these criteria.

He also wants to impose a further criterion in terms of *similarity of argument patterns*: i.e., the more the patterns can themselves be "unified" by seeing them as instantiations of even more abstract patterns, the greater their unifying power.

- Kitcher illustrates how his theory fares with some traditional problems:
    1. **asymmetry**:   pendulum's length and period; flagpole and shadow.
    2. **irrelevance**:   hexed salt dissolving.
    3. **accidental generalizations**:   e.g. Horace is bald because he's a member of the Greenbury School Board and all members of the Greenbury School Board are bald.
    4. **spurious unification**: notorious footnote 33.

VI. Questions for Kitcher

- Many worries that his explanation-as-unification leaves the relevance of cause/mechanism/general ontological significance in the lurch:

    1. (Woodward) Kitcher's "solution" of the flagpole asymmetry problem depends on the contingent truth that objects don't cast enough shadows. But if they *did*, we presumably still wouldn't say that we could explain the height of the flagpole in terms of the length of its shadow. (-> even clearer with, e.g., Newtonian mechanics in a closed system, say the solar system, and prediction/retrodiction.)

    2. (Woodward) Permits formal/mathematical unifications of phenomena which do not seem to provide the basis for genuine explanations of those phenomena (but which are "merely descriptive") – e.g. biological classification schemes.

- A further Woodward worry: explanation-as-unification seems to demand a "winner take all" conception of explanation – i.e., the only options are explanatory/unexplanatory, and don't leave room for more explanatory/less explanatory, on a continuum.