A Causal Perspective on Analogical Inferences

Wolfgang Pietsch,
Munich Center for Technology in Society, TU München

Workshop on the History of Statistics, LMU, 22. 3. 2016

Overview

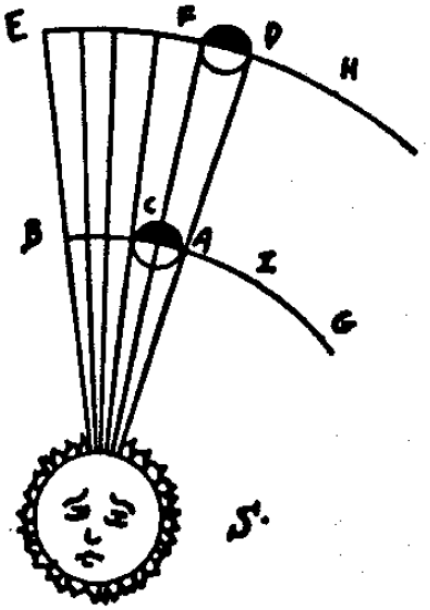
- Analogies in science: Kepler and Maxwell
- Historical perspectives
  - Carnap: the inadequacy of enumerative approaches
  - Keynes: the ubiquity of analogy
  - Hesse: the two-dimensional account
- Two kinds of analogy: predictive vs. conceptual
- A basic intuition: irrelevance of the negative analogy
- A deterministic account of causal irrelevance
- Complications: analogy and probability

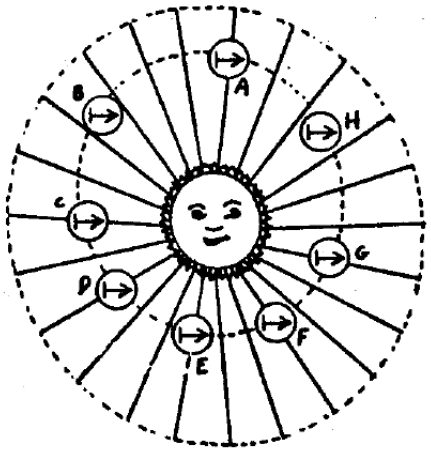
Analogies in science

- “I cherish more than anything else the Analogies, my most trustworthy masters. They know all the secrets of nature” (Kepler 1604)
- “It is by the use of analogies […] that I have attempted to bring before the mind, in a convenient and manageable form, those mathematical ideas which are necessary to the study of the phenomena of electricity.” (Maxwell 1855/56)

Kepler‘s use of analogies

- Central are the light/anima motrix analogy and the magnetism analogy (cp. Gentner et al. 1997)
  - “Since there is just as much power in a larger and more distant circle as there is in a smaller and closer one, nothing of this power is lost in traveling from its source, nothing is scattered between the source and the movable body. The emission, in the same manner as light, is immaterial, unlike odours, which are accompanied by a diminution of substance, and unlike heat from a hot furnace, or anything similar which fills the intervening space.” (Kepler 1609/1992, p. 381)
- Further aspects derived from the analogy: action at a distance and conservation law



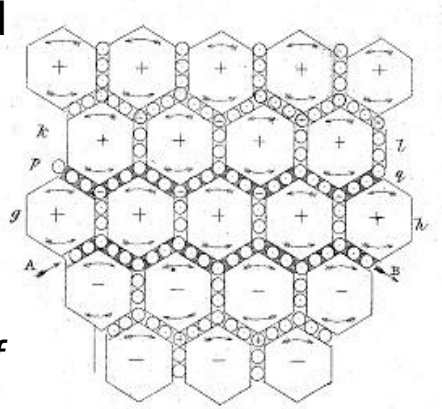


Analogies in Maxwell‘s methodology

- Analogical reasoning was a crucial ingredient of Maxwell’s scientific methodology
  - “We must therefore discover some method of investigation which allows the mind at every step to lay hold of a clear physical conception, without being committed to any theory founded on the physical science from which that conception is borrowed, so that it is neither drawn aside from the subject in pursuit of analytical subtleties, nor carried beyond the truth by a favourite hypothesis.
  - In order to obtain physical ideas without adopting a physical theory we must make ourselves familiar with the existence of physical analogies. By a physical analogy I mean that partial similarity between the laws of one science and those of another which makes each of them illustrate the other. Thus all the mathematical sciences are founded on relations between physical laws and laws of numbers, so that the aim of exact science is to reduce the problems of nature to the determination of quantities by operations with numbers.” (Maxwell 1855/56, 156)

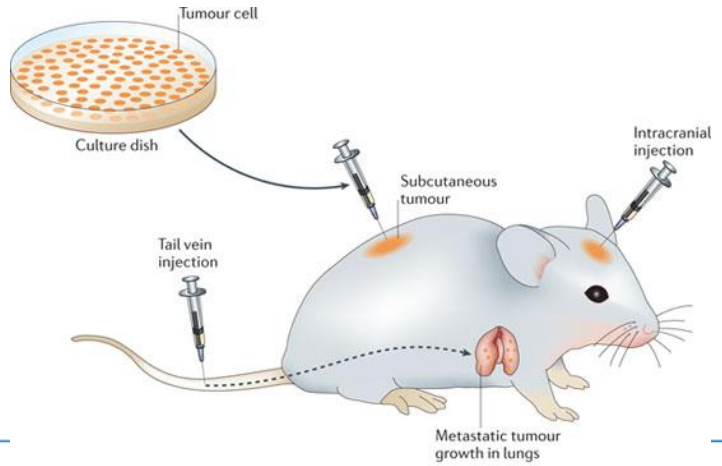
Analogies in Maxwell‘s methodology

- Especially in his studies of electromagnetism, Maxwell relied heavily on analogical reasoning.
- For example, the analogy between conduction of heat and action at a distance
  - “The laws of the conduction of heat in uniform media appear at first sight among the most different in their physical relations from those relating to attractions. The quantities which enter into them are *temperature*, *flow of heat*, *conductivity*. The word *force* is foreign to the subject. Yet we find that the mathematical laws of the uniform motion of heat in homogeneous media are identical in form with those of attractions varying inversely as the square of the distances. We have only to substitute *source of heat* for *centre of attraction*, *flow of heat* for *accelerating effect of attraction* at any point, and *temperature* for *potential*, and the solution of a problem in attractions is transformed into that of a problem in heat. […]
  - It is by the use of analogies of this kind that I have attempted to bring before the mind, in a convenient and manageable form, those mathematical ideas which are necessary to the study of the phenomena of electricity.” (Maxwell 1855/56, 157)



A different kind of analogy

- Both Kepler and Maxwell used analogies primarily for concept development in relatively early stages of a science.
- There is also a tradition in philosophy of science, involving among others Carnap and Keynes, that focuses on the role of analogies for making reliable predictions.
  - e.g. model organisms in biology or medicine



Carnap: the inadequacy of enumerative approaches

- Rudolf Carnap has developed one of the most extensive inductive frameworks in the 20th century, which explicitly aimed to include considerations of analogy.
  - The confidence in a hypothesis h based on evidence e is spelled out in terms of the confirmation function c(h|e).
  - This degree of confirmation c(h|e) corresponds to a logical probability.
  - Carnap defines analogical inferences as follows: “The evidence known to us is the fact that individuals b and c agree in certain properties and, in addition, that b has a further property; thereupon we consider the hypothesis that c too has this property.” (1945, 87)

The continuum of inductive methods

- Given a family of predicates P that applies to a number of individuals a
- Starting from the “straight rule” of induction, according to which the degree of confirmation is the relative frequency s_j/s of a property P_j in the first s individuals
- Extended by Carnap to a λ-γ system:
  - $c_j(s_j, \dots, s_k) = \frac{s_j + \lambda\gamma_j}{s+\lambda}$
  - corresponding to s real and λ virtual individuals; among the latter $\lambda\gamma_j$ have the property P_j
- The confirmation function can be rewritten in terms of an empirical and a logical part:
  - $c_j(s_j, \dots, s_k) = \frac{s}{s+\lambda}\frac{s_j}{s} + \frac{\lambda}{s+\lambda}\gamma_j$
  - For large s, the empirical part dominates; for small s, the logical part (essentially representing prior considerations).

The continuum of inductive methods

- Analogy is treated in terms of the mentioned γ corresponding to the width (or weight) of properties and an additional η corr. to the distance between properties.
  - instances with properties that are closer to the predicted property confirm better than those with more distant properties
  - the individual weight of the properties also has to be taken into account
- In general, the analogy influence
  - belongs to the logical / a priori part of the confirmation function
  - vanishes with a large number of instances

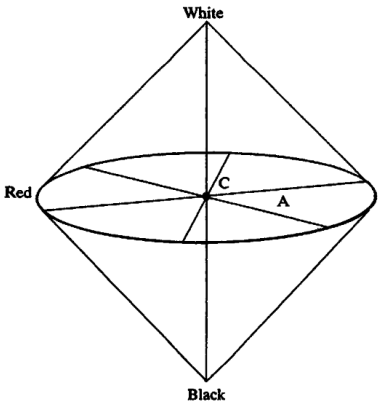


Fig. 14-1. The Color Space.

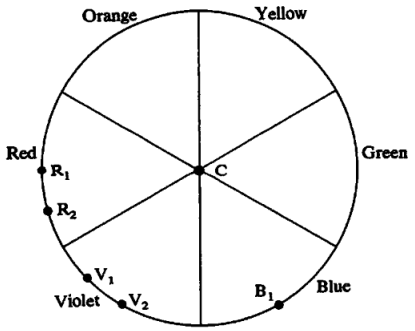


Fig. 14-2. Area A of the Color Space.

The continuum of inductive methods

- Some problems for Carnap‘s approach to analogy
  - There is wide-spread doubt whether his framework can cover all kinds of analogical reasoning, e.g. analogical inferences in the presence of a negative analogy (e.g. Hesse 1964)
  - The additional parameters seem somewhat ad hoc and the values are in general not fully determined by the evidence.
  - Analogy is usually confined to prior considerations and washes out as increasing evidence is gathered.

The continuum of inductive methods

- Some problems for Carnap‘s approach to analogy
  - There is wide-spread doubt whether this framework can cover all kinds of analogical reasoning, e.g. analogical inferences in the presence of a negative analogy (e.g. Hesse 1964)
  - The additional parameters seem somewhat ad hoc and the values are in general not fully determined by the evidence.
  - Analogy is usually confined to prior considerations and washes out as increasing evidence is gathered.

Keynes: the ubiquity of analogical reasoning

- There is a tradition of writers including maybe most notably John Stuart Mill and John Maynard Keynes who claim a much more substantial role for analogy in inductive reasoning
  - „I repeat that by emphasising the number of the instances Hume obscured the real object of the method. If it were strictly true that the hundred instances are *no* way different from the single instance, Hume would be right to wonder in what manner they can strengthen the argument. The object of increasing the number of instances arises out of the fact that we are nearly always aware of *some* difference between the instances, and that even where the known difference is insignificant we may suspect, especially when our knowledge of the instances is very incomplete, that there may be more. Every new instance *may* diminish the unessential resemblances between the instances and by introducing a new difference increase the Negative Analogy. For this reason, and for this reason only, new instances are valuable.“ (Keynes 1920, 233)
- => analogical reasoning based on well-defined similarity is key for inductive inferences rather than the enumeration of identical instances

The ubiquity of analogical reasoning

- Carnap's system implements
  - a clear distinction between enumerative induction and analogy
  - a principle of instantial relevance, which claims that any positive instance strictly increases the probability that the next instance is positive as well ("one of the basic characteristics of customary inductive reasoning")
    - $c(P_j a_{s+2}, e \,\&\, P_j a_{s+1}) > c(P_j a_{s+2}, e)$
- By contrast, Keynes argues that
  - all induction relies on analogy
  - identical instances do not confirm at all

Critique of enumerative approaches

- Pure induction and relatedly relative frequencies in statistics are of dubious value as they combine various instances that all differ individually in circumstances. It is not at all obvious why these instances should be counted with equal weight:
  - „I do not myself believe that there is any direct and simple method by which we can make the transition from an observed numerical frequency to a numerical measure of probability. The problem, as I view it, is part of the general problem of founding judgments of probability on experience, and can only be dealt with by the general methods of induction...” (Keynes 1920, 367)

The ubiquity of analogical reasoning

- According to Keynes, all inductive inferences are analogical inferences
  - „In an inductive argument, therefore, we start with a number of instances similar in some respects AB, dissimilar in others C. We pick out one or more respects A in which the instances are similar, and argue that some of the other respects B in which they are also similar are likely to be associated with the characteristics A in other unexamined cases. The more comprehensive the essential characteristics A, the greater the variety amongst the non-essential characteristics C, and the less comprehensive the characteristics B which we seek to associate with A, the stronger is the likelihood or probability of the generalisation we seek to establish.
  - These are the three ultimate logical elements on which the probability of an empirical argument depends,—the Positive and the Negative Analogies and the scope of the generalisation.“ (Keynes 1920, 219-220)

Keynesian terminology

- Analogical inferences are inferences based on similarity:
  - If two entities, source A and target A*, are similar and A has a property C, what is the probability that A* has C as well?
- Some terminology (cp. Keynes 1921):
  - Positive analogy: those properties which source and target have in common
  - Negative analogy: those properties in which source and target differ
  - Unknown analogy: those properties of which it is unknown whether they belong to the positive or negative analogy
  - Hypothetical analogy: those properties which are known of the source phenomenon and predicted of the target phenomenon

Commonsense Guidelines

- Paul Bartha lists some „commonsense guidelines“, many of which can be found in Keynes‘s work:
  - The more similarities (between the two domains), the stronger the analogy.
  - The more differences, the weaker the analogy.
  - The greater the extent of our ignorance about the two domains, the weaker the analogy.
  - The weaker the conclusion, the more plausible the analogy.
  - Analogies involving causal relations are more plausible than those not involving causal relations.
  - Structural analogies are stronger than those based on superficial similarities.
  - The relevance of the similarities and differences to the conclusion (i.e., to the hypothetical analogy) must be taken into account.
  - Multiple analogies supporting the same conclusion make the argument stronger. (Bartha 2010, 19)

The counting problem

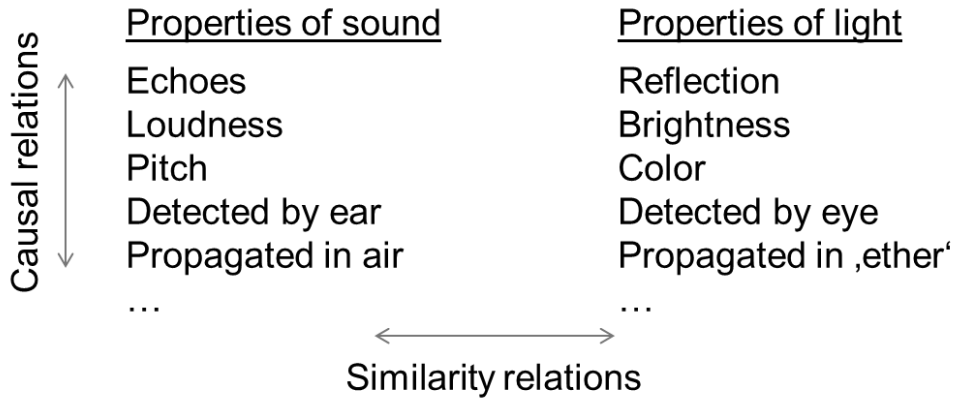
- Keynes‘s framework remains qualitative indicating how probabilities increase or decrease, but largely lacking the possibility of quantitative assessment.
- This problem originates in the shift from enumerative induction to analogical reasoning
  - In enumerative induction, the number of positive instances provides a natural quantitative measure.
  - In analogical reasoning, the positive analogy has to be weighed against the negative analogy.
- Leading to the counting problem: how are properties counted?

Hesse: the two-dimensional approach

- Presumably, there is no general solution to the counting problem.
- Rather, the solution is feasible only in terms of what has since been called the two-dimensional framework:
  - „Under what circumstances can we argue from, for example, the presence of human beings on the earth to their presence on the moon? The validity of such an argument will depend, first, on the extent of the positive analogy compared with the negative (for example, it is stronger for Venus than for the moon, since Venus is more similar to the earth) and, second, on the relation between the new property and the properties already known to be parts of the positive or negative analogy, respectively. If we have reason to think that the properties in the positive analogy are causally related, in a favorable sense, to the presence of humans on the earth, the argument will be strong. If, on the other hand, the properties of the moon which are parts of the negative analogy tend causally to prevent the presence of humans on the moon the argument will be weak or invalid.“ (Hesse 1966, 58-59; cited in Norton 2011, 8)

The two-dimensional approach

- A further example (Hesse 1966, 60):



- A general framework (Bartha 2010, 15):

SOURCE (S)	TARGET (T)	
P	P*	[positive analogy]
A	¬A*	[negative
¬B	B*	analogy]
Q		
---	---	
	Q* (plausibly)	

  - Bartha distinguishes four kinds of vertical relations: predictive, explanatory, functional, correlative

A moral and some doubts

- Approaching the counting problem by means of the two-dimensional framework seems to helplessly contextualize the assessment of analogical inferences.
- Thus, many doubt, whether a general formal framework for analogical reasoning is possible.
  - “Despite the confidence with which particular analogical arguments are advanced, nobody has ever formulated an acceptable rule, or set of rules, for valid analogical inferences.” (Bartha 2013, §2.4)
  - „Argument by analogy is a generally accepted form of inductive reasoning and many think that inductive reasoning can be represented using the probability calculus. From these facts one might expect that there would be accepted probability models that can represent inference by analogy, but no such model exists.” (Maher 2001, 183)

Overview

- Analogies in science: Kepler and Maxwell
- Historical perspectives
  - Carnap: the inadequacy of enumerative approaches
  - Keynes: the ubiquity of analogy
  - Hesse: the two-dimensional account
- Two kinds of analogy: predictive vs. conceptual
- A basic intuition: irrelevance of the negative analogy
- Causal irrelevance, a deterministic account
- Complications: analogy and probability

Predictive and conceptual analogies

	Predictive analogy	Conceptual analogy
Aim	Prediction	Development of a conceptual framework
Laws	Causal	Abstract, partly conventional
Evaluation	Reliable prediction; in terms of truth and probability	Useful framework; *not* in terms of truth and probability (due to underdetermination)
Framework	Carnap‘s continuum; eliminative induction (Mill, Keynes)	Gentner‘s structure-mapping theory (1983)
Example	Mouse model	Maxwell‘s heat/action-at-a-distance analogy

The core intuition

- A (predictive) analogical inference is valid, if and only if the negative analogy concerns only causally irrelevant circumstances.

The core intuition

- A (predictive) analogical inference is valid, if and only if the negative analogy concerns only causally irrelevant circumstances.

- Causation is crucial for predictive analogies, since only causal laws can establish reliable prediction and intervention:
  - “causal laws cannot be done away with, for they are needed to ground the distinction between effective strategies and ineffective ones.” (Cartwright 1979, 420)

Struggling with the notion of causal irrelevance

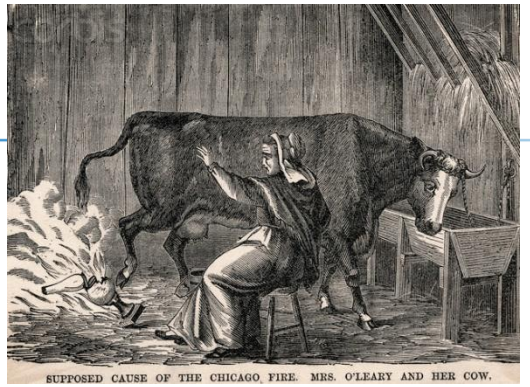
- Problem: the approaches to the notion of causal irrelevance that can be found in the literature are not up to the task
  - based on statistical irrelevance
    - statistical irrelevance seems neither a necessary (in deterministic contexts) nor a sufficient condition (when two influences exactly cancel) for causal irrelevance
  - as a counterpart to typical explications of causal relevance
    - Typical definition of causal relevance: „Factor A is causally relevant for the occurrence of an effect B, if and only if there exists at least one causal process, in which an event of type A (partly) causes the occurrence of an event of type B.“ (Baumgartner & Grasshoff 2004, 49; my translation)
    - A complementary notion of causal irrelevance requires to show that there is no such process. That seems impossible to establish empirically.

A suggestion

- Define both causal relevance and irrelevance with respect to a background or context:
  - In a context B, in which a condition A and a phenomenon C occur, A is causally relevant to C, iff the following counterfactual holds: if A had not occurred, C would also not have occurred.
  - In a context B, in which a condition A and a phenomenon C occur, A is causally irrelevant to C, iff the following counterfactual holds: if A had not occurred, C would still have occurred.
- Causal irrelevance thus is a three-place notion: A circumstance is causally irrelevant to a phenomenon C with respect to background B.

[Pietsch, http://philsci-archive.pitt.edu/11913/]

An example


SUPPOSED CAUSE OF THE CHICAGO FIRE. MRS. O'LEARY AND HER COW.

- (kicked-over lantern ∧ hay on the floor) ∨ lightning cause barn fire with respect to a context of other conditions B* =>
  - Kicked-over lantern is causally relevant with respect to background B* ∧ hay ∧ no lightning
  - Kicked-over lantern is causally irrelevant with respect to B* ∧ no hay ∧ no lightning
  - Kicked-over lantern is causally irrelevant with respect to B* ∧ hay ∧ lightning
  - Kicked-over lantern is neither relevant nor irrelevant with respect to B* ∧ no lightning
  - …

Context dependence

- According to the proposed definitions, both causal relevance and causal irrelevance are always relative to a context or background.
  - Roughly, a background is defined by (i) conditions that are irrelevant to the phenomenon and are allowed to vary; (ii) other conditions that are potentially relevant and remain constant.

- The deeper reason for context dependence is that by eliminative induction, empirical relationships can only be established with respect to a context.

Counterfactuals

- Traditional approaches to counterfactuals (e.g. Reutlinger 2012):
- Metalinguistic approach (e.g. Goodman):
  - A counterfactual of the form ‘If C were true, then A would also be true’ is true if and only if there is an auxiliary set S of true statements (including laws of nature and matters of fact) consistent with the antecedent C, such that the members of S, when conjoined with C, imply the consequent A. (Menzies 2014)
- Based on similarity between possible worlds (e.g. Lewis):
  - ‘If C were true, then A would also be true’ is true (at a world w), iff either (1) there are no possible C-worlds, or (2) some C-world where A holds is closer to the actual world than is any C-world where A does not hold. (Lewis 1973, 560)

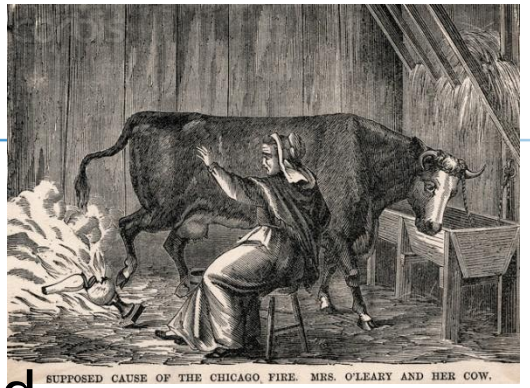
A difference making account of counterfact.

- Take the method of difference as a guiding idea to evaluate the truth value of counterfactuals.
  - Mill's formulation: "If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance save one in common, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon." (Mill 1848)

- Counterfactuals without possible worlds:
  - Show (i) that the examined counterfactual statement belongs to a class of statements with the same truth value, (ii) and that at least one of these statements describes a phenomenon that is observable.

A difference making account of counterfact.

- ‘If A were not the case, C would not be the case’ is true with respect to an instance in which both A and C occur in a context B, if (1) there exists at least one instance in which neither A nor C occurs in the same context B and (2) if B guarantees homogeneity.
  - Context B guarantees homogeneity, iff only conditions that are causally irrelevant to C can change, (i) except for A and (ii) conditions that are causally relevant to C in virtue of A being causally relevant to C.
- A condition X is causally relevant to C in virtue of A being causally relevant to C with respect to a background B, iff in all contexts within B, in which X is causally relevant to C, A is causally relevant to C as well (but not necessarily vice versa).


SUPPOSED CAUSE OF THE CHICAGO FIRE. MRS. O'LEARY AND HER COW.

The example

- Kicked-over lantern is causally relevant to the barn fire with respect to background B* ∧ hay ∧ no lightning
  - Evaluate the counterfactual „if the lantern had not been kicked over…“
  - Find an instance with the same context B* ∧ hay ∧ no lightning that differs only in circumstances that are causally irrelevant to the barn fire, except that
    - the lantern is not kicked over
    - and the causal processes leading to the kicked-over lantern may be different
    - as well as the processes leading from the kicked-over lantern to the barn-fire
- Indeed, the night before, everything was at the same place in the barn except that the cow did not raise its hoof…

Evaluating the causal counterfactual

- One might be tempted to think that the notion of causal irrelevance allows to introduce a similarity measure à la David Lewis, but:
  - comparison of events/phenomena in the actual world, not of different possible worlds
  - the ‘similarity measure’ is not continuous, but concerns only a two-valued function: differs only in terms of irrelevant circumstances or not
  - ‘Similarity’ depends on C and A!
- => not compatible with Lewis’s account of counterfactuals

Analogical inferences

- Main problem for predictive analogical inferences: similarity measure
  - => use the mentioned two-valued similarity measure
- More exactly:
  - Analogical inferences are valid, if and only if the negative analogy (i.e. the ◊-conjunction of all circumstances therein) is causally irrelevant to the hypothetical analogy with respect to a background constituted by the positive analogy.

Analogy and probability

- Possible complications:
  - There may be an unknown analogy that is causally relevant.
    - The analogical inference is valid with the probability that the relevant factor(s) belong to the positive analogy, i.e. are present in the target as well.
  - It may be uncertain, whether the negative analogy is irrelevant.
    - The analogical inference is valid with the probability that the corresponding factor(s) are irrelevant (translates into the probability that a contributing causal factor is absent or an inhibiting factor is present in the background)
  - Indeterminism: the boundary conditions may not fully determine the hypothetical analogy

$\Rightarrow$ In all these cases, the analogical inferences hold only with a certain probability.

$\Rightarrow$ What kind of probability? => better not frequentist

Kinds of analogies

- Several types of analogies are distinguished in the literature, e.g.
  - Similarity vs. identity: whether the corresponding properties of source and target are identical or merely similar
  - Entities vs. relations: whether the similarity can be expressed in terms of properties of source and target or whether one needs to take into account relations between those properties
- The claim is that all these types can be treated within the given framework, if sometimes a reformulation is required.

Conclusions

- A distinction between predictive and conceptual analogies was introduced.
- Several historical approaches to predictive analogies were presented, in particular the frameworks of Carnap and Keynes.
- Only the latter proved adequate for dealing with analogical inferences, but it faces the problem how to weigh the negative and the positive analogy.
- A formal treatment for predictive analogical inferences was proposed: ‚An analogical inference holds, if and only if the negative analogy is causally irrelevant to the hypothetical analogy with respect to a background consisting of the positive analogy.‘
- A counterfactual definition of causal irrelevance is employed, where the truth-values of counterfactual propositions are evaluated in terms of difference-making.