

# Statistical Relevance and Explanation

Wesley Salmon

# Salmon: Flaws of Hempel's I-S view

## Overview:

1. Allows **irrelevant** factors to count as explanatory. (conditions *insufficient*)
2. Does not impose **directionality** on explanation. (conditions *insufficient*)
3. An explanation does not have to confer **high probability** on explanandum. (condition *unnecessary*)
4. An explanation is not an **argument**. (condition *unnecessary*)

# Irrelevant Factors

Consider again:

John took his wife's birth control pills.

$P(\text{not pregnant}/\text{take bcp}) = .95$

===== .95

John did not get pregnant.

Intuitively: not an explanation.

Apparently: Hempel's theory implies it is.

Can Hempel escape?

# Cause and Explanation

## Relevance

**Cause vs. indicator of cause** as explanans

Today the barometer suddenly read low.

$P(\text{storm/low bar. reading}) = .9$

=====

There was a storm today.

## Directionality

1) **Cause vs. effect:** as explanans

Flagpole shadow

Flagpole height

geometry

-----

Sun height

# Directionality of Explanation

## 2) Time: **Retrodiction**

We can use the laws of motion to predict where given objects in the solar system were at time  $t$  given where they were at an earlier time  $t - n$ . (Think of a solar eclipse.) Where they *were* plus the laws do explain where they are now.

Hempel's D-N model captures this.

But we can use the same equations in time-reverse. Hempel's theory must count that as an explanation of the past positions of the objects (why?), and it's not (why?).

# Relevance, not high probability

John has latent, untreated syphilis.

$P(\text{ paresis} / \text{l.u. syphilis}) = .3$

$P(\text{ paresis} / \text{not-(l.u. syphilis)}) = \text{nil}$

===== .3

John has paresis.

Hempel's theory implies this is not an explanation at all.

Salmon: our intuition that the syphilis is explanatory comes from the fact that it's **relevant**, that is, that it **raises** the probability of paresis.

→ What's important isn't the **high probability** of the conclusion, but the **probabilistic relevance** of the factor cited.

→ Otherwise, how do we explain low-probability events?

# Statistical Relevance - Definitions

The property C is ***statistically relevant*** to B within A if and only if  $P(B/A.C) \neq P(B/A.-C)$ . (Otherwise, statistically irrelevant.) *Note:* Statistical relevance is symmetric.

A ***partition*** of a class is a set of mutually exclusive, exhaustive subclasses of that class.

A property C may partition a class A by dividing it into the two subclasses A.C and A.-C.

Suppose we are judging the probability of B, given other properties. (B is going to be the explanandum.) A partition (such as C in class A) is a ***place selection*** iff we can tell whether the members have C without knowing whether they have B.

# Statistical Relevance - Definitions

A is a **homogeneous reference class** (with respect to B) if and only if every property determining a place selection in A is statistically irrelevant to B.

This implies that there is no way in principle (no property) to effect a partition of A that is statistically relevant to B without using B itself.

In other words, within A there is no other property than B that makes B more or less probable.

**Epistemically homogeneous** – homogeneous as far as we know

**Practically homogeneous** – even though we may know a property for partitioning exists we are unable to use it or it is too much trouble.

Note: There are *many* homogeneous reference classes in A for B. If A is homogeneous for B, then so is every subclass of A. (Do Venn Diagram<sub>8</sub> of John and birth control pills.)

# Explanation - Salmon

We explain the presence of property B on or with property A on an individual by:

- 1) noting  $P(B/A)$
- 2) Partitioning A via a property C.
- 3) Noting  $P(B/A.C)$
- 4) If  $P(B/A.C) > P(B/A)$  then C ***explains*** B.
- 5) We continue partitioning with more properties, until we reach a homogeneous reference class. Then we have a full explanation and we stop. (Using another property to go to a smaller homogeneous subclass would introduce irrelevance.)

# Example – Salt Solution

Suppose we have just witnessed salt dissolving in water. We ask: Why does this stuff dissolve in water?

Attribute (B): dissolves in water

Reference Class (A): substances

Suppose  $P(B/A) = .2$

Suppose C = being table salt.

The presence of property C increases the probability of B in A, from .2 to ~1. I.e.  $P(B/A.C) > P(B/A)$ . Thus, C explains B.

(Redo with D = being hexed)

# Stat. Rel. Explanation Applications – high probability issue

Q: Why does John have paresis?

A: He has (untreated) syphilis.

Salmon wants the theory of explanation to imply that this answer is explanatory.

On Salmon's stat. rel. theory, this means we want:

though the probability of paresis given syphilis is not high, it is higher than the probability of paresis (without syphilis).

Verify.

Does the property of syphilis give a homogeneous reference class?

# Stat. Rel. Applications – relevance: causation

How does the SR theory apply to the barometer case? Barometer reading is an indicator of the storm, but not a cause, and not an explanation?

Salmon: we must employ a further probabilistic relevance notion to see the difference between a cause and an indicator.

**D screens off C from B in A iff**

$$P(B/A.D.C) = P(B/A.D.-C)$$

*Note:* Screening off is not symmetric.

# Stat. Rel. Applications – relevance: causation

B = storm happens

A = it is today in temperate region X

C = barometer reads low

D = pressure is low

How do we apply the screening off notion to represent the fact that the barometer reading doesn't do anything to explain *why the storm came?*

**A cause screens off an indicator. An indicator does not screen off a cause.**

**Screening-off Rule:** When a property D screens off another property C (from B in A), then C is not part of the explanation of why B is the case.

So, C **explains** why the A is a case of B iff C is statistically relevant to B (in A) and no property screens off C from B (in A).

# Stat. Rel. Application— directionality: dependency

The shadow length and flagpole height do not explain the Sun height, but the Sun and flagpole height do explain the length of the shadow.

This is because the shadow is dependent on the Sun and flagpole, but not vice versa. Can we use the probabilistic relevance notions to capture this asymmetry?

SL = Shadow length

FH = Flagpole height

SH = Sun height

# Stat. Rel. Application— Directionality in time

**Retrodiction** – Where the Sun and Moon were, plus the laws of physics, together explain why and eclipse happened today. How do we capture that reversing the equation in time does not provide an explanation?

The present and future don't explain the past because the past does not *depend* on the present or future. The future does depend on the past and present, which is why the latter explain it.

Can we capture this asymmetry in terms of statistical relevance?

# Determinism and Indeterminism

**Determinism:** Every event is dictated to be exactly as it is by previous states of the universe plus the laws of nature.

**Indeterminism:** Not every event is dictated to be exactly as it is by previous states of the universe.

Salmon: Hempel's requirement that the explanans make the explanandum highly probable gives a misleading picture. We know that the world is indeterministic, and some events just are **not** probable! Nothing in the past dictates that they must be as they actually turn out. Why should we require of an explanation that it give grounds for *expecting* this improbable event that was undetermined by the past?

Consider a carbon atom. Its probability of decaying in the next year is low. If it *does* decay it will not be appropriate to explain that by something that makes that outcome probable. Nothing does make it probable!

What are Hempel's possible avenues of reply?

# Stat. Relevance and Functional Explanation

A **functional explanation** cites the function of, say, an organ/feature of the animal's body, to explain why the organ/feature is present.

Q: Why does this animal have hemoglobin in its blood?

A: Because it has oxygen transported in its blood, and hemoglobin transports oxygen.

Problem: The information in A does not imply or make probable the explanandum in Q.

*Rather.*

Hemoglobin transports oxygen.

The animal has hemoglobin in its blood.

-----

The animal has oxygen transported in its blood.

Hempel can't count A as an explanation.

Can Salmon?