The Quine-Duhem Thesis and Underdetermination

Pierre Duhem (1861 – 1916) **Willard Van Orman Quine (1908-2000)**

1

Duhem against inductivism

- Inductivism (Bernard, Bacon): free our minds when making experiments
- Duhem: This is impossible, especially in physics:
  - we rely on physics in using each measurements apparatuses
  - "The physicist is obliged to trust his own theoretical ideas or those of his fellow physicists. [...] The statement of the result of an experiment implies, in general, an act of faith in a whole group pf theories."

2

Duhem against falisificationism

- Falsificationism/ HD model
  - T predicts O; experiment finds E=/=O; T is falsified
- Duhem: this is impossible
  - when a physical theory T is tested by an experiment, it is not T alone that is tested, but a large collection of theory, auxiliary hypotheses, and assumptions that are tested
  - $T \& A_1, \ldots, A_N \rightarrow O$
  - Finding E=/=O does not imply T is false: at least one of the T, A_1, ..., A_N is false and logic alone does not tell us which
    - HOLISM IN PHYSICS

3

The theory ladenness of observation as a general argument fro holism is physics

- The physicist, in order to connect the predictions of the theory with direct observation, needs to **translate** from the everyday language to the theoretical language
- (theory-ladenness of observation) This translation is affected by using theories about how the measuring apparatus works
- Therefore: (holism) when a physical theory is tested by an experiment, it is not the theory alone that is tested, but a large collection of theory, auxiliary hypotheses, and assumptions that are being put to the test

4

Consequence of holism: crucial experiments in physics are impossible

- E is a crucial experiment between T_1 and T_2 if T_1 predicts that E will give the result O and T_2 predicts that E will have the result not-O:
  - If we perform E and obtain O, T_2 is eliminated
  - If we perform E and obtain not-O, T_1 is eliminated
  - (examples…)
- Duhem:
  - we cannot derive O from T_1 alone
  - So crucial experiments are impossible

5

Underdetermination

- Crucial experiments are impossible
- Theories cannot be falsified by empirical evidence
- The same empirical data is compatible with infinite many incompatible theories
- It is underdetermined which theory is supported/corroborated/confirmed by a given piece of evidence

6

Clarifications about Duhem's view

- His view was restricted to physics
- He attacked just the extreme view that experiments can refute with certainty theories as a matter of logic
- He left open the possibility that experiments (in conjunction with other considerations) could lead rationally to the rejection of theories as false and that successful experiments could confirm theories

7

Clarifications about Duhem's view

- He never denied that *in fact* theories get refuted in science
- He described how scientists could protect their theory from refutation by modifying some of the assumptions
  - substitute $(T\&A_1\&...\&A_N)$ with $(T\&B\&A_2...\&A_N)$
- but he never said that any modification is reasonable
  - the new system must be consistent
  - B cannot be false
  - B cannot be ad hoc
  - …

8

Poincare's Conventionalism: the argument

- Many (mutually incompatible) theories can cope with the same data
- If that is true, then there is no way to find out which theory is correct (crucial experiments are impossible)
- If there is no way to find out which theory is correct, then there is no fact of the matter whether one theory is true or not

9

Poincare's Conventionalism

- Theories, theoretical terms and theoretical statements are neither true or false, they do not refer to anything
- They can only be classified as useful or not
  - Ex: "when a gamma ray hits a photographic places it leaves a mark" is neither T nor F
- They are instruments
  - Ex: a thermometer is neither true or false, but it is useful

10

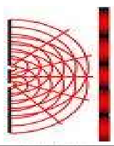
Duhem's rejection of Poincare's conventionalism

- Logical alone cannot force you to abandon a theory
- But "good sense" in science can
- Scientist A and scientist B can logically adopt different strategies wrt to T when experiments contradict it:
  - A: modifies the fundamentals of the theory
  - B: modifies some auxiliary hypotheses
- Good sense is telling when an experiment is crucial

11

Duhem's rejection of Poincare's conventionalism

- Example of good-sense at work: Jean Biot (1774-1862)
  - defender of the particle theory of light
  - more and more difficult to defend after the work of Thomas Young (1773-1829) and August Fresnel (1788-1827)



12

Duhem's rejection of Poincare's conventionalism

- Even after most scientists opted for the wave theory of light, Biot kept modifying the assumption in the particle theory
- But then followed the Foucault experiment (light travelled more slowly in water than in air) and he abandoned it
  - " … it may be that we find it childish and unreasonable … to maintain obstinately at any cost, at the price of continual repairs and many tangled-up stay, the worn eaten columns of a building tottering in every part, when by razing these columns it would be possible to construct a simple, elegant, and solid system. "

13

Quine's attack on the two dogmas of empiricism

- Holism, again
- The two dogmas of empiricism:
  - Analyticity: The analytic /synthetic distinction
  - Reductionism: Every meaningful synthetic statement is logically equivalent to some sentence containing only observational terms (joined together with logical connectives)
- They are dogmas because:
  - 1=the analytic/synthetic distinction is an unsupported article of faith
  - 2=reductionism is also unsupported because it is based on the analytic /synthetic distinction

14

Analytic-synthetic distinction: the basic idea (pre Quine)

- Synthetic statements: observational statements
  - Positivism (verifiability criterion of meaning = Reductionism):
    - A statement has **meaning** only if it logically implies a group of statements about experience/ they are confirmed by experience
    - Ex: this table is round
- Analytic statements: they have no observational content/ their meaning (if any) does not come from observation but from their linguistic component/ they are confirmed no matter what
    - Ex: all bachelors are unmarried males

15

Frege's definition of analyticity

- Frege's definition of analyticity (reduction to a tautology): a statement is analytic iff it is a tautology or can be reduced to it by means of definitions
  - No bachelor is unmarried→ No unmarried man is married
- Definitions are acceptable only when they preserve the existing meaning of the term in question
- So a satisfactory account of analyticity depends on a account of synonymy (sameness in meaning)
  - What is synonymy?

16

Q's rejection of the analytic-synthetic distinction

- Attempt 1: two terms, X and Y, are synonyms when they are interchangeable *salva veritate:*
  - Without changing the truth or falsity of the sentences in which they occur
    - X=bachelor
    - Y=unmarried man
      - All bachelors are unmarried men → All unmarried men are unmarried men

17

Q's rejection of the analytic-synthetic distinction

- But this does not guarantee the sameness in meaning:
    - X=creature with a heart
    - Y=creature with a kidney
      - All creatures with a heart are creatures with a heart → All creatures with a heart are creatures with a kidney
  - X and Y are interchangeable salva veritate because they the same extension (they refer to exactly the same objects) but they do not mean the same thing.
- For synonymy between X and Y we need more

18

Q's rejection of the analytic-synthetic distinction

- Attempt 2: X is a synonym of Y = necessarily, X iff Y
- But this just amounts to say that "X iff Y" is analytic, and this is circular
- A possible fix? Perhaps the failure can be traced to the vagueness of ordinary language
  - Artificial language in which the semantic rules are generating the analytic sentences
- But this is circular again:
  - what distinguishes these semantical rules (to generate analytic statements) from other semantical rules (such as those specifying all the truths of the language)?
  - These are the ones that picks out all and only the analytic sentences

19

Q's rejection of the analytic-synthetic distinction

- Quine's conclusion: the idea that there is an analytic-synthetic distinction is unsupported → meanings are not independent of other statements that we accept → <u>we cannot decide whether a given statement is analytic or synthetic **without** considering our entire web of beliefs</u> → Q's holism

20

Q's rejection of the analytic-synthetic distinction

- Q's holism: the connection with falsificationism/crucial experiments/…
  - Duhem: T&A→O, not just T
  - Quine: <u>make sure to include in A also what you thought were analytic truths</u>
    - QUD (Quinean UnderDetermination): any theory can be reconciled with any recalcitrant evidence by making suitable adjustments in our other assumptions about nature
    - If an experiment fails, even ***logic*** can be to blame

21

Duhem and Quine: Differences

- Context
  - Quine: in the context of analytic/synthetic distinction
  - Duhem: nothing like that
- Type of holism
  - Quine's <u>Semantic holism</u>: any expression in a language cannot be understood in isolation
  - Duhem's <u>Confirmation holism</u>: a theory cannot be tested alone by experience

22

Duhem and Quine: Differences

- Scope
  - Duhem: holism in physics; Quine: global holism – "the totality of our so-called knowledge or beliefs"
- Ways of saving a theory
  - Duhem: unreasonable and contrary to good sense to stick with a theory beyond a certain point
  - Quine: <u>pragmatic factors</u> are important (but no detail provided)
    - The only ground for choosing which explanation to believe is " the degree to which they expedite our dealings with sense experiences"

23

Criticism of Duhem

- Quine seems right (and Duhem wrong) in thinking that the whole science, and not just physics, should be subjected to the holistic thesis
- Other sciences:
  - Ambiguity of falsificationism is avoided in them in using instruments because the chemist, say, accepts many auxiliary hypotheses as established truths on the presumed infallibility of physics
    - Just a difference in the psychology of testing, not in the logic

24

Criticism of Quine

- The claim that "the whole science" is the unit of empirical significance is implausible
  - When a physical theory is combined with other theories and assumptions to generate a prediction, theories from other sciences play no role in the derivation
  - Quine, later on, tones down his thesis in this respect
    - "little is gained by saying that the unit is in principle the whole science"

25

A possible formulation of the Quine-Duhem thesis

- The holist thesis applies at any (high) level (contrary to Duhem, in light of Quine)
- The group of hypotheses under test in any given situation is in practice limited and does not extend to the whole human knowledge (contrary to Quine, in light of Duhem)
- Q's claim that "any statement can be held to be true come what may…" is true from a logical point of view but scientific good sense concludes in many situation that it would be perfectly unreasonable to hold to particular statements (addition to Quine, in light of Duhem)

- Donald Gillies

26

Underdetermination, Induction, Confirmation

27

Two types of underdetermination

- Underdetermination of theory by data: there is an infinite number of theories compatible with the same empirical data
- Deductive underdetermination
  - it limits itself to what can be established about the status of theories, given some evidence, through deductive logic
    - Hume
- Ampliative underdetermination
  - it permits the use of non-deductive inferences as well
    - Quine, Goodman, Kuhn, Hesse, Bloor

28

Induction and its role in science

The problem of induction

- J: Problem of justification: explaining why the general principles we follow when we make inductive inferences are reliable
  - discussed by Hume
- D: Problem of description: identifying the general principles we follow when we make inductive inferences
- Lipton: both problems arise because in every inductive arguments the conclusion is underdetermined by the premises
  - Ampliative nature of induction

29

Problem of justification

- Inductive arguments are deductively underdetermined
- Hume's argument: the inductive inference from (DATA) to (THEORY) is not deductively valid
- Ex: universal generalization (UG)
  - All observed As have been B
  - Therefore, UG: All As are B

30

The Problem of Induction

- This argument can become deductive if we add UN: nature is uniform
  - All observed As have been B
  - UN: nature is uniform
  - Therefore, UG: All As are B
- However, UN is underdetermined by the data

31

The Problem of Induction

- Hume's argument that one cannot justify UG:
- (1) if UG can be shown to be justified, then there is an argument that shows it;
- (2) Arguments are either deductive or inductive;
- (3) No deductively valid argument can justify UG (because of underdetermination);
- (4) No inductive argument can justify UG (because of circularity);
- Thus, UG cannot be justified.

32

The Problem of Induction

- Counterinductivist: Someone who will accepts DN (disuniformity of nature/counterincution)='it's time for a change instead of UN (uniformity of nature)
  - counterinductive (rather than inductive) inference from the same DATA → very different conclusion
  - EX:
  - Inductivist:
    - Observed metals expand when heated → All metals expand when heated
  - Counterinductivist:
    - Observed metals expand when heated → non-observed metals do not expand when heated

33

The Problem of Induction

- Bottom line:
- no one is more justified than the other to believe the conclusion they arrived at using their non-deductive inference because:
  - There is no non-circular justification for UN, and
  - There is no non-circular justification for DN

34

Proposed solutions: Hume

- Our habits of inductive inference and our acceptance of UN are non-rational:
  - No positive compelling reason can be provided for doing things our way rather than, say, the counterinductivist's way
- But our acceptance of UN is not irrational:
  - it is not contrary to reason. The principle is not logically inconsistent or incoherent, and no positive argument can be raised against it.
- Our acceptance of UN is not optional:
  - It is a matter of custom, habit, or (better) instinct

35

Proposed solutions: Hume

- Hume's "solution" is then this:
  - It is a natural, inevitable, not irrational belief: we are rationally justified in it
  - The counterindictivists are different from us, they do not reason like we do
- Consequences of Hume's solution
  - What this implies is that what we are rationally entitled to believe depends on what sort of beings we are, and not just on the available evidence and argument
  - A Counterinductivist could be equally rational

36

Proposed solutions: Strawson

- Peter Frederick Strawson: Induction is rational by definition
- Version 1: inductive inference conforms to our standards of rationality:
  - It is rational to make inferences from past observations to the future

37

Proposed solutions: Strawson

- Peter Frederick Strawson: Induction is rational by definition
- Version 1: inductive inference conforms to our standards of rationality:
  - It is rational to make inferences from past observations to the future
- Objection: calling something rational does not establish that the reasoning in question has the properties we want

38

Proposed solutions: Strawson

- Version 2:
- we are more certain of induction than of the arguments against it.
  - We can think of Hume's argument as a paradox and conclude that one of the premises has to be false

39

Proposed solutions: Strawson

- Version 2:
- we are more certain of induction than of the arguments against it.
  - We can think of Hume's argument as a paradox and conclude that one of the premises has to be false
  - Objection: But where is the flaw? What is the alternative to Hume's account?

40

Proposed solutions: Russell

- UN is a prori

41

Proposed solutions: Russell

- UN is a prori
- Reply:
- It just seems an unjustified stipulation, UN does not seem to be a priori!

42

Proposed solutions: laws

- Science gives us laws that explain *why* things have been predictable in the past, and the same laws guarantee that things will carry on the same way in the future

43

Proposed solutions: laws

- Science gives us laws that explain *why* things have been predictable in the past, and the same laws guarantee that things will carry on the same way in the future
- Reply:
  - But we have exactly the same problem concerning how we can know that the same laws will hold in the future as in the past.
  - Also, how do we know which generalizations are laws, and which are not? → see later (Grue)

44

Confirmation Theories

45

Evidence and hypotheses

- Problem of description
- We believe that certain inductive inferences are stronger than others, that certain evidence is more reliable than others
- Confirmation theory: general principle to determine the quality of the evidence

46

Hempel's criteria of Confirmation

- Theory evaluation used by scientist in deciding whether an hypothesis is acceptable or probably true
- Hempel ("Criteria of Confirmation and Acceptability"):
  - survey of criteria
  - no systematic, unified theory to justify them (unlike Bayesianism)
  - rejection of Falsificationism: hypothesis can be made more or less probable by evidence

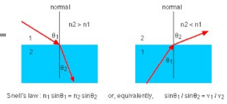
47

Criteria of Confirmation - evidence

- Quantity: the more evidence there is, the better for a theory
- Precision: important when we test identity claims or null results (ex gravitational and inertial mass in GR)
- Diversity: repetition on different kinds of tests, on a diverse sample, on a variety of conditions…
  - why a diverse sample?
    - Falsifiability: the more thorough is the test (i.e. the greater the power of the test to falsify a hypothesis), the more support a favorable outcome, the greater the confirming power of the evidence it generates (twist on Popper)

48

Criteria of Confirmation



- Ex: Snell's law in optics
  - $N_i \sin i = N_r \sin r$
- S_1: $N_{air} \sin 30 = N_{water} \sin r$
  - Preform T_1 (a test using air and water and i=30)→ test S_1
  - Add more angles → test S_2
  - Add more media → test S_3 ….
    - More we add variety, more of S we test
- S implies S_1, but S_1 could be true and S false, and T_1 wouldn't show it

49

Hypothetico-Deductive model

- HD model – did you forget? ;-)
- Articulate an H from which certain observational consequences can deductively be drawn
  - Observational consequence O=a statement whose truth or falsity can be established by making an observation
- Check O with an experiment
- If O is true, H is confirmed; if O is false, H is disconfirmed

50

Hypothetico-Deductive model

- Problems:
  - 1-statistical hypotheses cannot be confirmed
    - You cannot deduce E from H

51

Hypothetico-Deductive model

- 2-the problem of alternative hypotheses
  - Whenever an experimental result E confirms H, it also confirms infinitely many hypotheses that are incompatible with H
    - One prefer one over the others? HD does not say

52

Hypothetico-Deductive model

- 3-taking paradox (paradox of the irrelevant conjunction)
  - H→E, thus E confirms H
  - Consider H'=H&G
    - where G=some contingent statement independent of H and irrelevant to E, i.e. 'the moon is made of cheese'
  - H&G →E, thus E confirms H&G to the same amount it confirms H

53

Hypothetico-Deductive model

- 3- SCC and CCC cannot be both true SCC: H→G, E confirms H, then E confirms G
  - CCC: H→G, E confirms G, then E confirms H
  - H= all ravens are black; all swans are purple; E=a black raven; H'=H&G
    - CCC: H'→H, E conf H, thus E conf H'
    - SCC: E conf H', H'→G, thus E conf G
    - → a black raven confirms all swans are purple …???!!!

54

Hypothetico-Deductive model

- Other problems for the HD model:
  - Hempel's raven paradox: there is utterly irrelevant evidence such a white socks that must confirm a generalization
  - Goodman's paradox: a piece of evidence can be compatible with an infinite amount of hypotheses
    - → see later

55

The ravens paradox

- A black raven is an instance of H: "all ravens are black"
- A white raven is a counterinstance
- A white shoe is irrelevant
- However:
  - H:"all ravens are black" is logically equivalent to H': "all objects that are non-black are non-ravens"
  - Intuitively, a white shoe is an instance of H', so it confirms H'
  - But then it must also confirm H
  - But this seems counterintuitive, if not paradoxical

56

The ravens paradox – Hempel's solution

- The white shoe does confirm "All ravens are black"
- The result appears paradoxical because we possess prior information which we do not ignore:
  - We already know that a shoe is a non-raven
    - so why should we believe it would confirm 'all ravens are black'?
  - Instead, if we didn't know what the object was, to find it not to be black would add confirmation to 'all ravens are black':
    - it's an object that we now find is not a raven, and it also is not black

57

The ravens paradox – Hempel's solution

- H:"all sodium salts burn yellow"
- E: some compound that does not burn yellow
- E confirms H': "whatever does not burn yellow is not a sodium salt" and so H
  - This does not sound (too) odd

- H:"all sodium salts burn yellow"
- E: this piece of pure ice (i.e. not sodium salt) does not burn yellow
- E confirms H': "whatever does not burn yellow is not a sodium salt" and so H
  - This sounds odd!

58

The ravens paradox – Hempel's solution

- H:"all ravens are black"
- E: some object that is not black
- E confirms H': "whatever is not black is not a raven" and so H
  - This does not sound (too) odd

- H:"all ravens are black"
- E: this shoe (i.e. not a raven) is not black
- E confirms H': " whatever is not black is not a raven" and so H
  - This sounds odd!

59

The ravens paradox – Hempel's solution

- Moral: inductive inference, because it is ampliative, is sensitive to the context
- What looks to be a good induction in isolation, turns out to be not so good when the context, including background information, is taken into account
- The inference from "*a* is a white shoe" to "all ravens are black" is not so much unsound but uninteresting and uninformative

60

The ravens paradox- closer look

- P1: Instance confirmation (Nicod's Criterion, NC): a proposition of the form "all P are Q" is supported by the observation of a particular P that is a Q
- Thus, a non-black non-raven, e.g. a white shoe, confirms H'="all objects that are non black are non ravens"
- P2: Equivalence condition (EC): if E provides evidence for H, then E also provides evidence for any proposition logically equivalent to H
- H="all ravens are black" is logically equivalent to H'="all objects that are non black are non ravens"
- PC (paradoxical conclusion): Thus, a white shoe confirms H="all ravens are black"

61

The ravens paradox- closer look

- Two apparently reasonable premises:
  - Equivalence condition (EC): if E provides evidence for H, then E also provides evidence for any proposition logically equivalent to H
  - Instance confirmation (Nicod's Criterion, NC): a proposition of the form "all P are Q" is supported by the observation of a particular P that is a Q
- can be combined to reach the paradoxical conclusion:
- PC: the observation of a green apple or a white shoe provides evidence that all ravens are black

62

Solutions which accept PC- Hempel

- Inductive inference, because it is ampliative, is sensitive to the context
- What looks to be a good induction in isolation, turns out to be not so good when the context, including background information, is taken into account
- The inference from "*a* is a white shoe" to "all ravens are black" is not so much unsound but uninteresting and uninformative

63

Solutions which reject NC

- Quine on natural kinds:
- any set of objects forms a kind only if (and perhaps if) it is "projectible," i.e. judgments made about some members of that set can plausibly be extended by induction to other members
  - "raven" and "black" are natural kinds:
    - any black raven provides some evidence that all ravens are black
  - "non-black" and "non-raven" are not:
    - a nonblack non-raven is *not* evidence that all non-black things are non-ravens

64

The Grue paradox-the new riddle of induction

- There are other faulty inductions that do not seem to be accountable for by reference to background information
- Some generalizations seems to be conformed by their instances, some do not
- Ex: generalizations
  - H1:"Everybody in this room is safe from freezing"
    - Confirmed by (1): Joe is in this room and is safe from freezing
  - H2:"Everybody in this room is third son"
    - NOT Confirmed by (2): Joe is in this room and he is a third son
- Explanation:
  - H1 is a law-like generalization, while H2 is accidental
- The problem is raised by the Grue paradox: what makes a generalization law-like?

65

The Grue paradox-the new riddle of induction

- Old problem of induction:
  - The problem of justification: reliability of inductive inferences
- Goodman's solution:
  - We need not give any independent justification for the validity of inductive reasoning:
  - Inductive reasoning – like deductive reasoning - is a basic method of inquiry
    - Inductive validity: An argument from premises P, Q... to a conclusion C is inductively valid just in case anyone who accepts the premises while rejecting the conclusion is being unreasonable
    - To say that and argument is inductively valid is just to say that it is correct by our own standards

66

The Grue paradox-the new riddle of induction

- Compare:
  - we have standards for deciding when a sentence is grammatical; there is nothing more to being grammatical than obeying to our standards
  - Similarly: the implicit rules we use to classify arguments as inductively valid determine what really is valid

67

The Grue paradox-the new riddle of induction

- Linguistics produces a set of formal rules for determining whether a sentence is grammatical
- Deductive logic produces a set of explicit formal rules for deductive validity
- Inductive logic produces the set of rules concerning how well a given body of evidence supports any given hypothesis (logic of confirmation)
- We have *tacitly accepted* a formal logic of induction. Now the project is to make these tacitly understood principles explicit

68

The Grue paradox-the new riddle of induction

- Compare induction with deduction:
  - An inference is deductively valid iff it follows a deductively valid rule of inference
  - Deductive inferences are justified as being in conformity to the valid rules of deduction
  - It is circular, but the circle is virtuous, rather than vicious, since the rules and instances are being brought to harmony

69

The Grue paradox-the new riddle of induction

- We can do the same for inductive reasoning:
  - an inference is inductively valid iff it follows an inductively valid rule of inference
- How do we know which rules are inductively valid?
  - We check the general rules of inference by looking at the particular inferences that **we** are inclined to find acceptable and we check the particular inferences by looking at whether they follow the general rules of inference
  - → the research project of developing a logic of induction

70

The Grue paradox-the new riddle of induction

- Goodman's riddle is an argument against the possibility of a formal inductive logic - so the new riddle is the problem of description:
  - How can we distinguish between 'valid' and 'invalid' inductions?
- A rule of inductive inference looks roughly like this:
  - (INDUCTION): if before a certain time t I have observed that all F are G, then I have -to some degree- confirmed the hypothesis that all F are G
- Goodman's riddle aims to show that we cannot give any general answer to whether this is 'valid' or not

71

The Grue paradox-the new riddle of induction

- H:"all emeralds are green"
- Instance: green emerald → confirms H
- Now define "grue" :
  - X is **grue** iff X is green and observed before T, or x is blue and never observed before or at T
  - Grue x = (Gx&Ox) v (Bx&~Ox)
  - T=some fixed future time
- Each observed green emerald also supports H': "all emeralds are grue"

72

The Grue paradox-the new riddle of induction

- However, this means that the <u>same observations</u> of green emeralds support incompatible hypothesis:
- How should we *project* our inferences about the **next** emerald (call it Bob) that will be observed at or after T?
  - If H: Bob will be green
  - If H': Bob will be blue
- We have two inductive arguments, which lead to logically incompatible predictions about Bob
- If both are inductively valid, then we must conclude that a reasonable person would believe **both**
- but no reasonable person would do that

73

The Grue paradox-the new riddle of induction

- By our standards, the argument for the conclusion that Bob (the first emerald observed after t) is green is valid (and thus H was confirmed), the other is not
- Since both arguments have the same form, we are forced to conclude that inductive validity is not a formal property of arguments
- This is a striking disanalogy with deduction

74

The Grue paradox-the new riddle of induction

- This is this problem of distinguishing between confirmable and non-confirmable hypothesis
- This is the problem of "projectability":
  - a predicate is projectible iff figures in an inductively valid instance of (INDUCTION).
- "Green" is projectible, "grue" is not
- The problem is then to state a rule for distinguishing the good projectible predicates from the bad un-projectible ones

75

Possible solutions -clarifications

- Clarification: This does not entail that emerald change in color
  - 'or' in the definition is inclusive (A or B or both)
  - So, if emeralds change color at T they'd be grue
  - But to be grue it is merely sufficient that either
    - x is observed only <u>before</u> T and found green (and will always remain green)
    - x is observed only <u>after</u> or at T and found blue (and has always been blue)

76

Possible solutions -clarifications

- Lifetime (a-temporal) vs time-related predicates:
- A-temporal predicates: they apply throughout the entire existence of the object
  - Ex: human, soluble, …
- Time-related predicates: they apply only to a time, or a period of time
  - Ex: color - This car is yellow new but it was red before the spray job
- Goodman thinks of grue as a lifetime predicate
  - If it's true at t, it's always true

77

The Grue paradox-the new riddle of induction

- More clarifications:
- The paradox cannot be resolved looking at background knowledge
- Bleen and grue are not colors (two objects can have the same shade of blue and yet one of them be bleen and the other grue, depending on when each is first observed) , they are schmolors.

78

Objections and replies

- Grue is not projectible because it refers to a particular time in its definition (it's a positional predicate)

79

Objections and replies

- Grue is not projectible because it refers to a particular time in its definition (it's a positional predicate)
- Goodman's Reply: this is relative to language.
  - Green: iff examined before t and grue; or not examined before t and bleen
- The situation is completely symmetrical

80

Objections and replies

- Nature is not uniform if emeralds are grue: not all emeralds will be of the same color

81

Objections and replies

- Nature is not uniform if emeralds are grue: not all emeralds will be of the same color
- Goodman's Reply: they will have the same schmolour (=the grue/bleen property)

82

Goodman's own solution

- Old riddle: dissolved because, just as deduction, induction is justified by conformity to accepted practices
- New riddle: confirmed generalizations contain predicates that are well entrenched in our practice, we have successfully used it many times before in our inductive inferences
  - 'grue' is not entrenched, while 'green' is

83

Feature of this solution

- The distinction between projectible and non-projecible depends on contingent accidental facts about our history and our practice
- There is nothing to rule out an alternative development of our inductive practices
- What makes and inference good or bad depends on us
- This seems disturbing:
  - had we born in a different culture with other entrenched predicates we would consider other predicates to be projectible

84

Objection?

- Science uses new (thus unentrenched) predicates all the time

85

Objection?

- Science uses new (thus unentrenched) predicates all the time
- Goodman: what is entrenched is the extension, not the predicate
- Novel terms inherit their entrenchment from the old ones
  - Q inherit entrenchment from P if P and Q have the same extension

86

Objection?

- Science uses new (thus unentrenched) predicates all the time
- Goodman: what is entrenched is the extension, not the predicate
- Novel terms inherit their entrenchment from the old ones
  - Q inherit entrenchment from P if P and Q have the same extension
  - However: how do we know that grue and green have NOT the same extension?

87

Grue paradox, and underdetermination

- The paradox has not so much about time indexing but much more to do with constraints that we wish to put on inductive inferences:
- When we infer from a sample to a population (such as from observed As being B to All As being B), we presuppose that As in our sample would have been B if they weren't in it
- This is not the case for grue:
  - We don't believe the counterfactual conditional to apply because blue things (thus grue if observed >=T) would NOT have been observed to be grue at the earlier time at which the sample was drawn

88

Grue paradox, and underdetermination

- For any prediction we can always find a regularity that licenses that prediction
- Ex 1: the three boxes
  - in the first box there is a green insect, in the second a yellow ball, in the third a purple feather
  - The fourth contains a mask
  - You are asked to predict its color
- You must look for regularities, and **with little ingenuity** we can always find one
- Suppose you want to predict the mask is red. Here is how to induce such regularity: …

89

Grue paradox, and underdetermination

- Define "snarf": something presented you in a box and that is either an insect, or a ball, or a feather, or a mask
  - Now you have observed three snarfs
- Define "murkle": a thing X is murke just when it is an insect and it is green, or it is a ball and it is yellow, or it is a feather and it is purple, or it is another object and is red
- Now there regularity is: "all observed snarfs are murkle"
- If we project that into the future we obtain the required prediction

90

Grue paradox, and underdetermination

- Another example
- graphs and curves:
  - for the same points there are infinitely may curves that could fit the points, each that would lead to different predictions

91

Grue paradox, and underdetermination

- Another example again:
- different series of numbers could be completed in different ways, all compatible with the observable numbers in the series:
  - i: 1,2,3,4,5,...
  - ii: 2,4,6,8,10,...
  - iii: 1,3,5,7,9,...
  - generating function of i: k
  - generating function of ii: 2k
  - generating function of iii: 2k-1
  - But also:
  - (k-1)(k-2)(k-3)(k-4)(k-5)+k is a generating function for i, and this leads the result 126 for k=6.

92

Grue paradox, and underdetermination

- The problem of formulating precise rules for determining projectibility is the new riddle
- It is just an instance of the ubiquitous **underdetermination of hypothesis** by data
- We need to restrict the set of hypotheses that we are to consider, and focus on those with projectible concepts is one way to do this

93

Other solutions to the grue paradox – Natural kinds

- Quine, Kripke, Putnam: Only generalizations which include natural kinds are projectible
  - Each member of a natural kind share a common nature in virtue of which it belongs to that kind
    - Natural kinds ←→ essential properties ←→ microstructural properties.
    - Ex: being H2O is an essential property of water, having atomic number 79 is an essential property of gold...
    - It is impossible for gold not to have 79 as its atomic number. So the law that gold has atomic number = 79 is necessary (even if its discovery is empirical)

94

Other solutions to the grue paradox

- Bayesian Conformation Theory: "grue" is less confirmed than "green", almost zero:
  - The degree of confirmation depends in part of the prior probabilities of the hypothesis relative to our background knowledge.
  - "all emeralds are grue" has a much lower prior probability than "all emeralds are green", hence it receives little confirmation

95

Bayesian Confirmation Theory

- The Bayesian approach to confirmation:
  - It provides a unified explanation of a wide range of accepted principles of scientific methodology including:
    - Statistical hypotheses can be confirmed
    - Not every theory is confirmed equally by the evidence it entails
    - Novel predictions have a special confirmatory value
    - Simple hypotheses are to be preferred
    - Ad hoc hypotheses should be avoided
    - A diverse set of evidence lend stronger support to theories than narrow ones
    - Sometimes it is good to put the blame on some auxiliary hypotheses when the theory is disconfirmed by evidence

96

Bayesian Confirmation Theory

- The Bayesian approach to confirmation:
  - It provides a unified explanation of a wide range of accepted principles of scientific methodology including:
    - Statistical hypotheses can be confirmed
    - Not every theory is confirmed equally by the evidence it entails
    - Novel predictions have a special confirmatory value
    - Simple hypotheses are to be preferred
    - Ad hoc hypotheses should be avoided
    - A diverse set of evidence lend stronger support to theories than narrow ones
    - Sometimes it is good to put the blame on some auxiliary hypotheses when the theory is disconfirmed by evidence
  - It solves the paradoxes of confirmation (raven, Goodman)

97

Bayesian Confirmation Theory

- The Relevance Criterion of Confirmation
- unlike HD or instance confirmation, here the confirmation is quantitative
  - E confirms H iff P(H/E)>P(H)
  - E disconfirms H iff P(H/E)<P(H)
- Notation:
  - P(H): prior probability of H
  - P(H/E): posterior probability of H → probability of H given E

98

Bayesian Confirmation Theory

- Bayes' Theorem:
  - it is a deductive consequence of the three basic axioms of probability theory
- Axioms of probability theory:
  - 1 - every probability is a real number between 0 and 1
  - 2 - if A is a necessary truth, then P(A)=1
  - 3 - if A and B are mutually exclusive (that is, if it is impossible for both A and B to be true) the P(AvB)=P(A)=P(B)

99

Bayesian Confirmation Theory

- Several important theorems:
  - Negation rule: P(~A)=P(A)
  - Implication rule: if A logically entails B, then P(B)>=P(A)
  - Equivalence rule: if A and B are logically equivalent, then P(A)=P(B)
  - General addition rule: P(AvB)=P(A)+P(B)-P(A&B)

100

Bayesian Confirmation Theory

- Conditional probability rule: P(A/B)=P(A&B)/P(B) where P(B)>0
- General multiplication rule: P(A&B)=P(A/B)P(B)
- Special multiplication rule: when A and B are independent, P(A&B)=P(A)P(B)
- Total probability rule: P(A)=P(A/B)P(B)+P(A/~B)P(~B)

101

Bayesian Confirmation Theory

- Bayes' theorem: $P({}^{B}/_{A}) = P({}^{A}/_{B})\frac{P(B)}{P(A)}$, where P(A)>0
- Bayes' theorem (BT) and scientific reasoning:
  - $\boldsymbol{P({}^{H}/_{E}) = P({}^{E}/_{H})\frac{P(H)}{P(E)}}$
    - P(H/E): posterior probability of H given evidence E
    - P(H): prior probability of H
    - P(E): probability of evidence E (expectedness)
    - P(E/H): probability of E given H (likelihood of H on E)
    - All probabilities are **subjective degrees of beliefs**
  - Bayes's theorem tells us how to revise our degrees of belief in H given that we have acquired the evidence E

102

Bayesian Confirmation Theory

- More generally:
  - $P(H/E) = \frac{P(E/H)P(H)}{[P(E/H)P(H)+P(E/{\sim}H)P({\sim}H)]}$
  - or, even more generally:

$$P\left(H_k/E\right) = \frac{P\left(E/H_k\right)P(H_k)}{\sum_{i=1}^{N} P\left(E/H_i\right)P(H_i)}$$

103

Bayesian Confirmation Theory

- How to compute the probabilities?
  - P(E/H) (likelihood): not problematic
    - if H is deterministic, it deductively entails E, thus P(E/H)=1
    - if H is a statistical hypothesis, then H will specify the probability of E

104

Bayesian Confirmation Theory

- How to compute the probabilities?
  - P(E) (expectedness): more problematic
  - Purely subjective measure of how expected (i.e. not surprising) the evidence is
    - → so there is not only one P(E)
  - (Also, we need to know all the P(E/Hi) and all the P(Hi), and we usually do not know what all the alternative Hi are)

105

Bayesian Confirmation Theory

- How to compute the probabilities?
  - P(H) (prior): also problematic
  - it is the subjective degree of belief that a given person has in that moment in H
    - → so there is not only one P(H)

106

Bayesian Confirmation Theory

- Appeal:
- 1) BCT can handle statistical hypotheses:
- P(E/H), the likelihood of the evidence, can be less than 1

107

Bayesian Confirmation Theory

- 2) BCT provides degrees of confirmation (thus solves the tacking problem):
- when comparing two theories T1 and T2:
  - T1=H; T2=H&I, I=irrelevant conjunction
  - priors ? P(H)>>P(H&I)
  - Bayes theorem:
    - P(H/E)=P(E/H)P(H)/P(E) =(det. theories)=P(H)/P(E)
    - P(H&I/E)=P(E/H&I)P(H&I)/P(E) =(d.t.)=P(H&I)/P(E)
  - → posteriors: P(H/E)>>P(H&I/E) so, E confirms H better than it confirms H&I

108

Bayesian Confirmation Theory

- 3) Surprising predictions confirm more:
  - P(H/E)=P(E/H)P(H)/P(E) =(det. theories)=P(H)/P(E)
  - S=surprising; O=old
  - P(S)<<P(O) (by def.), P(E) is at the denominator
  - So, P(H/S)>>P(H/O)

109

Bayesian Confirmation Theory

- 4) Preference to simple hypotheses
- simpler theories have greater priors:
- $P(H_{simple})>>P(H_{compex})$
  - ex: T1: NM with "force", and T2: NM with "gorce & morce"

110

Bayesian Confirmation Theory

- Problem: how do we judge that T2 is worse off??
- (as: the simplest equations that entails the data is the best choice)
  - either there is a measure of simplicity for equations that can be proven to have higher priors
  - or argue that the simpler has higher likelihoods

111

Bayesian Confirmation Theory

- 5) evidence from a diverse sample confirms better:
  - Diverse evidence is more likely to reveal which competitors are false → more rival hypotheses are eliminated → the probability of survivors increases

112

Bayesian Confirmation Theory

- 6) BCT explains why sometimes it's rational to blame the auxiliary hypothesis (A):
- P(H)>P(A)
- Thus:
- P(H/E)/P(A/E)=P(H)/P(A) > 1 → P(H/E)>P(A/E)

113

Bayesian Confirmation Theory

- 7) BCT explains why we dislike ad hoc hypotheses:
- $P(H_{ad\ hoc})<P(H_{not\ ad\ hoc})$
- Thus:
- $P(H_{ad\ hoc}/E)<P(H_{not\ ad\ hoc}/E)$

114

Bayesian Confirmation Theory

- 8) Solution of the ravens paradox:
  - BCT defines comparative (A confirms H better than B) and quantitative (a confirms H to degree p) confirmation
  - It is then capable of differentiating support for the two hypothesis in question
  - The amount of confirmation provided to "all ravens are black" by a white shoe is positive but very small, due to the large discrepancy between the number of ravens and the number of non-black objects
  - The conclusion appears paradoxical because we intuitively estimate the amount of evidence provided by the observation of a white shoe to be zero

115

Bayesian Confirmation Theory

- 9) Solution of the grue paradox
- "grue" is less confirmed than "green", almost zero:
  - The degree of confirmation depends in part of the prior probabilities of the hypothesis relative to our background knowledge.
  - "all emeralds are grue" has a much lower prior probability than "all emeralds are green", hence it receives little confirmation

116

Bayesian Confirmation Theory

- Salmon:
- Bayesian Algorithm for Theory Choice
- $\frac{P(T_1/E)}{P(T_2/E)} = \frac{P(T_1)}{P(T_2)}\frac{P(E/T_1)}{P(E/T_2)}$
- Choose the theory with higher posterior probabilities

117

Bayesian Confirmation Theory

- Bayesian Algorithm for Theory Choice
- Problem:
- For deterministic theories we have P(E/T)=1
- Thus, Ratio prior probabilities=Ratio posterior probabilities
  - Thus no amount of evidence can change our degree of belief in a theory
  - Priors are completely **dominant**
  - Salmon: Kuhn's criteria are relevant in selecting them

118

Bayesian Confirmation Theory

- Problems for BTC:
- If P(T)=0, it will remain zero no matter what

119

Bayesian Confirmation Theory

- Problems for BTC:
- If P(T)=0, it will remain zero no matter what
- Response:
  - Require strict coherence
  - No contingent proposition is assigned either 0 or 1

120

Bayesian Confirmation Theory

- Problems for BTC:
- How can subjective degrees of belief (in P(H)) be reconciled with scientific objectivity?

121

Bayesian Confirmation Theory

- Problems for BTC:
- How can subjective degrees of belief (in P(H)) be reconciled with scientific objectivity?
- Response: the theorem of washing out the priors
  - as evidence accumulates, the values of P(H/E) calculated by different people with different priors will converge

122

Bayesian Confirmation Theory

- Problems for BTC:
- How can we motivate our use of Bayes' theorem to update our degrees of beliefs?

123

Bayesian Confirmation Theory

- Problems for BTC:
- How can we motivate our use of Bayes' theorem to update our degrees of beliefs?
- Response: construct a further Dutch book argument, that is diachronic
  - make a series of bets, if one does not follow BT one can always construct a Dutch book against her (theorem)

124

Bayesian Confirmation Theory

- Problems for BTC:
- Problem of old evidence: old evidence does not confirm
  - $P(E_{old})=1$, thus $p(H/E_{old})=P(H)$
  - Ex: Newton with Kepler and Galileo; Einstein and Mercury
- Responses:
- 1) "Fine, old evidence does not confirm, and everyone who thought otherwise was wrong"

125

Bayesian Confirmation Theory

- 2) Deny that $P(E_{old})=1$
- $P(E_{old})$ is actually $P(E_{old}/B)$
- B=background knowledge you would have if E were not yet known (counterfactual background knowledge)
- How to compute this?
  - Present proposal: try to imagine what this B could be
  - Historical proposal: consider what B was in the past before E was known
    - Both have troubles, since it is difficult to pin down what they mean…

126

Bayesian Confirmation Theory

- 3) New criterion of confirmation:
  - E confirms H iff P(H/H→E)>P(H)
  - Confirmation is done not by E, but by the discovery that H entails E

127

Bayesian Confirmation Theory

- Problem: sometimes we think E_{old} conf H even if we know well in advance that H→E; sometimes H is designed to account for E_{old} (so, that H→E_{old} is hardly a discovery)
  - Ex: Einstein's theory was designed to account for the motion of Mercury's perihelion and we thought it was confirmed

128