From Textbook Bayesianism to Naturalized Bayesianism

Stephan Hartmann

London School of Economics

Outline

I. Bayesianism vs. Reliabilism
II. Motivation
III. Modeling in Science
IV. Textbook Bayesianism
V. Bayesian Networks
VI. Example 1: Variety-of-Evidence Thesis
VII. Example 2: Testimony
VIII. Example 3: Scientific Theory Change
IX. Naturalized Bayesianism
X. Conclusions

I. Bayesianism vs. Reliabilism

Scope

- Glymour 1: "One of Stephan's colleagues complains that *Tetrad* does not apply everywhere, and that's not a good criticism."
- Glymour 2: "Bayesianism does not work well for all search problems, and that's not good for Bayesianism."
- Glymour 1 and Glymour 2 seem *inconsistent*.
- My position: Let's see how far we get with one program.

Confirmation

- Glymour 1: “Scientists agree that the anomalous perihelion of Mercury *confirms* GTR.”
- Glymour 2: “There is no need to explicate the concept of confirmation because it is not a generic concept. All there is are (more or less good) tests.”
- Question: Is confirmation a generic concept, or not?
- The Bayesian thinks yes, and aims at explicating it. The Bayesian also thinks that other concepts such as explanation, reduction etc. need to be explicated.

How typical is data mining?

1. Reliabilist methods work well for data mining problems.

   But not all of science is data mining. Einstein, for example, didn’t use any data to come up with GTR.

2. Alternative hypotheses have to be known, and then we can apply reliabilist methodologies and (hopefully) get convergence.

   But could one come up with GTR in this way? Like so often in “discovery”, the important work has to be done before. (Note the analogy to Glymour’s argument against Bayesianism.)

Truth?

- In the reliabilist framework, the truth can only be obtained if we consider the true hypothesis as one of the candidates. If we do not, we won't arrive at a truth hypothesis.
- This can also be done, in principle, in a Bayesian framework. But Bayesians are *modest* and reluctant to use the big word "truth."
- The Bayesian is content with convergence and agreement amongst actors. This gives her the best available hypothesis. Can we ask for more?

II. Motivation

Some questions I am interested in

- How are scientific theories and models confirmed?
- How is evidence for or against a theory evaluated?
- How do different theories hang together?
- Can aspects of scientific theory change be explained philosophically?

I am not the first to address these questions…

Two traditions…

1. Normativism
   - Examples: Falsificationism, Bayesianism
   - Typically motivated on *a priori* grounds
2. Descriptivism
   - Examples: Kuhn, naturalized philosophies of science
   - Examine specific case studies that contribute to a better understanding of science

…and their problems

1. Normativism
   - challenged by insights from the history of science (e.g. Popper and the stability of normal science)
   - often “too far away” from real science
2. Descriptivism
   - not clear how generalizable insights or normative standards can be provided

Desiderata for a methodology of science

1. It should be *normative* and provide a defensible *general* account of scientific rationality.
2. It should provide a framework to *illuminate* "intuitively correct judgments in the history of science and explains the incorrectness of those judgments that seem clearly intuitively incorrect (and shed light on 'grey cases')" (John Worrall)

How can this goal be achieved?

- Mimic successful scientific methodology and…
- construct (philosophical) models in the framework of a (philosophical) theory.

I'll explain what I mean by this.

III. Modeling in Science

The ubiquity of models in science

- Highschool: Bohr model of the atom, model of the pendulum,…
- University: Standard Big Bang Model, Standard Model of particle physics,…
- Even later: Physicists like Lisa Randall (Harvard) use models to learn about the most fundamental question of the universe.
- We observe a shift from theories to models in science, and this shift is reflected in the work of philosophers of science in the last 25 years.

Theories and models

- Scientists often use the words *theory* and *model* interchangeably.
- Example: the Standard Model of particle physics.
- So how can one distinguish between theories and models?

Features of theories

- Examples: Newtonian Mechanics, Quantum Mechanics
- General and unversal in scope
- Abstract
- Often difficult to solve (example: QCD)
- No idealizations involved (ideally...)

Features of models

- Examples: the model of a pendulum, the Bohr model of the atom, gas models, the MIT Bag model,...
- Specific and limited in scope
- Concrete
- Intuitive and visualizable
- Can be solved
- Involve idealizing assumptions

Two kinds of models

- Models of a theory

  - example: the treatment of the pendulum is a model of Newtonian Mechanics

  - the theory acts as a modeling framework (and not much follows from the theory w/o specifying a model)

- Phenomenological models

  - example: the MIT Bag model

  - there is no theory into which a model can be embedded.

Modeling in philosophy

- So maybe we should construct models in philosophy as well!
- Some authors do this already, see, e.g., the work of Brian Skyrms on the social contract and Clark Glymour et al.'s work on causal discovery.
- I will show that models are also valuable tools for *understanding the methodology of science*.

IV. “Textbook Bayesianism”

What is Bayesianism?

- Quantitative confirmation theory
- When (and how much) does a piece of evidence *E* confirm a hypothesis *H*?
- Typically formulated in terms of probabilities = subjective degrees of belief
- Normative theory (see Dutch books)
- Textbooks: Howson & Urbach: *Scientific Reasoning*, and Earman: *Bayes or Bust*.

The mathematical machinery

- Confirmation = positive relevance between *H* and *E*
- Start with a prior probability $P(H)$
- Updating-rule: $P_{new}(H) := P(H|E)$
- By mathematics ("Bayes' Theorem") we get:

$$P_{new}(H) = P(E|H)\, P(H)/P(E)$$

- *H* confirms *E* iff $P_{new}(H) > P(H)$
- *H* disconfirms *E* iff $P_{new}(H) < P(H)$

Examples of *Textbook Bayesianism*

- Let's assume that H entails E: $P(H|E) = 1$.
- Then $P_{new}(H) = P(H)/P(E)$
- (i) It is easy to see that Bayesianism can account for the insight that surprising evidence (i.e. if $P(E)$ is small) confirms better.
- (ii) Let's assume we have different pieces of evidence: $P(E) = \mathrm{P}(E_1, E_2, \ldots, E_n) = P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_2, \mathrm{E}_1) \cdot \ldots \cdot P(E_n|E_{n-1}, \ldots, \mathrm{E}_1)$. Hence, less correlated pieces of evidence confirm better: the variety-of-evidence thesis.

We conclude

- Textbook Bayesianism *explains* very general features of science.
- However, it turns out to be too general as it does not take into account *de facto constraints* of the scientific practice (such as dependencies between measurement instruments). Taking them into account might lead to different conclusions.
- Moreover, Textbook Bayesianism does not have an account of what a scientific theory is.

Jon Dorling's version

- Jon Dorling aims at reconstructing specific episodes in the history of science and fitting them into the Bayesian apparatus.
- To do so, one has to assign specific probability values to the hypotheses and likelihoods in question.
- This variant of Textbook Bayesianism is too specific.

Upshot

- We need a version of Bayesianism that is not too general (to connect to the practice of science), and not too specific (to gain some philosophical insight).
- It would also be nice to have an account that has a somewhat wider scope, i.e. an account that reaches beyond confirmation theory.

  Modeling will do the job!

Modeling and Bayesianism

- Models help bridging the gap between a general theory (here: Bayesianism) and the scientific practice.
- Bayesianism is taken to be a modeling framework (just like Newtonian Mechanics is).
- The models we will construct are *models of a theory*.
- These models will not be too specific, i.e. we will not assign specific numbers to the probability variables, but explain general features of the methodology of science.

V. Bayesian Networks

An example from medicine

- T: Patient has tuberculosis
- X: Positive X-ray
- Given information:

  $t := P(\mathrm{T}) = .01$

  $p := P(\mathrm{X}|\mathrm{T}) = .95 = 1- P(\neg\mathrm{X}|\mathrm{T}) = 1-$ rate of *false negatives*

  $q := P(\mathrm{X}|\neg\mathrm{T}) = .02 =$ rate of *false positives*

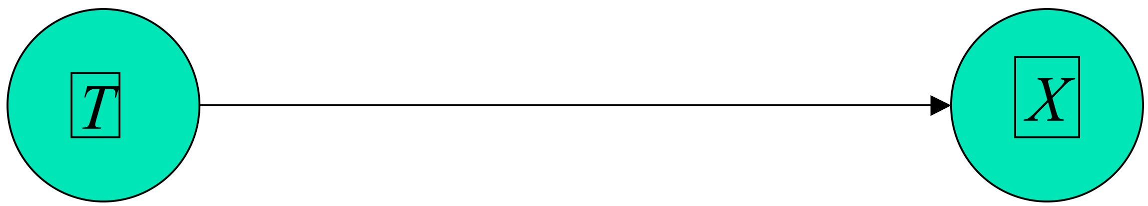
- What is $P(\mathrm{T}|\mathrm{X})$? $\Rightarrow$ Apply Bayes' Theorem

$$P(\mathrm{T}|\mathrm{X}) = P(\mathrm{X}|\mathrm{T})\, P(\mathrm{T}) / [P(\mathrm{X}|\mathrm{T})\, P(\mathrm{T}) + P(\mathrm{X}|\neg\mathrm{T})\, P(\neg\mathrm{T})] =$$

$$= p\, t / [p\, t + q\,(1-t)] = t/[t + (1-t)\, x] \quad \text{with } x := q/p$$

$$= \underline{.32}$$

A Bayesian Network representation



$P(T) = .1$

$P(X|T) = .95$

$P(X|\neg T) = .2$

Parlance:

"*T* causes *X*"

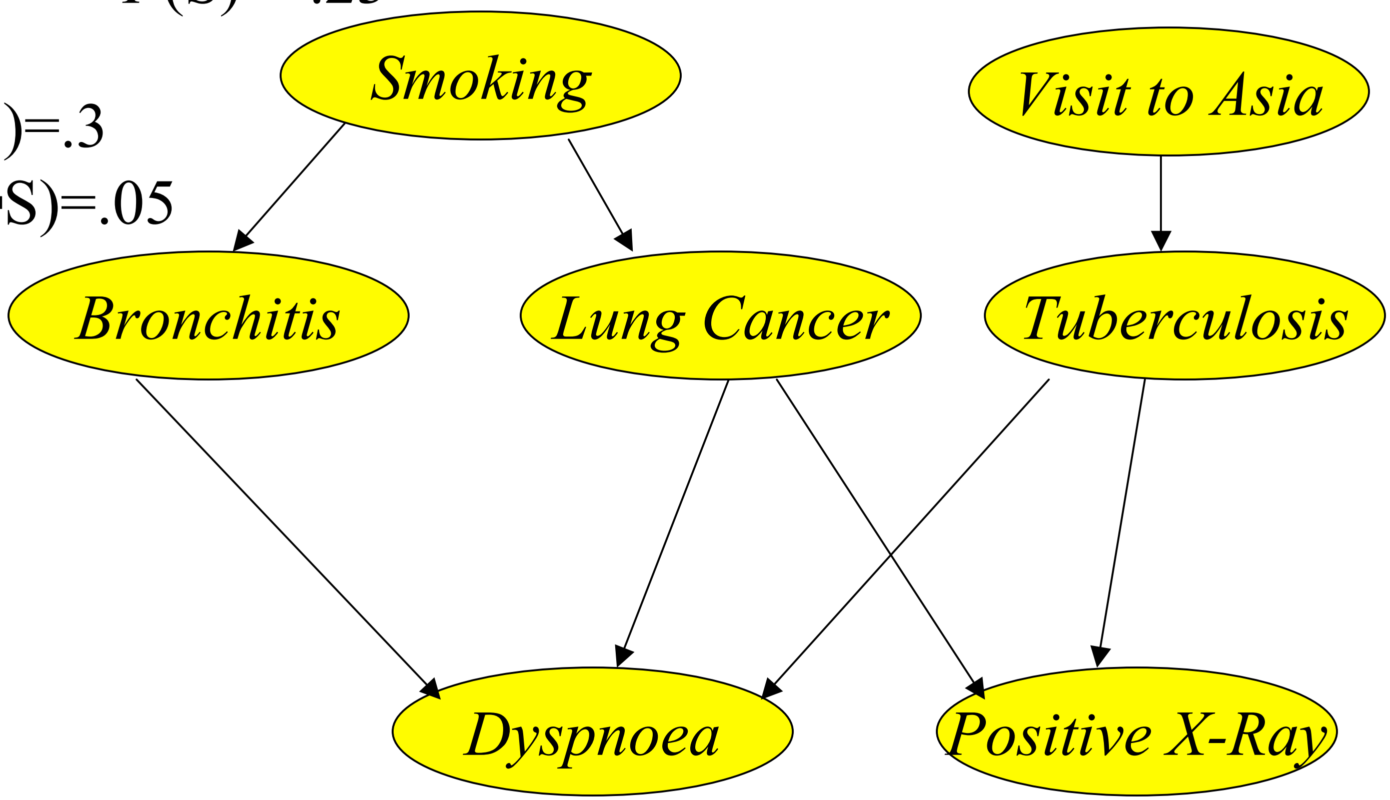
"*T* directly influences *X*"

A more complicated (=realistic) scenario

$P(\mathrm{S}) = .23$

$P(\mathrm{B}|\mathrm{S}) = .3$

$P(\mathrm{B}|\neg\mathrm{S}) = .05$



What is a Bayesian Network?

- A BN is a directed acyclic graph (dag, for short) with a probability distribution defined over it.
- BNs represent conditional independencies that hold between (sets of) variables, which helps reducing the number of probabilities that have to be fixed to specify the joint probability distribution.
- The *Parental Markov Condition* is implemented in the network.

What's so good about BNs?

- Bayesian Networks are intuitive
- Specifying the joint probability distribution $P(A_1, A_2, \ldots, A_n)$ over n binary variables requires the specification of 2^n-1 probabilities.
- However, in a BN one only has to specify the probability of each node given its *parents*: $P(A_i|par(A_i))$.

Algorithms

- One can show that the joint probability $P(A_1, A_2, \ldots, A_n) =$ $P(A_1|\mathrm{par}(A_1)) \cdot P(A_1|\mathrm{par}(A_1)) \ldots \cdot P(A_n|\mathrm{par}(A_n))$

- There are even more efficient algorithms to compute whatever probability one is interested in.

How to construct a Bayesian Network model

- Fix a set of variables.
- Specify the probabilistic independencies that hold among them.
- Construct a Bayesian Network and specify all $P(A_i|\mathrm{par}(A_i))$.
- Calculate the requested (conditional or unconditional) probabilities.

The proof of the pudding is in the eating…

VI. Example 1: The Variety-of-Evidence Thesis

The question

- Does the variety-of-evidence thesis *always* hold?
- What happens, for example, if the measurement instruments (that provide us with the evidence for the hypothesis we test) are only partially reliable?

One reading of the thesis

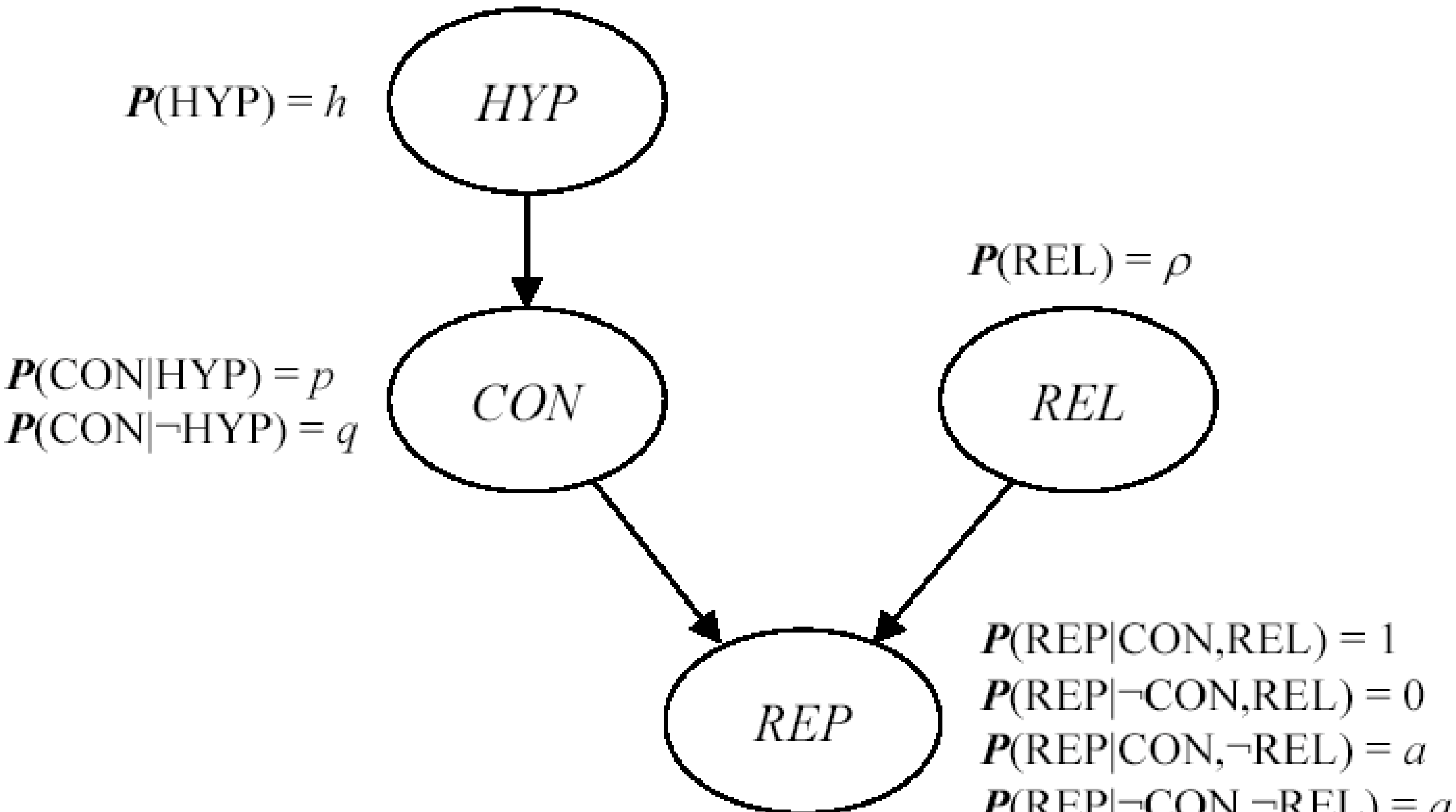
- A hypothesis *H* was positively tested with an instrument I_1.
- Now, a second test should be performed.
- Which scenario is to be preferred?

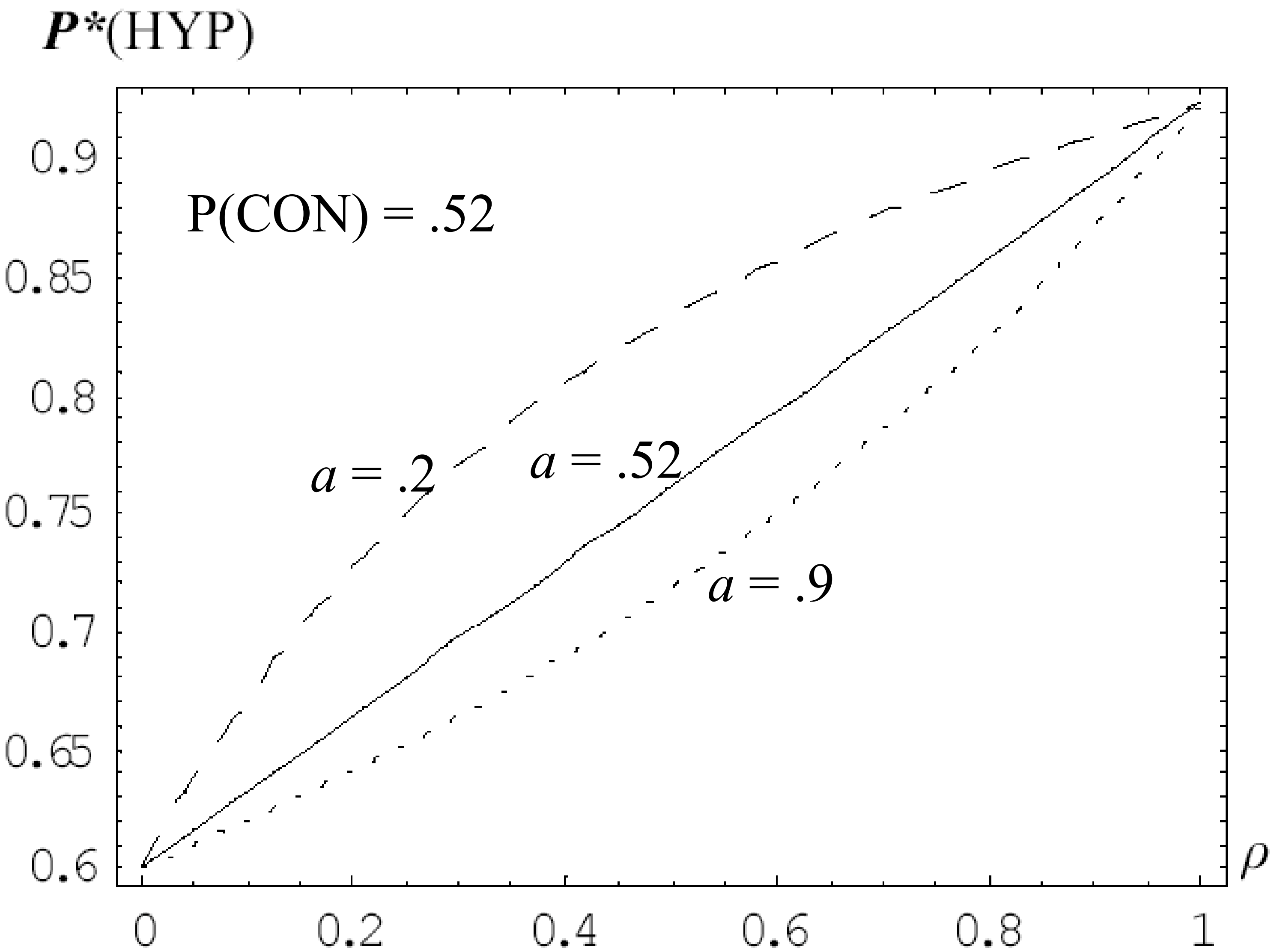
  Scenario 1: use another instrument I_2 (i.e. the tests are independent)

  Scenario 2: use the same instrument again (i.e. the tests are dependent)
- *Variety-of-Evidence Thesis*: Scenario 1 leads to a higher posterior probability and is to be preferred.

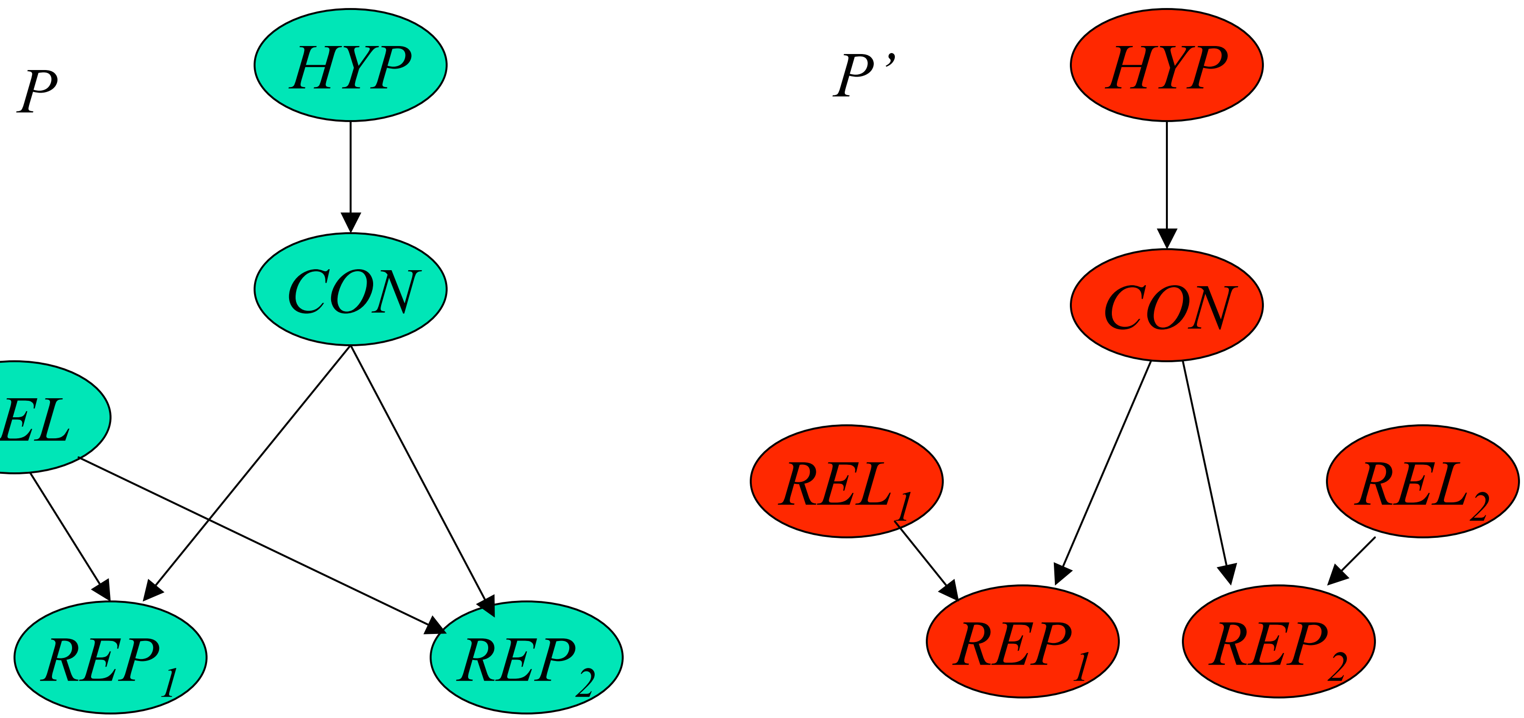
But is the thesis true? → construct a model!

The Basic Model



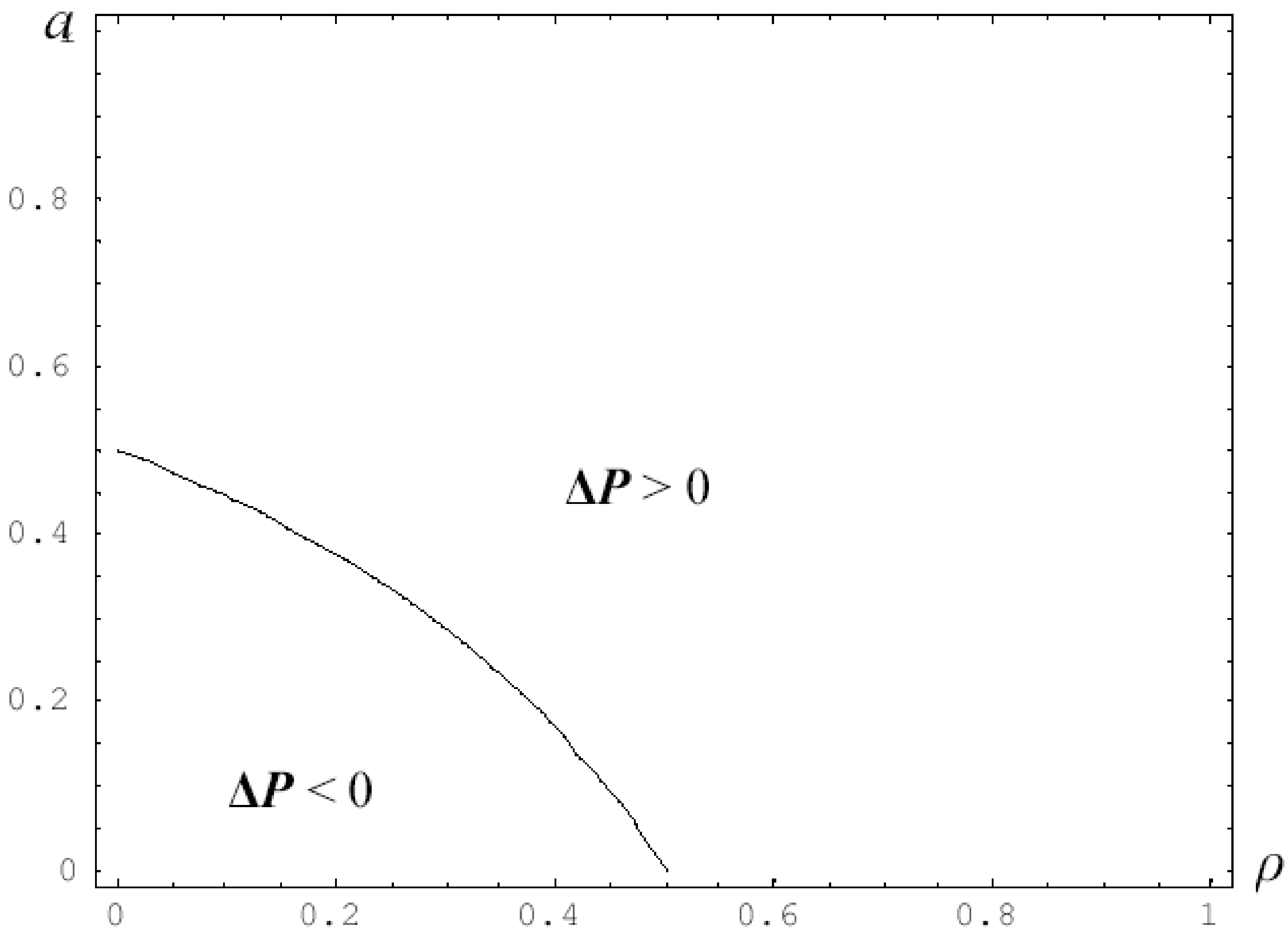

P*(HYP)
P(CON) = .52
a = .2
a = .52
a = .9
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0
0.2
0.4
0.6
0.8
1
ρ

Single vs. Multiple Instruments



The Relative Strength of Confirmation

- Use the theory of Bayesian Networks to calculate the posterior probability of the hypothesis for both cases!
- To find out which procedure leads to more confirmation, calculate the difference
  $\Delta P = P'(\text{HYP}|\text{REP}_1, \text{REP}_2) - P(\text{HYP}|\text{REP}_1, \text{REP}_2)$
- After some algebra, one obtains:
  $\Delta P > 0$ iff $\quad 1 - 2(1 - a)(1 - \rho) > 0$


a
0.8
0.6
0.4
0.2
0
$\Delta P > 0$
$\Delta P < 0$
0
0.2
0.4
0.6
0.8
1
ρ

Interpretation

There are two conflicting considerations:

1. Independent test results from two instruments yield stronger confirmation than dependent test results from a single instrument.
2. Coherent test results obtained from a single instrument increase our confidence in the reliability of the instrument, which increases the degree of confirmation of the hypothesis.

The *variety-of-evidence thesis* challenged

Under certain conditions, test results from a single test instrument provide greater confirmation than test results from multiple independent instruments.

Upshot

- We have constructed and analyzed a *toy model* (which is still *very* far away from real science).
- The model shows that the variety-of-evidence thesis is not sacrosanct. It can fail, and it does fail under certain circumstances.
- The model is used as an exploratory tool. It suggests something, but the results need to be interpreted (ideally) model independently.

Outline

I. Bayesianism vs. Reliabilism
II. Motivation
III. Modeling in Science
IV. Textbook Bayesianism
V. Bayesian Networks
VI. Example 1: Variety-of-Evidence Thesis
VII. Example 2: Testimony
VIII. Example 3: Scientific Theory Change
IX. Naturalized Bayesianism
X. Conclusions

VII. Example 2: Testimony

Line up

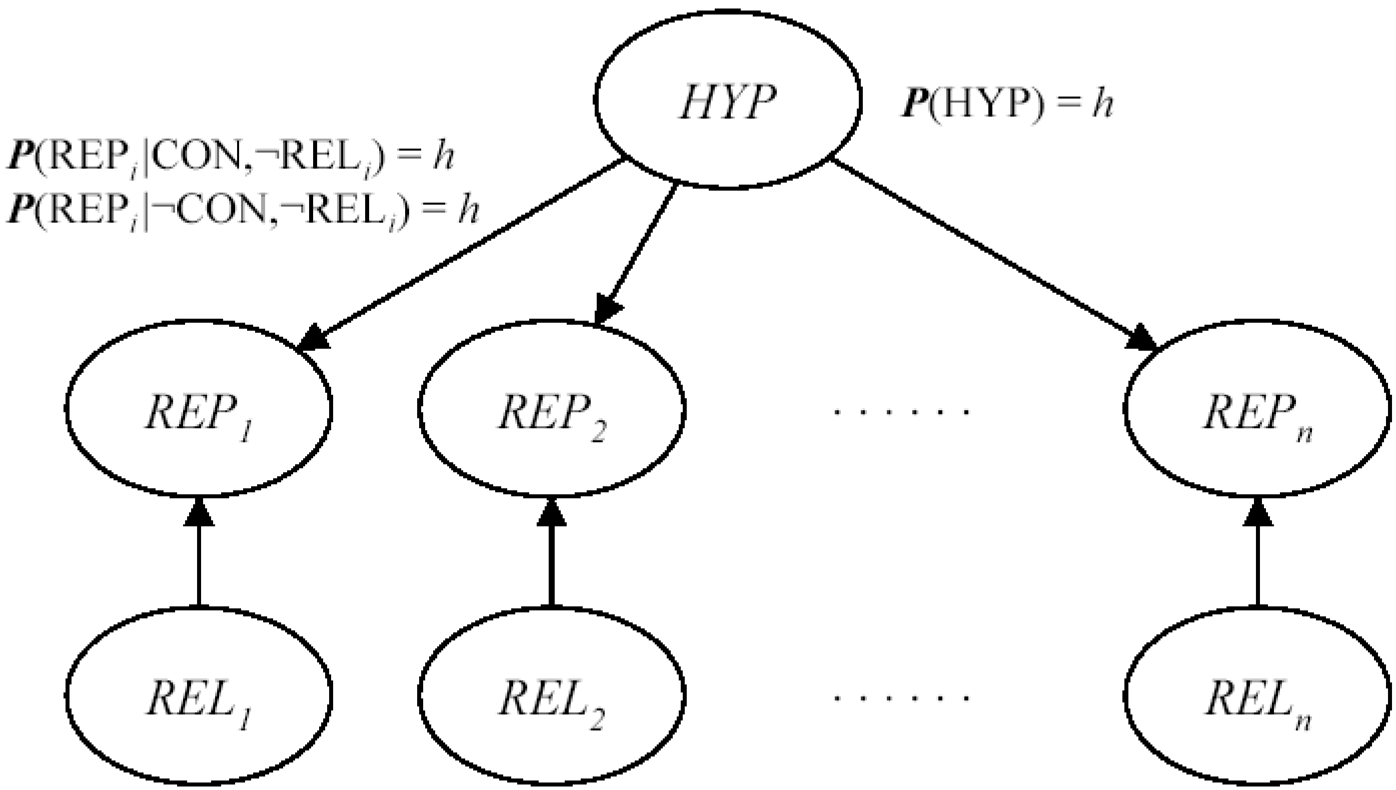
- Suppose that there are n equiprobable suspects for a crime and exactly one suspect is the culprit.
- The witnesses are either reliable, i.e. truth-tellers, or they are unreliable.
- If unreliable the chance that they say x is the culprit is $1/n$.
- What is the chance that x is the culprit given that all witnesses say that x is the culprit?

Ekelöf's Claim

- The more surprising it is that *x* is the culprit – i.e. the greater the number of suspects – the greater the chance is that *x* is the culprit (given resp. positive reports).

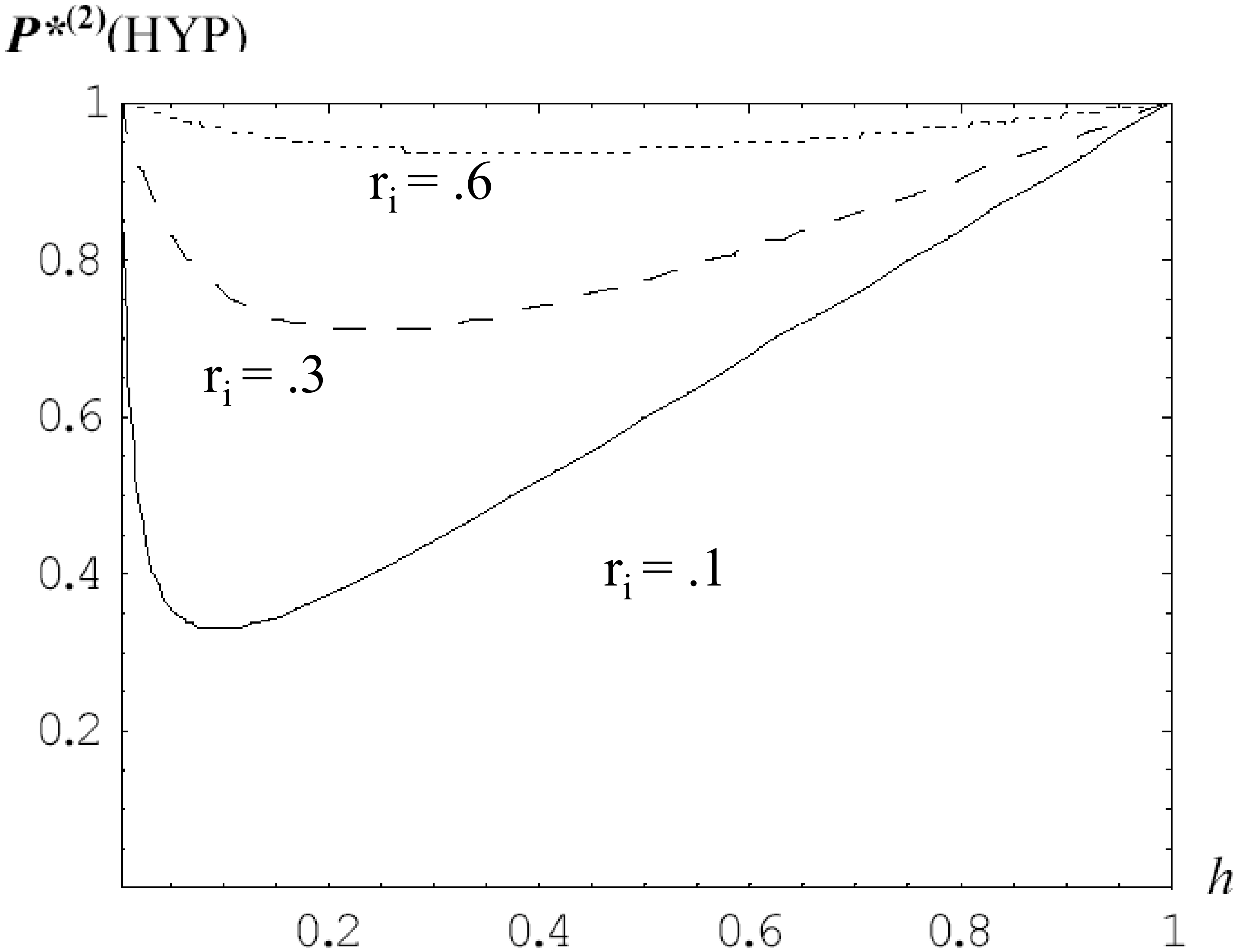
  - Compare: 3 suspects vs. 100 suspects
- The smaller the prior probability that *x* is the culprit, the greater the posterior probability that *x* is the culprit.
- Let's see if we can model this!

The Model


HYP
P(HYP) = h
P(REP_i|CON,¬REL_i) = h
P(REP_i|¬CON,¬REL_i) = h
REP_1
REP_2
REP_n
REL_1
REL_2
REL_n

After some math…

$$\boldsymbol{P}^{*(2)}(\mathrm{HYP}) = h^* = \frac{(h + \bar{h}r_1)(h + \bar{h}r_2)}{h + \bar{h}r_1 r_2}$$


P*(2)(HYP)
1
0.8
0.6
0.4
0.2
0.2
0.4
0.6
0.8
1
h
r_i = .6
r_i = .3
r_i = .1

Find Minima

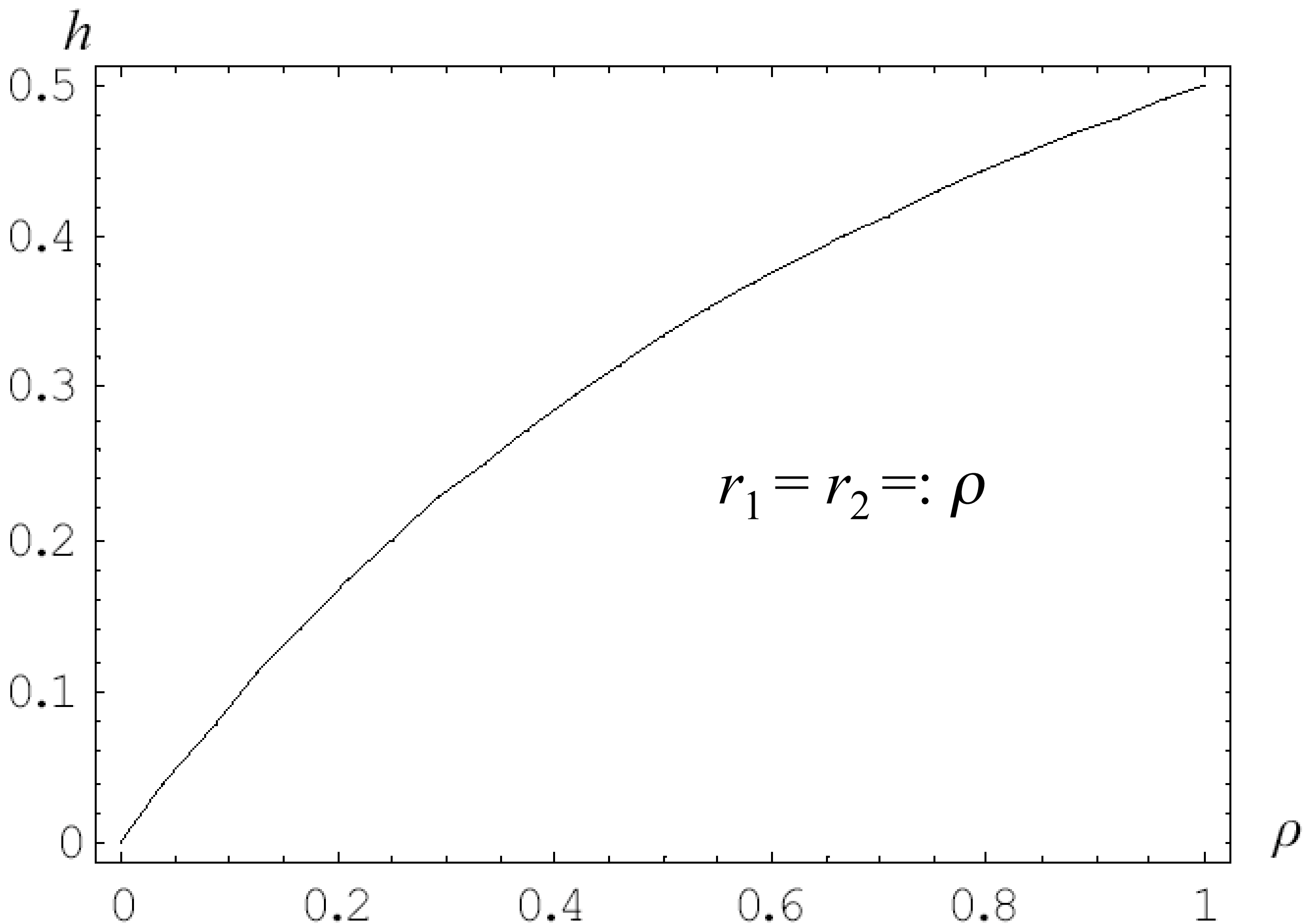
- Find partial derivatives with respect to *h:*

$$\frac{\partial h^*}{\partial h} = \bar{r}_1 \bar{r}_2 \frac{h^2 - \bar{h}^2 r_1 r_2}{(h + \bar{h} r_1 r_2)^2}$$

- Set partial derivative at 0 and solve for *h*:

$$h_{min} = \frac{\sqrt{r_1 r_2}}{1 + \sqrt{r_1 r_2}}$$

Ekelöf's claim holds true for area under the curve for two witnesses…

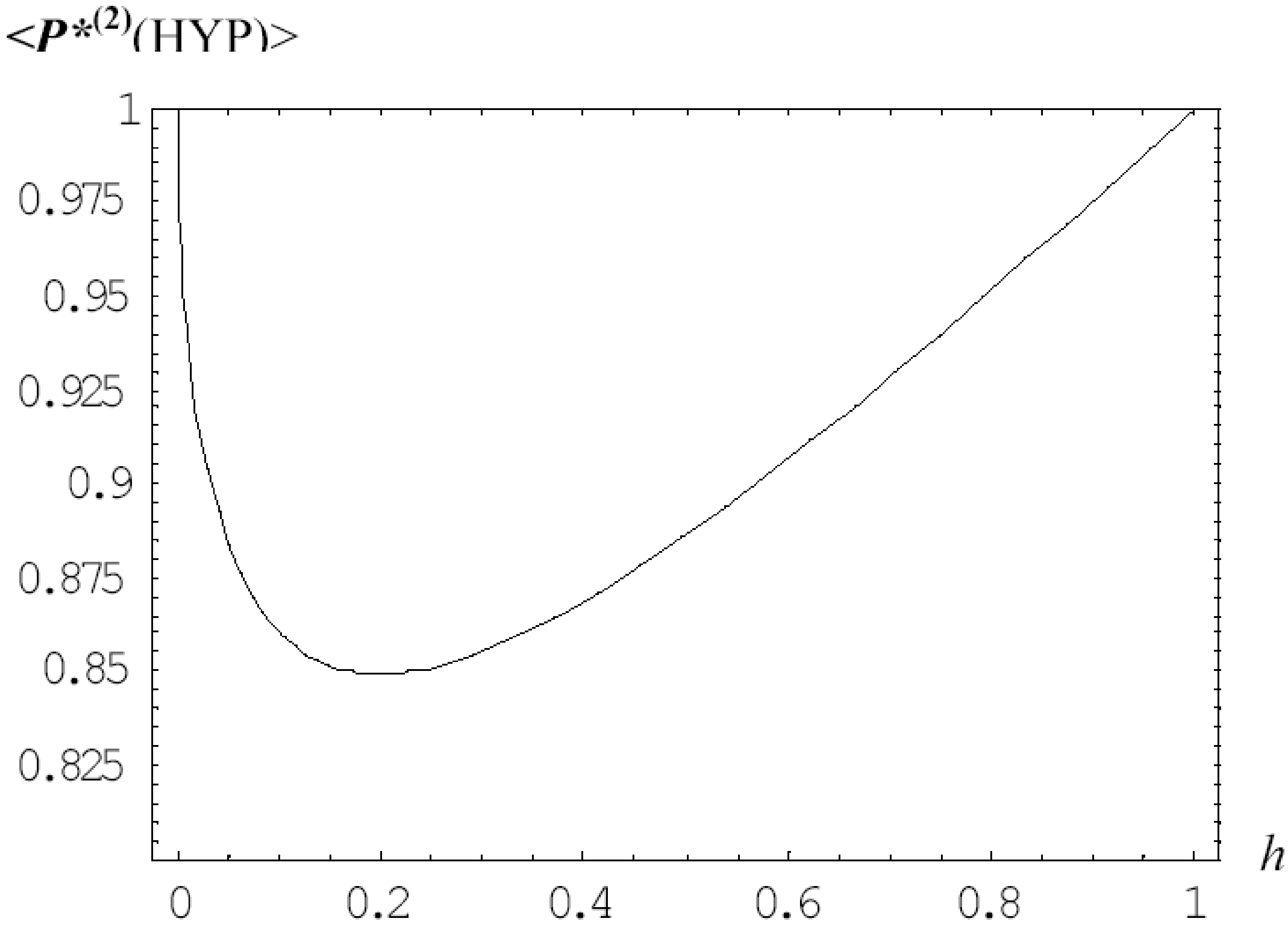


From Risk to Uncertainty

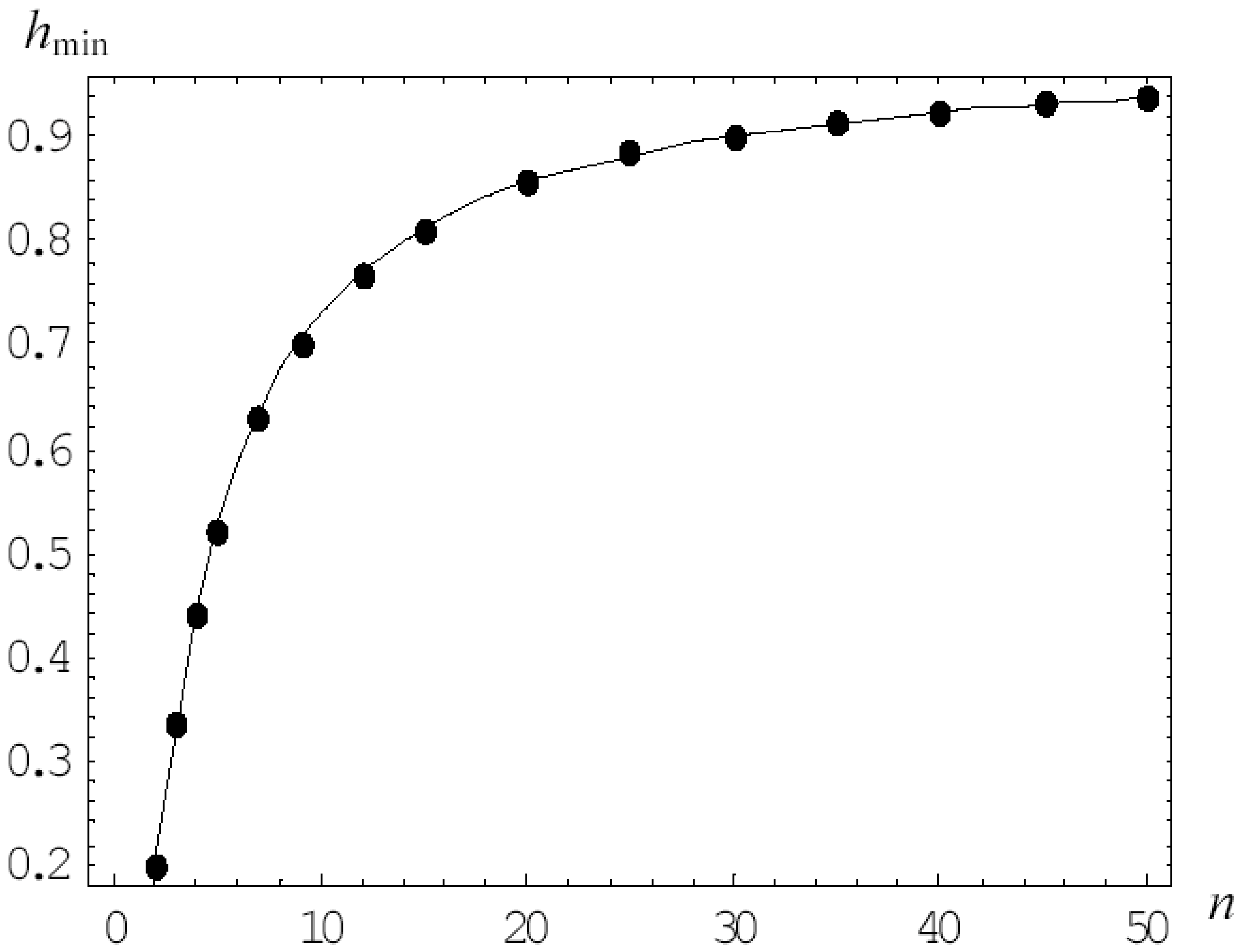
- Decision making under uncertainty: no information about the reliability. Assume a uniform distribution and „average the reliability out“:

$$h^* = \int_0^1 \int_0^1 \frac{(h + \bar{h} r_1)(h + \bar{h} r_2)}{h + \bar{h} r_1 r_2} dr_1 dr_2$$

Uncertainty and $n = 2$



h_{min} for Uncertainty and for n Witnesses



Ekelöf Partially Vindicated

- When there are five or more independent witnesses whose degree of reliability is unknown, then the smaller the prior probability h for $h \in (0, 1/2)$ that x is the culprit, the greater the posterior probability that x is the culprit.
- Hence, agreement amongst five or more witnesses that one out of 3 possible suspects is the culprit is better than that one out of 2 possible suspects is the culprit.

VIII. Example 3: Scientific Theory Change

The problem

- Scientific theories change over time. What changes, and what remains?
- The philosophical debate so far has been too much focused on the very big question as to whether a rational reconstruction of theory change can be given.
- My aim is more modest: I would like to explain some aspects of scientific theory change.
- Here are two examples…

Intermezzo: What is a scientific theory?

- Textbook Bayesianism has no account of what a scientific theory is. This is one of its (many) shortcomings.
- I propose a Bayesian account.

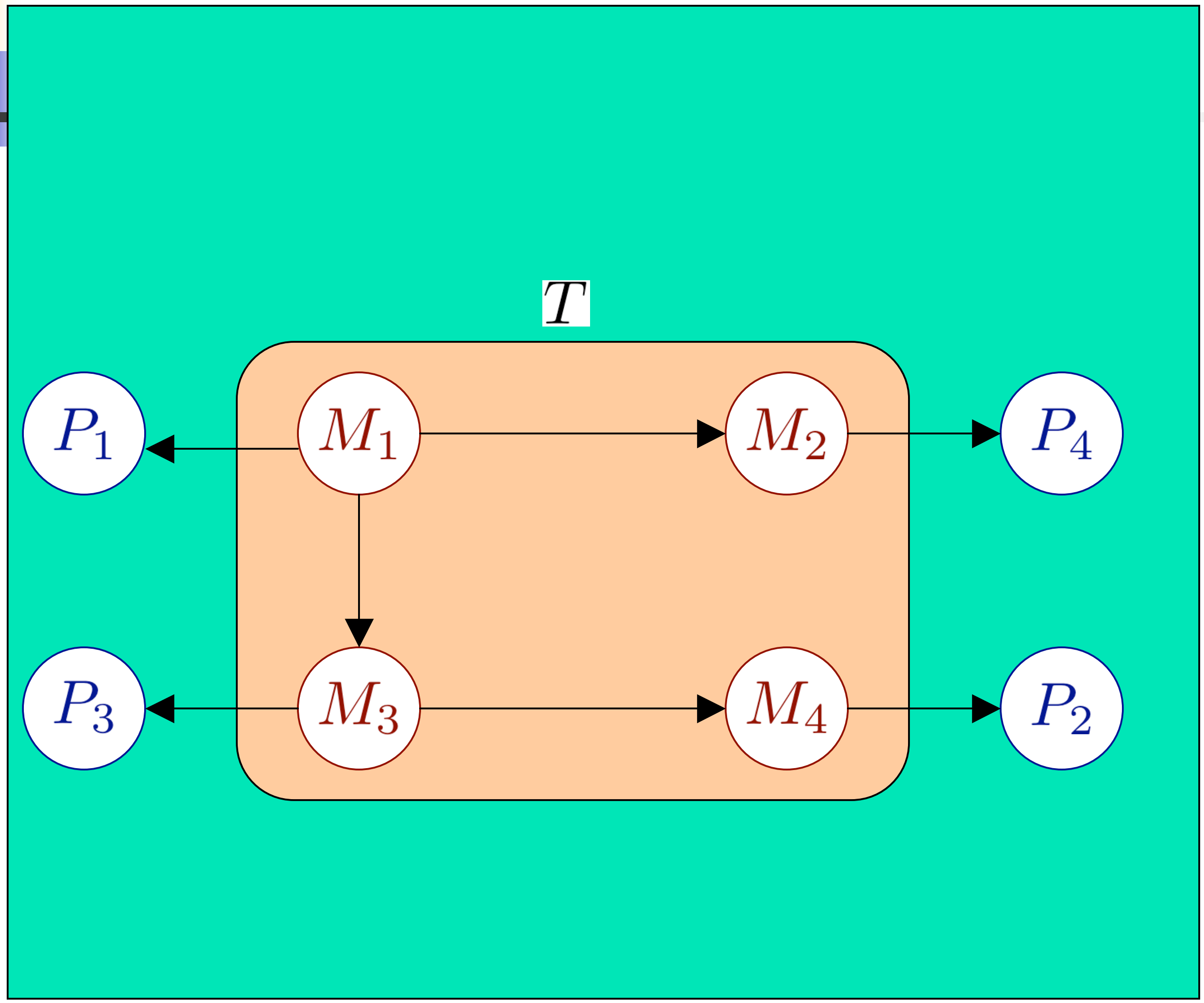
Received views

- Syntactic view:
  - theories are linguistic entities
  - sets of assumptions (and their consequences)
- Semantic view:
  - theories are non-linguistic entities
  - realizations of an abstract formalism

The probabilistic view

- Theories are networks of interrelated models.
- Models (M_i) are conjunctions of propositions that account for a specific phenomenon P_i (e.g. instantiations of laws). One model for each phenomenon. Conditional independence assumptions hold.
- There is a joint probability distribution over all propositional variables M_i, P_i.
- From this, the posterior probability of the theory (given the phenomena) can be obtained.

Representing theories by Bayesian Networks



The empirical and the non-empirical

- What I just presented represents the *empirical part* of the theory.
- Additionally, a theory has a *non-empirical* (or heuristic) part that helps us to construct models (cf. theories as modeling frameworks), might be axiomatized, contains (non-probabilified) laws, etc.

(i) The stability of normal science

- In normal science, more and more applications of a theory are considered. The goal here, however, is not to test the theory. It is taken for granted and applied. It is rather the scientist who is tested.
- An immediate problem for the Bayesian:

  - $P(M_1, M_2, M_3) < P(M_1, M_2)$

  - so by adding more and more applications of a theory, the joint probability will sooner or later be below the threshold (if there is any). I refer to this as the *conjunction problem*.
- So can (the rationality of) normal science be defended in a Bayesian framework?

Coherence

- The conjunction problem can be (dis-)solved by taking evidence (P_i) into account. Then it is possible that

  $P(M_1, M_2, M_3|P_1, P_2, P_3) > P(M_1, M_2|P_1, P_2)$
- Question: When is this the case?
- Answer: There must be enough positive relevance among the models (and the evidence must confirm the models).
- One way to measure how relevant the models are for each other is by examining how well they fit together or cohere.

A theorem

Let T = {$M_1,\ldots, M_n$} and T' = T ∪ {M_{n+1}}. Each model M_i is independently supported by a piece of evidence E_i. Let E = {$E_1,\ldots, E_n$} and E' = E ∪ {E_{n+1}} and let the strength of the evidence be characterized by $r_i = 1 - P(E_i|\neg M_i)/P(E_i|M_i)$ (for $i = 1,\ldots, n+1$). Let $a_0 := P(T)$ and $b_0 := P(T')$.

If

(i) all models are supported with the same strength of evidence (i.e. $r_1 = r_2 = \ldots = r_{n+1} > r_0 := (a_0 - b_0)/[a_0(1 - a_0)]$), and

(ii) T' is more coherent than T,

Then

the posterior probability of T' is greater than the posterior probability of T, i.e. P(T'|E) > P(T|E). (Proof omitted)

How does this theorem help?

1. It is plausible that, within normal science, theories become more and more coherent.
2. I is also plausible that the prior of the theory does not change much if a model is added, and that all models are (more or less) equally well supported.
3. Then the theorem tells us that the posterior probability of the theory increases in the course of normal science. QED.

(ii) Why anomalies hurt so much

- Here I depart from Kuhn.
- Anomalies (such as the discovery of line spectra in the late 19th century) hurt the whole theory, and not just a part of it.
- As a consequence, the whole theory is (and, perhaps, has to be) given up (or at least its domain of applicability has to be restricted).
- What is the rationale for this?

Coherence and confirmation transmission

- In recent work, Dietrich & Moretti (2006) showed that sufficiently coherent sets of propositions *transmit* confirmation.
- I.e. if *E* confirms one of the propositions of a theory (which is, remember, a highly coherent system), it also confirms any other proposition of the theory as well as the whole theory (i.e. the conjunction of all propositions).
- This is a rationale for indirect confirmation.

…and disconfirmation

- The same holds, of course, also for disconfirmation.
- So an anomaly (= there is no model for a specific phenomenon or all models fail to account for the phenomenon) hurts the whole theory in normal science (which is, again, highly coherent), and this can be fatal for it.

But what about theory change?

- Here more case studies are needed…
- Examine case studies and extract features that survive theory change, and those that do not, and try to explain this in a formal model.
- Current research:

  \- formulate an account of *intertheory relations* that avoids the extremes - reductionism à la Nagel and an untamed pluralism (à la Cartwright and Dupré).

  \- coherence and unification

IX. Naturalized Bayesianism

- Naturalized Bayesianism is a middle ground between

  - aprioristic Textbook Bayesianism, and

  - naturalized philosophies of science
- Bayesianism is a modeling framework, and
- (generalizations from) case studies provide the necessary "empirical" input (which I hope to get from historical studies).

X. Conclusions

- Philosophical models are heuristically important and help us explain features of the methodology of science.
- They differ from scientific models as we only have our *intuitions* to test them.
- There is a *reflexive equilibrium* between our intuitions and the consequences of the model.
- Though Naturalized Bayesianism provides good explanations, some features of science might resist a Bayesian explanation. We then have to find another account to explain them. Just like in science.

XI. Coherence

1. Is Coherence Truth-Conducive?

- According to the Coherence Theory of Justification, coherence is an indicator of truth.
- Or: The more coherent a set of information is, the higher its degree of confidence.
- My goals:

  - Make this claim more precise

  - Challenge it!

The Problem

When we receive information from independent and partially reliable sources, what is our degree of confidence that this information is true?

- Independence?
- Partial reliability?

Independence

REP_i

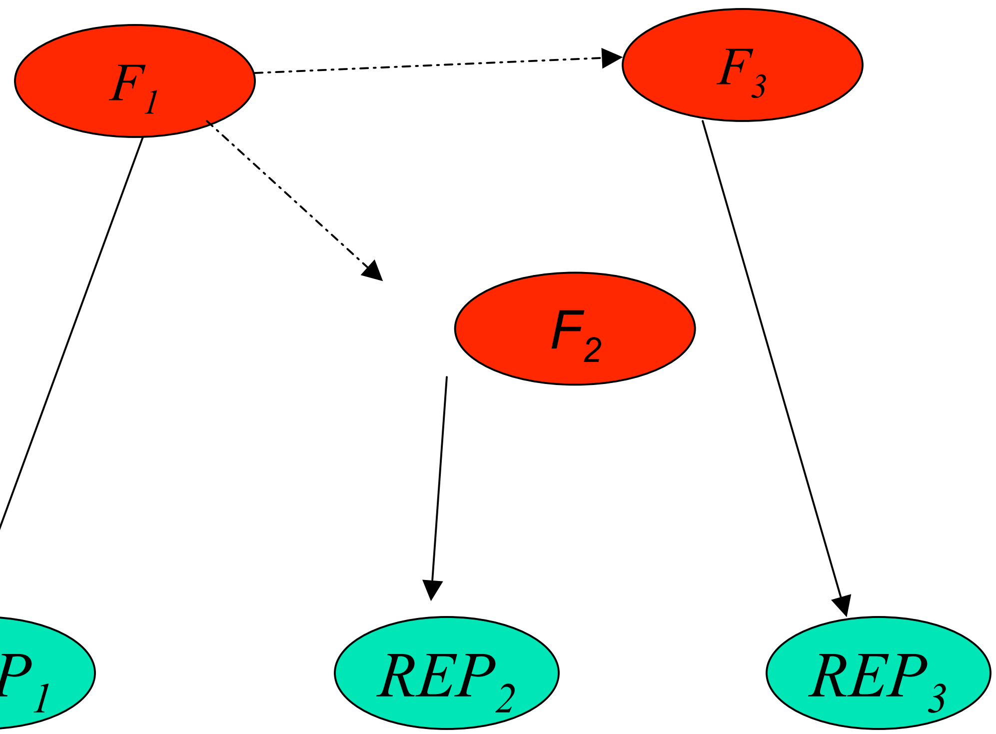
is independent of

$F_1, REP_1, F_{i-1}, REP_{i-1}, F_{i+1}, REP_{i+1}, F_n, REP_n$

given F_i

A Bayesian Network Model

$P(\mathrm{REP_i}|\mathrm{F_i}) = p$
$P(\mathrm{REP_i}|\neg \mathrm{F_i}) = q$
for $p > q$



Determinants of the Degree of Confidence

1. How *expected* is the information?
2. How *reliable* are the sources?
3. How *coherent* is the information?

Expectancy

- Example: Medical tests: Locus of gene
- There is some *prior knowledge* about the locus of the gene
- Compare two cases with two tests each
- Case 1: Overlapping area is expected
- Case 2: Overlapping area is not expected

- Case 1 has a higher degree of confidence.

Reliability

- Same setup
- Only difference:

  The tests in case 1 are more reliable than in case 2.

- Case 1 has a higher degree of confidence.

Coherence

- Same setup
- Only difference:

  Let the overlapping area be the same in cases 1 and 2, yet the non-overlapping area in case 2 is larger than the non-overlapping area in case 1.

- Case 1 has a higher degree of confidence.

A Conjecture

The three determinants are *separable*, i.e.

1. The more reliable the information sources are, the greater our degree of confidence, *ceteris paribus*.
2. The more plausible the information is, the greater our degree of confidence, *ceteris paribus.*
3. The more coherent the information is, the greater our degree of confidence, *ceteris paribus*.

My Goal

Challenge conjecture 3

(i.e. a central claim of the Coherence Theory of Justification)

Reliability

$P(\mathrm{REP}_i|\mathrm{F}_i) = p$ (for all $i = 1,\ldots, n$)

$P(\mathrm{REP}_i|\neg \mathrm{F}_i) = q$ with $p > q$

$r := 1 - q/p$

Randomization Full Reliability

0---1

Expectancy and Coherence

a_i with i counting the # of negative values

$a_0 = .05$

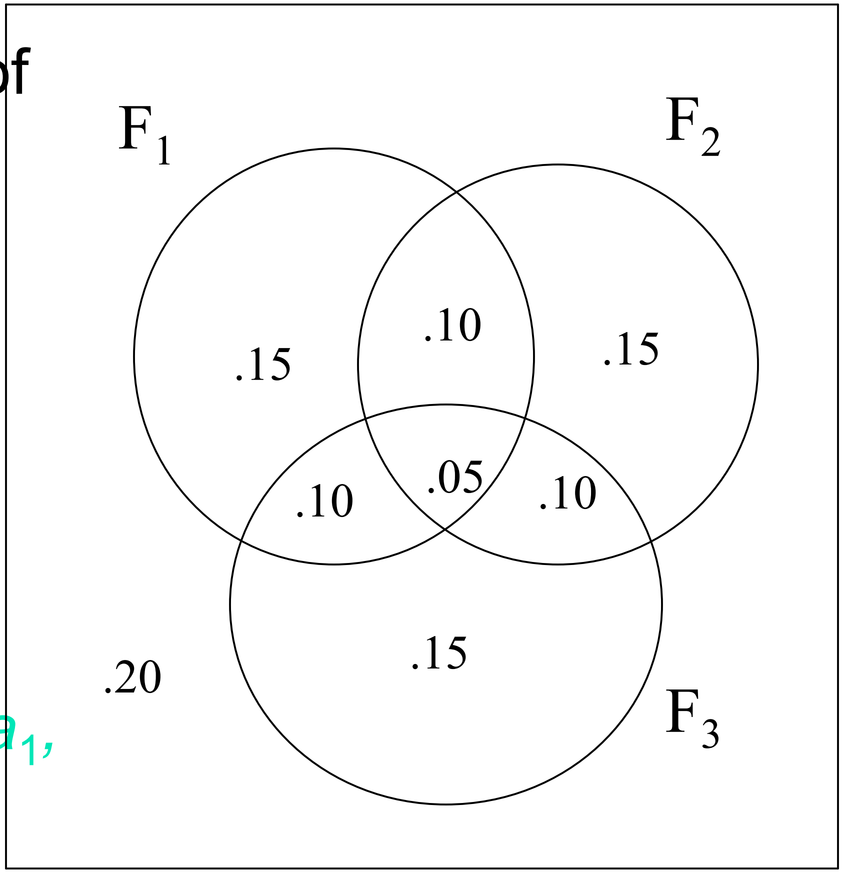
$a_1 = 3 \times .10 = .30$

$a_2 = 3 \times .15 = .45$

$a_3 = .20$

Expectancy → a_0

Coherence → $f(<a_1, a_2, a_3>)$



Degree of Confidence

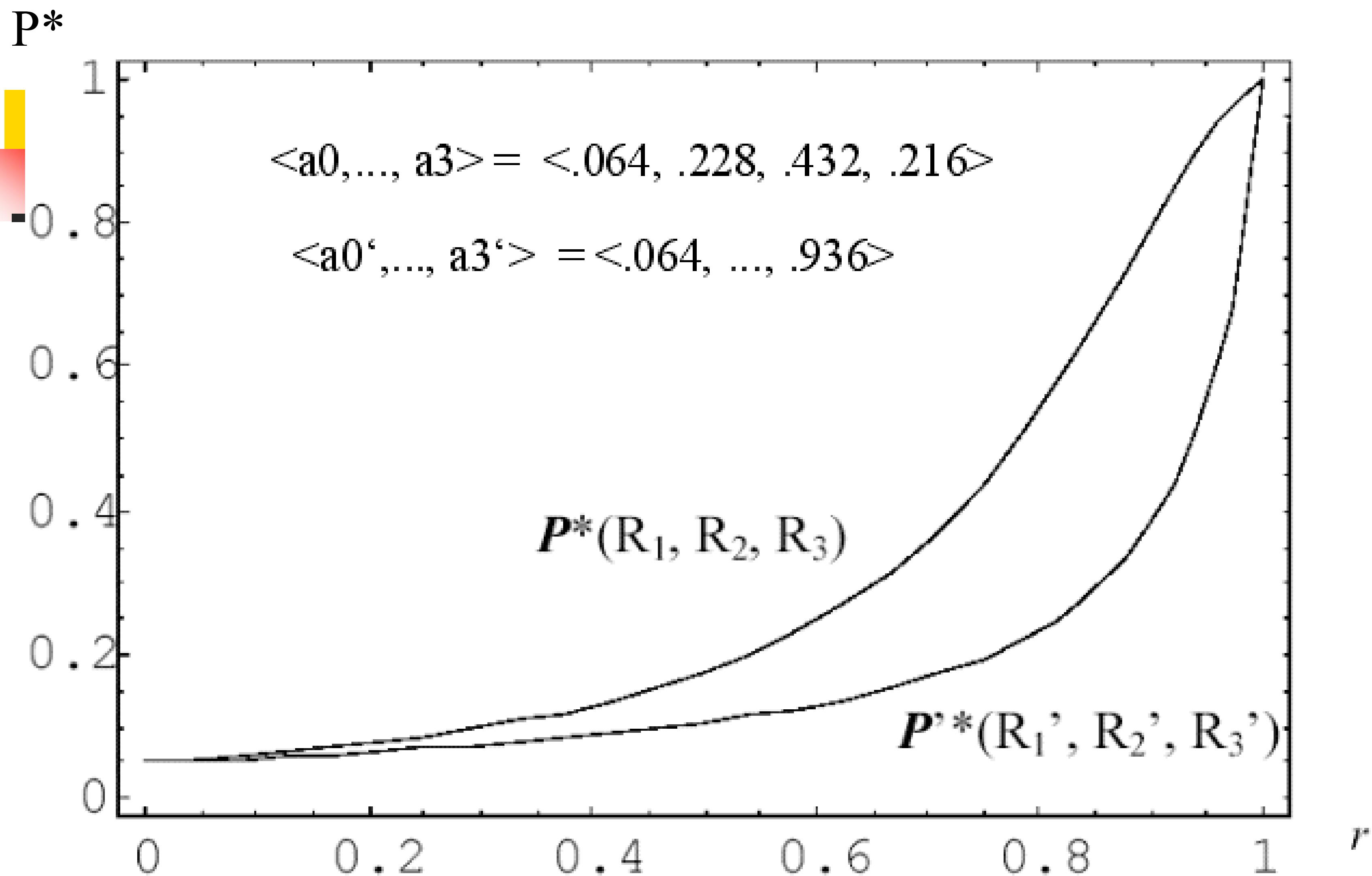
P^* $= P^*(F_1,\ldots, F_n)$ $=$
$P(F_1,\ldots,F_n|REP_1,\ldots, REP_n) =$

$$\frac{a_0}{\sum_{i=0}^{n} a_i (1-r)^i}$$

An Example for n = 3

F_1 = The culprit was wearing Coco Chanel shoes.
F_2 = The culprit had a French accent.
F_3 = The culprit drove a Renault.

F_1' = The culprit was a woman.
F_2' = The culprit had a Flemish accent.
F_3' = The culprit drove a Ford.

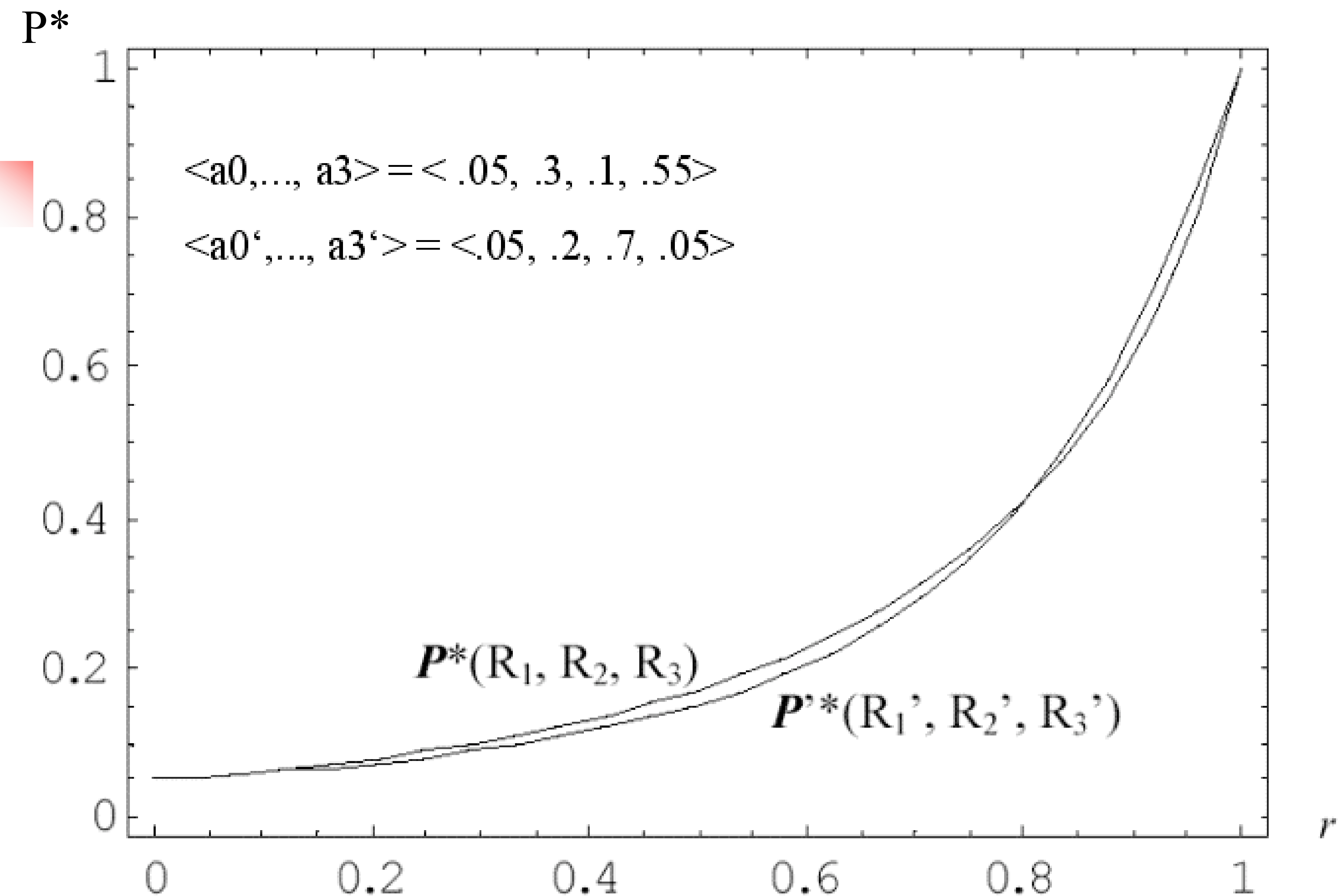

P*
<a0,..., a3> = <.064, .228, .432, .216>
<a0',..., a3'> = <.064, ..., .936>
P*(R1, R2, R3)
P'*(R1', R2', R3')
0
0.2
0.4
0.6
0.8
1
r

Note: One counter example is enough to show that conjecture 3 is false.

Assumptions:

(1) There are no further determinants.

(2) The coherence only depends on $<a_0,\ldots, a_n>$.


P*
<a0,…, a3> = < .05, .3, .1, .55>
<a0‘,..., a3‘> = <.05, .2, .7, .05>
$P^*(R_1, R_2, R_3)$
$P'^*(R_1', R_2', R_3')$
0
0.2
0.4
0.6
0.8
1
r

Separability Challenged

Separability: The more coherent the information is, the greater our degree of confidence, *ceteris paribus*.

(1) Suppose that S is more coherent than S':
Separability fails when we set a_0 = .05 and r = .90.

(2) Suppose that S' is more coherent than S:
Separability fails when we set a_0 = .05 and r = .10.
∴Separability fails

An Example

- F_1 = The culprit is a woman.
- F_2 = The culprit has a Flemish accent.
- F_3 = The culprit drove a Ford.

- F_1‘ = The culprit wore Coco Chanel shows.
- F_2‘ = The culprit has a French accent.
- F_3‘ = The culprit drove a **Ford**.

What to Do?

Do we have to give up separability? This would have severe consequences for the Coherence Theory of Justification.

Check all assumptions we made!

Four *prima facie* plausible assumptions characterize *Bayesian Coherentism*...

(i) Separability

The more coherent the information is, the greater our degree of confidence, *ceteris paribus*.

(ii) Probabilism

The coherence of new information items is a function of probabilistic features of the information items.

(iii) Holism

The relation ‘is more coherent than’ is defined over information sets of size $n \geq 2$.

(iv) Ordering

The relation ‘is more coherent than’ is an ordering, i.e. it is reflexive, transitive and complete.

Which Assumption Should be Given Up?

- *Separability*: Drop the Coherence Theory of Justification.
- *Probabilism*: Drop the Bayesian Coherence Theory of Justification.
- *Holism*: Lehrer > Bonjour
- *Ordering*: $\Rightarrow$ Quasi-Ordering

3. How Can one Measure the Coherence of an Information Set?

Goal of this section: Construct a measure that induces a quasi-ordering of information sets.

Case I

- S = {[the culprit was a woman], [the culprit had a Flemish accent], [the culprit drove a Ford]}
- S' = {[the culprit was wearing Coco Chanel shoes], [the culprit had a French accent], [the culprit drove a Renault]}.
- S" = {[the culprit was wearing Coco Chanel shoes], [the culprit had a French accent], [the culprit drove a Ford]}.

$\Rightarrow$ S' > S $\Rightarrow$ S" > S and S > S"

Case II

- B' = {[all ravens are black], [this bird is a raven], [this bird is black]}
- B = {[this chair is brown],[electrons are negatively charged], [today is Thursday]}

$\Rightarrow$ B' > B

Case III

- T' = {[Tweety is a bird], [Tweety cannot fly], [Tweety is a penguin]}
- T = {[Tweety is a bird], [Tweety cannot fly]}

$$\Rightarrow \qquad T' > T$$

Lewis's Proposal

$\{R_1, \ldots, R_{i-1}, R_{i+1}, \ldots, R_n\}$ is coherent (or congruent)

iff

$\boldsymbol{P}(R_i | R_1, \ldots, R_{i-1}, R_{i+1}, \ldots, R_n) > \boldsymbol{P}(R_i)$ for all $i = 1, \ldots, n$.

Olsson's Proposal

$\{R_1, \ldots, R_m\} > \{R_1', \ldots, R_n'\}$

iff

$$\frac{P(R_1, \ldots, R_m)}{P(R_1 \vee \ldots \vee R_m)} \geq \frac{P(R_1', \ldots, R_n')}{P(R_1' \vee \ldots \vee R_n')}$$

Shogenji's Proposal

$\{R_1,\ldots, R_m\} > \{R_1',\ldots, R_n'\}$

iff

$$\frac{P(R_1,\ldots,R_m)}{\prod_{i=1}^{n} P(R_i)} \geq \frac{P(R_1',\ldots,R_n')}{\prod_{i=1}^{n} P(R_i')}$$

Fitelson's Proposal

K-O Measure of Confirmation:

$$\boldsymbol{F}(R_1,R_2)=\frac{\boldsymbol{P}(R_1 \mid R_2)-\boldsymbol{P}(R_1 \mid \neg R_2)}{\boldsymbol{P}(R_1 \mid R_2)+\boldsymbol{P}(R_1 \mid \neg R_2)}$$

$\{R_1, R_2\} > \{R_1, R_2'\}$

iff

$$\frac{\boldsymbol{F}(R_1,R_2)+\boldsymbol{F}(R_2,R_1)}{2} \geq \frac{\boldsymbol{F}(R_1',R_2')+\boldsymbol{F}(R_2',R_1')}{2}$$

Moral

Intuitive proposals give us no more than a partial elucidation of certain aspects of coherence (positive relevance, overlap,…) but do not lead to a unitary account.

We need a measure that that weights these factors in an appropriate way, and there is no principled way to do this.

Way out: Think about the *function* of coherence.

Coherence of What?

Ant hills ~ fitness

Law firms ~ productivity

Families ~ happiness

Coherence of … is the property of … which increases ___ and is the neighborhood of our pre-theoretical notion of coherence.

Coherence of Information Sets

The coherence of an information set $\{R_1,\ldots, R_n\}$ is the property of this information set which increases the confidence boost that results from being informed by independent of partially reliable sources that respectively $R_1, \ldots, R_n$ are the case.

Tentative Proposal

We measure the coherence by the degree-of-confidence boost, i.e. the ratio of the prior over the posterior probability

$$c_r^{tent}(S) = \frac{\boldsymbol{P}^*(\mathrm{F}_1, \ldots, \mathrm{F}_n)}{\boldsymbol{P}(\mathrm{F}_1, \ldots, \mathrm{F}_n)}$$

Problem #1

- Let S contain highly coherent information with $\boldsymbol{P}(F_1, F_2) \approx 1$ and let S' contain highly incoherent information with $\boldsymbol{P'}(F_1', F_2') \approx 0$

- Then for any value of r, $c_r^{tent}(S') > c_r^{tent}(S)$

Solution: Normalization

- We measure the coherence by the degree of confidence boost that actually obtained over the degree of confidence boost that would have obtained, had the same information been presented as maximally coherent information.

The Measure

$$c_r(\Sigma) = \frac{\frac{P^*(\mathrm{F}_1,\ldots,\mathrm{F}_n)}{P(\mathrm{F}_1,\ldots,\mathrm{F}_n)}}{\frac{P^{max*}(\mathrm{F}_1,\ldots,\mathrm{F}_n)}{P^{max}(\mathrm{F}_1,\ldots,\mathrm{F}_n)}}$$

$$= \frac{a_0 + (1 - a_0)(1 - r)^n}{\sum_{i=1}^{n} a_i (1 - r)^i}$$

Problem #2

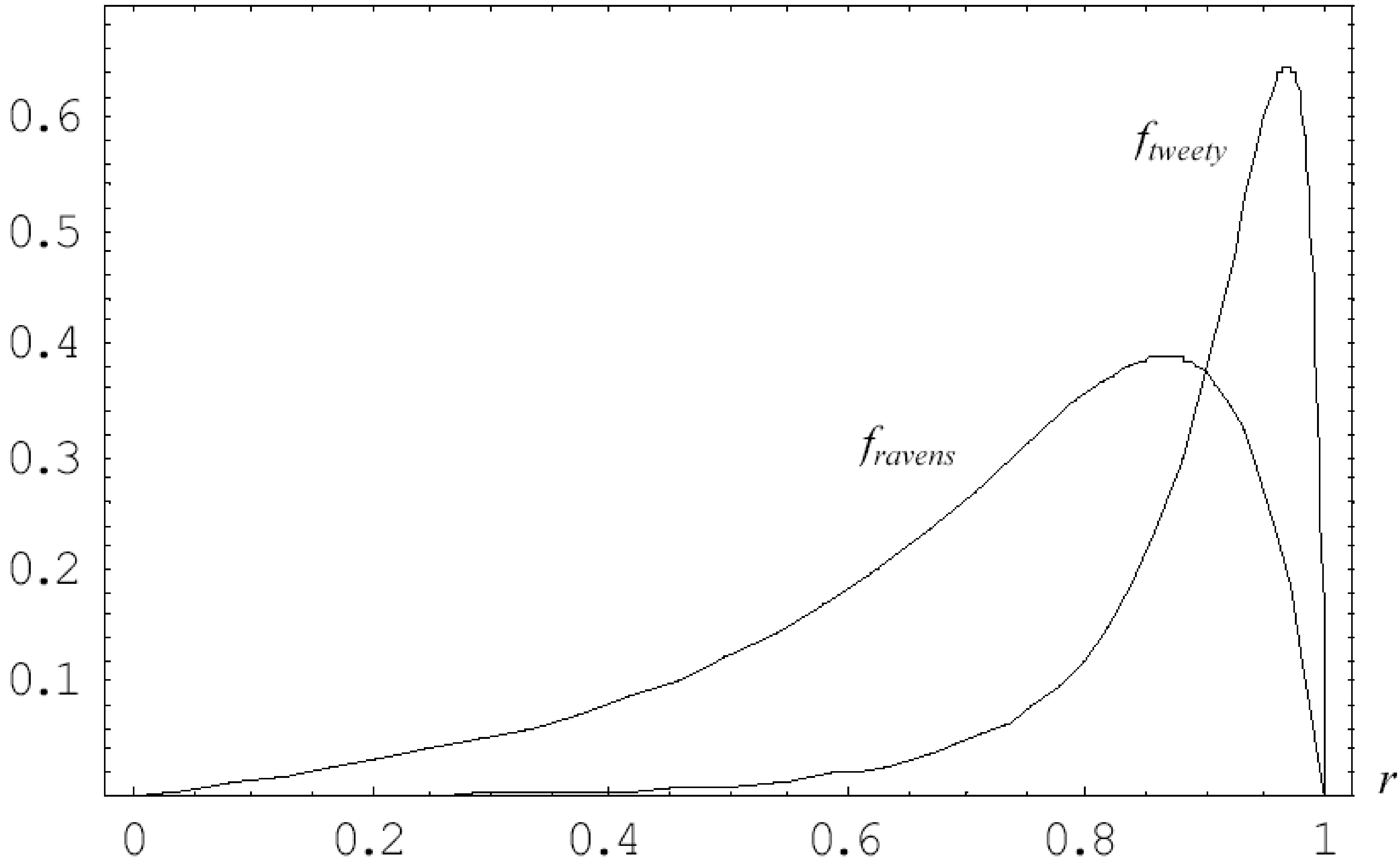
A coherence measure should not be dependent on the reliability of the informants *r.*

Solution

S > S' iff

$c_r(S) > c_r(S')$ for all $r \in (0, 1)$

Construct $f_r(S, S') := c_r(S) - c_r(S')$



0.6
0.5
0.4
0.3
0.2
0.1
0
0.2
0.4
0.6
0.8
1
r
f_{tweety}
f_{ravens}

Taking Stock

- We proved an impossibility theorem and discussed various ways out.
- We suggest to give up that there is a coherence ordering over information sets.
- We proposed a measure for the coherence of an information set that induces a partial ordering.
- The measure does not obtain from intuitive ideas about the nature of coherence, but about its function.

3. Open Problems

1. Come up with alternative coherence measures.
2. Testimony: Various independent witnesses report the same unlikely event.
3. Systematically explore dependencies between reports: how do they affect the degree of confidence?
4. Belief revision: Represent an information set by a Bayesian Network; new and possibly conflicting information comes in. What shall one do?

…