1.32 Pattern Subsumption

According to the pattern subsumption approach, a good explanation shows that the phenomenon to be explained is an instance of a more general pattern. To understand an event, then, is to see that it is an instance of a pattern of events. To understand a law is to see that it is an instance of a pattern of laws; in the Newtonian explanation of Kepler's laws, for example, the elliptical orbits of the planets are shown to be just one instance of the many kinds of motion generated by the inverse square law of gravitation.

Pattern subsumption accounts of explanation have been offered by Hempel (in some moods), Cartwright (1983), and, under the rubric of the unification approach, by Friedman (1974) and Kitcher (1981, 1989). (The distinctive features of unification accounts are identified below.) Salmon's SR account of explanation might also be interpreted as a pattern subsumption view (section 1.34).

Not every pattern explains: seeing that a black object has its blackness in common with other black objects in no way explains the blackness. The paradigmatically explanatory pattern rather associates one property with another. Patterns of this sort are captured by universal generalizations with an if/then or *All Fs are G* structure: *All ravens are black*, for example, or *If an apple falls from a tree of height d at time t it will hit the ground at time* $t + \sqrt{2d/g}$ (where g is gravitational acceleration). That the consequent of such an if/then pattern obtains in some particular instance is explained by seeing that it is accompanied by the antecedent, and that there is a general pattern of such antecedents being accompanied by such consequents. In the simplest case, then, an object's G-ness is explained by pointing to, first, the object's F-ness, and second, a general pattern of G-ness in Fs.

The formal requirements of the DN account provide a simple framework for representing subsumptions. The following DN explanation, for example, represents the "simplest case" from the end of the last paragraph:

All *F*s are *G*
x is an *F*
---
x is a *G*

As you can see, the "law" states the pattern and the deduction demonstrates that the explanandum, *x*'s *G*-ness, falls under the pattern. It is quite straightforward, then, to interpret the DN account of explanation as a pattern subsumption account. (To do so, it must be argued that the DN form is capable of representing subsumption under any explanatory pattern; in what follows, however, I will for simplicity's sake confine my attention to patterns of the if/then variety typically found in event explanation.)

The pattern subsumption interpretation illuminates two aspects of the DN account that on the nomic expectability interpretation remained obscure. First, it is quite clear on the pattern subsumption interpretation why something like a law must figure in an explanation: an explanation must cite a pattern, and, on certain empiricist views of the nature of laws, including that presented in Hempel and Oppenheim (1948), to state the existence of a pattern just is to state a law, and vice-versa.

Second, the pattern subsumption interpretation of DN explanation carries no implication of epistemic relativity. For an explanandum to be an instance of a pattern, the pattern must be real, but it need not be known to be real. Thus on the pattern subsumption interpretation, it is clear why the premises of a DN explanation must be true, but need not be known to be true.

The DN account's formal requirements give a satisfactory account of which states of affairs instantiate which explanatory if/then patterns, but they say too little about which if/then patterns are explanatory. A genre of problem cases devised by Salmon, Kyburg and others shows that not every if/then pattern—not even of the simplest, most straightforward, *All Fs are G* variety—is explanatory.

Consider, for example, the following scenario (Kyburg 1965). A sample of salt is "hexed" by intoning some magical-sounding words in its vicinity, and is then placed in water. The salt dissolves. The dissolving is just one instance of a general pattern of dissolving: all salt that is hexed in this way dissolves. Thus, if no restrictions are placed on explanatory patterns, the dissolving can be explained by observing that the salt was hexed and placed in water. But this cannot be right. The dissolving is explained by the salt's being placed in water, but not by its being hexed. The unqualified DN account fails to distinguish between two elements in the pattern: an element that is genuinely explanatorily relevant, the placing in water, and an element that is not relevant, the hexing. Salmon calls this simply the problem of relevance.[2]

To solve the relevance problem, a pattern subsumption account must put some restriction on the if/then patterns that are to qualify as explanatory. Nancy Cartwright's simulacrum account of explanation shows how this can be done (Cartwright 1983). Cartwright's account brings the additional benefit of exemplifying an approach on which non-linguistic scientific constructions can function as explanations.

A simulacrum explanation begins with a scientific model, which may be understood as representing the sort of if/then pattern introduced above. (Note that *model* here is used in its informal, scientific sense.) Think of the model as having "inputs" and "outputs", where the inputs represent initial or other boundary conditions of the modeled system and the outputs represent the behavior of the system in the conditions represented by the inputs. A model for the apple generalization above, for example, will take the distance fallen as an input and produce the time taken as an output.

Almost any formal construct may serve as a model: a universally quantified sentence, such as the apple generalization or "All gases expand when heated"; a set of equations, such as the ideal gas law, $PV = nRT$; a full-blown theory, such as statistical mechanics; a computer simulation; or a mechanical representation

2. The relevance problem affects the expectability version, as well as the pattern subsumption version, of the DN account. For attempts to deal with the problem that could be appropriated by a Hempelian, but which make no explicit appeal to pattern subsumption, see Salmon (1979) and Fetzer (1981).

such as an orrery, a working model of the solar system. Understood correctly (and this may require a certain interpretative finesse—or an instruction manual), any of these models will represent a pattern of input/output pairs, which I call the pattern *generated* by the model.

For the purpose of a simulacrum explanation, the internal workings of the model need not correspond to anything in the represented system: "A model is a work of fiction" (Cartwright 1983, 153). Thus the explanatory power of a model is determined entirely by the pattern it generates; nothing else counts. Further, as you will see shortly, even this pattern need not precisely reflect what is found in the real world, though the more closely it resembles a real world pattern, the better.

A model explains a particular output just in case it replicates that output. For example, an apple's taking 1.4 seconds to fall 10 meters is explained by showing that, in or according to the model, an apple takes takes 1.4 seconds to fall 10 meters. I understand simulacrum explanations as proceeding by pattern subsumption: a model encapsulating the apple generalization—*If an apple falls from a tree of height d at time t it will hit the ground at time* $t + \sqrt{2d/g}$—represents a pattern of distances fallen and times taken in real apples, and explains particular distance/time pairs by subsuming them under the pattern.

The resources to deal with the relevance problem are provided by a ranking that Cartwright imposes on models, and therefore, on explanatory patterns: "the [explanatory] success of the model depends on how much and how precisely it can replicate what goes on" (also p. 153).

There are two parts to the ranking, which I will call the criteria of accuracy and generality. The pattern generated by a model is *accurate* to the degree that it reflects a pattern in the real world, that is, to the degree to which the input/output pairs generated by the model match the relevant actual input/output pairs.[3] The pattern generated by a model is *general* in proportion

3. You might measure accuracy either by the match between pattern and the whole world, or by the match between the pattern and the parts of the world that are to be explained. Nothing

to the number of actual input/output pairs it generates. To put it another way, a model is general in proportion to the frequency with which the inputs it accepts—the cases that satisfy the antecedent of the if/then—are realized in the actual world. Its accuracy then measures the degree to which the outputs generated by the model given these inputs match the real world outputs. The generality of the apple generalization, for example, depends on the number of actual apples that actually fall in our world—the more falling apples, the greater the generality. The generalization's accuracy depends on the accuracy with which it yields the actual time taken by these apples to fall. Factors such as air resistance, then, slightly diminish the generalization's accuracy.[4]

There are two ways to interpret the doctrine that a more accurate and general model is a more explanatory model. You might say that an accurate, general model has great explanatory power simply because it generates a pattern with many actual instances, that is, actual input/output pairs, thus because it explains a great number of phenomena. Or you might say that, in addition, a accurate and general model explains each one of those phenomena better—that is, that, all other things being equal, of two models both generating a given explanandum, the more accurate and general model provides the better explanation *of that very phenomenon*. It would follow that there is a certain holism to the explanatory enterprise: the success with which a pattern explains one particular phenomenon depends on its ability to successfully subsume many other phenomena. This second interpretation is by far the more interesting and fruitful, I think; I take it that it is what Cartwright intends.

Some versions of the pattern subsumption account mandate an even stronger claim. According to what Woodward (2003, §8.8) aptly calls the "winner-takes-all" approach, the pattern that maximizes generality and accuracy is not just the *best* explanation of the phenomena it subsumes: it is their

too important turns on the choice; I note, however, that when the kairetic account invokes a notion of accuracy, it is of the latter variety.

4. For more on accuracy, generality, and other properties of models, see Weisberg (2003).

only explanation.

Salmon's relevance problem can now be solved. The difficulty, recall, was that the pattern articulated by the universal generalization

1. All hexed salt dissolves when placed in water

is at best a flawed explanation of its instances. The simulacrum account shows why. Any instance of hexed salt's dissolving in water is not only an instance of (1), but also of

2. All salt dissolves when placed in water.

Since (2) is considerably more general than (1)—it makes predictions about all salt, not just all hexed salt—and equally accurate, it provides a better explanation than (1) of the dissolution of any particular sample of hexed salt. Further, on a winner-takes-all version of the simulacrum account, an explanation that cites hexing is not merely explanatorily dominated; it has no explanatory power at all—a preferable conclusion, I think.

The pattern subsumption account is still, however, incomplete. Return to the case of hexed salt. Model (2) above—the "theory" that all salt dissolves in water—is general, but not as general as the theory obtained by conjoining it with another true, but irrelevant generalization, say *All ravens are black*. That is, the "theory" *All salt dissolves in water and all ravens are black* is more general, because generating more actual phenomena (dissolving salt *and* black ravens) than the dissolution law alone. Why, then, is it not a better explanation of any particular episode of dissolution?[5]

Such examples show that there is still more to say about explanatory patterns. Why is the set of events generated by *All salt dissolves in water and all*

5. This problem was first noted by Hempel himself. He provides a purely syntactic solution for the case in which the explanandum is an event in the original DN paper (Hempel and Oppenheim 1948, §7). He is, however, unable to find a formal solution for the case in which the explanandum is a law (Hempel and Oppenheim 1948, n33); see section 7.1. The solution for events was later challenged, then amended to meet the challenge, and so on. See Salmon (1990b), §1.1 for a short history.

ravens are black not an explanatory pattern? Prima facie, because it is not a pattern at all: there is little or nothing that dissolving salt and black ravens have in common, so there is no single pattern instantiated by all and only salt and ravens. It seems, then, that a further requirement should be added to the theory of explanatory patterns: the phenomena generated by a model do not constitute an explanatory pattern unless they resemble one another in some respect, or as I will say, unless the putative pattern is *cohesive*.[6]

What sort of resemblance is required? Intuitively, a pattern generated by a model is cohesive only if the output of each input/output pair is related to the input in the same or a similar way. That is, there should be a single rule, presumably embodied by the model itself, that determines the output for each input. There should also be a single rule that determines what counts as a possible input, that is, that determines the scope of the model. All systems to which the model applies should, in other words, have single, shared nature. It is in virtue of this second aspect of the cohesion requirement that a pattern such as that specified by *All ravens and lumps of coal are black* is considered insufficiently cohesive for explanatory duty. In itself, this is hardly a substantive account of cohesion, without some further account of what constitutes a "single" rule. No doubt a standard for similarity, among other things, would be needed to fully characterize cohesion, bringing with it the usual problems (Goodman 1971). My aim here, however, is not to define a pattern subsumptionist notion of cohesion, but simply to point to its importance. (I will have more to say about cohesion in sections 3.6 and 7.34.)

When the simulacrum account is augmented by adding a cohesion desideratum, the result is the following sophisticated pattern subsumption account of explanation. A phenomenon is explained by showing that it is generated by a model that optimizes the following desiderata:

---

6. A somewhat different attempt by a proponent of the pattern subsumption approach to say why the subsumption of a dissolution event under the salt/ravens law is not explanatory is Kitcher's desideratum of *stringency* (Kitcher 1981, §8). The stringency of a model is a complex property; I will not try to define it here.

1. Generality: maximize the number of actual phenomena that qualify as inputs.

2. Accuracy: maximize the goodness of fit between generated outputs and actual outputs.

3. Cohesion: maximize the degree to which the phenomena generated by the model share a resemblance both in their inputs and in the way that their outputs are related to their inputs.

There is a point where the drive for cohesion and accuracy acts in opposition to the drive for further generality. This is the point at which accuracy, cohesion, and generality are jointly maximized; the optimally explanatory subsuming pattern is found here.

An additional desideratum is sometimes added to the three above:

4. Simplicity: maximize the simplicity of the model.

Unlike the other desiderata, this is, at least on the face of things, a condition on the model rather than on the generated pattern.

Observe that the simplicity desideratum duplicates some of the work done by the cohesion desideratum, because simpler models tend to generate more cohesive sets of phenomena. But neither desideratum renders the other redundant. Of two equally cohesive models, one may be simpler, and of two equally simple models, one may be more cohesive.

A pattern of great generality, accuracy, and cohesion is a pattern that has considerable unifying power: it brings together many perhaps disparate-seeming phenomena and shows that they fit a single template. If explanation is a matter of pattern subsumption, then, it is thereby also a matter of unification. For this reason I classify unification accounts of explanation (Friedman 1974; Kitcher 1981, 1989) as a species of the pattern subsumption approach.

Is there anything that sets unification views apart from other sophisticated pattern subsumption views? Two things. First, unification accounts tend to impose a simplicity desideratum. Second, and more importantly, unification accounts tend to require even more holism than is already inherent in the pattern subsumption approach. On Kitcher's unificationism, for example, a model explains a phenomenon that it generates just in case it belongs to the set of models that best unifies (that is, optimizes accuracy, generality, cohesion, and simplicity over) *all* the phenomena. A lesser holism might require unification of phenomena only on a domain by domain basis. Kitcher has weakened his holism in this way in recent years.

1.33 Causality

The explanatory relation is a causal relation, according to the causal approach. On a two-factor causal view, such as the kairetic account advocated in this study, the explainers of a phenomenon must both bear the causal relation to the phenomenon and pass a test for explanatory relevance. On the prevalent one-factor approach, the explainers of the phenomenon must simply bear the causal relation to the phenomenon.

One-factor accounts differ in many ways in their conception of the nature of the causal relation. Most obviously, they differ in their selectivity. A completely non-selective causal view classifies as explainers of a phenomenon anything that, from the point of view of fundamental physics, exerts a causal influence on the phenomenon, up to and including the gravitational influence of the distant stars. Salmon (1984)'s causal theory of explanation is of this sort.

A selective causal view, while perhaps allowing that the ever-present influence of the distant stars constitutes a causal relation of a sort, holds that the kind of causal relation required for explanation is a far rarer thing. Of the myriad physical influences on a phenomenon, only a few will count as the phenomenon's *causes*, and so qualify for admission to its explanation. Lewis (1986a) and Woodward (2003) offer relatively selective views of causal explanation.

Two-factor approaches can also have their selectivity assessed. The kairetic account begins with an extremely non-selective view of the causal relation, but