A Critique of the Bayesian Solutions to Some Paradoxes of Confirmation: Irrelevant Conjunction, Hempel's Paradox and Old Evidence

A Thesis Submitted to the University of Hyderabad in Fulfilment of the Requirements for the Award of the Degree of

DOCTOR OF PHILOSOPHY
IN
PHILOSOPHY

by

Philose Koshy

Department of Philosophy
School of Humanities
University of Hyderabad
Hyderabad 500046

JUNE 2013

Department of Philosophy
School of Humanities
University of Hyderabad
Hyderabad 500046, India

CERTIFICATE

This is to certify that the thesis titled **'A Critique of the Bayesian Solutions to Some Paradoxes of Confirmation: Irrelevant Conjunction, Hempel's Paradox and Old Evidence'** submitted to the University of Hyderabad in fulfillment of the requirements for the award of the Degree of Doctor of Philosophy in Philosophy is a *bona fide* record of original work done by **Mr**. **Philose Koshy** during the period of his study in the Department of Philosophy, University of Hyderabad, under my supervision and that the thesis has not been submitted to any other University or Institute of learning for the award of any degree.

Supervisor
Prof. Prajit K.Basu
Department of Philosophy

Co supervisor
Dr. B. Shobha
School of Mathematics and Statistics

Head,
Department of Philosophy

Dean,
School of Humanities

Department of Philosophy
School of Humanities
University of Hyderabad
Hyderabad 500046, India

DECLARATION

I hereby declare that this thesis titled **A Critique of the Bayesian Solutions to Some Paradoxes of Confirmation: Irrelevant Conjunction, Hempel's Paradox and Old Evidence'** submitted for the award of the Degree of Doctor of Philosophy in Philosophy to the University of Hyderabad, embodies the result of *bona fide* research work carried out by me under the supervision of **Prof. Prajit K.Basu.** It has not been submitted to any other University or Institute of learning for the award of any degree.

S*ignature of the Candidate*

Date: **28.06.2013**

Place: **Hyderabad** **Philose Koshy**
06HPPH01

Acknowledgments

Writing this page necessarily took me several years back. I remember the various moments which determined and defined my path. A few among them were my solitary moments. But most of them were the moments shared with my teachers, friends and family. It seems that a few of the moments tell what they mean to me, to my work, to the path which I chose and travelled.

Many years back, my supervisor told me, apparently criticizing the way I work, that ‘philosophy thrives on criticism’. I cannot encapsulate his supervision in any other better way than this. My philosophy-days and this thesis thrive on my supervisor’s critical analyses which have guided me to find my ways of philosophizing. Unfamiliar ways, the minute focus and its meticulousness have often surprised and even disappointed me but only to revive later with greater spirit and insight. I learned a new and inspiring sense of thinking under his supervision, especially the ability to think through writing. The years of guidance have drastically redefined the ways of my writing. Now, often it seems to me that my pen thinks for me. Such was the drastic change that his guidance brought into the ways of my writing. I express my utmost gratitude to my supervisor and teacher Professor Prajit K Basu.

Reading *The Meditations on First Philosophy* gives an impression that Descartes could easily denounce and reject most of the beliefs which he holds. His painstaking effort was to build a system of belief on a firm foundation. In the ways of philosophizing, I often find that it is difficult to denounce or reject the positions or arguments or intuitions which I consider dearer to me. Especially, I could not be sceptical about my own positions. Interactions with my co-supervisor taught me the art of precise characterizations of philosophical positions, which eventually challenged my own positions and intuitions and helped to rethink for improved positions. Each meeting with Madam challenged my positions. Each notes on my draft to show the philosophical claims surprised me to realize that I am learning an inspiring way of philosophizing. She taught me probability theory and taught me to understand and analyse claims of probability theorems. Learning the probability theory was an inspiring and challenging exercise. I thank my co-supervisor Dr. B.Shobha for helping me to learn theory of probability and checking and analysing each and every claim in my drafts as well as challenging interactions.

I thank Professor Stathis Psillos for his kind words and his valuable interaction (at HBCSE Mumbai) that helped me choose the topic of Bayesian confirmation theory as my research topic. I thank Prof. S.G. Kulkarni for his valuable comments and inspiring interactions. I also extend my sincere gratitude to all faculty members of Philosophy department. I thank Prof. M.I. Beg and Dr. Usha Mohan for their probability and statistical inference classes which substantially helped me in the work of my thesis.

I extend my heartfelt gratitude to my friends Bipin, Pranav, Dinesh, Shinu, Shiby, Alam, Roshan, Sunkanna, Jentle, Yazer, Salman,Themeem, Hanu, Abhilash, Jayasuryan, Chitti babu, Dhanaraj, Prabhakar and Vincet.

I thank Shinod for his passionate narrations of scientific theories and his trademark way of unrelenting criticisms. I thank Jolly for always trying hard to make his fairy tales of possible worlds sensible to me and for his coffee which made our historic meetings alive. I thank my cheerful and inspiring friend, Jayashree. I extend my gratitude to Sreejith for being my best critic and opponent in the philosophical discussions of our evening walk. Thanks a lot to Varghese and Jaya for their love and support.

Initial days in this campus are unforgettable for various reasons, so as Prasu and Sreekumar too. Prasu, thank you so much for everything. Your love, support and interactions helped me to grow in unimaginable ways and shaped the philosophy student and person in me. I owe a lot to my humble and generous friend. My interactions with Sreekumar are inevitable constituents of every contribution I have made in this work. Sreekumar, thank you so much for a warm, influential and challenging friendship.

I thank my Pappa and Amma for their love and support. Their vision and support helped me to reach here. Thanks a lot to my sisters, brothers, Louischan, Beenachi, John, Rufu and Adri.

Philose Koshy

CONTENTS

Introduction

Concept of justification plays an important role in all philosophical analyses. It occupies a central role especially from the beginning of modern philosophy starting from Descartes. Any analysis or debate on a topic necessarily assumes some concept of justification. Any knowledge claim necessarily employs a concept of justification. That is why analysis of knowledge considers justification as a necessary condition of knowledge. And it is interesting to note that the contemporary debates of analysis of knowledge, especially in the wake of Gettier kind counter- examples is centred on the precise characterization of the concept of justification.

1. Analysis of the Concept of Justification in Philosophy of Science

There can be various kinds of justification. There can be justification by a priori truths which can be seen in Kant's position of synthetic a priori truths or in Descartes statement like ' I think therefore I am' or there can be justification by assumption (God is not a deceiving demon therefore, my perceptions are about the actual world). The justification in the formal systems is the justification by assumptions where the assumptions are construed as axioms and definitions. Another kind of justification is justifying by appealing to authority/expert. Our testimonial knowledge comes under this category.

In this dissertation, my attempt is to analyze the concept of justification in relation with the debates in philosophy of science. In philosophy of science, the question of justification is how to justify theories and laws in science. I solely focus on the justification of theories by observation statements.

Structurally, there are only two kinds of justification: deductive and inductive. In the case of deductive justification, the premises provide a conclusive support to the conclusion; such relation is called deductive entailment relation. In the inductive justification, though truth of the premises does not entail the truth of the conclusion, it supports the truth of the conclusion. While the justification of the formal statements is deductive, most of the cases of justification in science are inductive in nature. Inductive justification is a much weaker relation than a deductive one and more inclusive too. While the deductive justifications of formal statements are called proofs, the inductive justification of theories by observation statements is known as confirmation.

2. Analysis of Confirmation

In this study, my attempt is to analyse/characterize the inductive-justificatory relation between hypothesis and observation statements, which are known as evidence. When a hypothesis or a theory is inductively justified, we consider it as a confirmed hypothesis or theory. My attempt is to analyse the notion of inductive-justificatory relation, which is known as confirmation-relation. As we have seen, confirmation is an inductive relation between hypothesis and evidence.

The question is that on what basis we can determine that a hypothesis or a theory is justified. In other words, the question is how we can precisely characterize the inductive justification or inductive support for a hypothesis provided by a set of evidence. In the case of deductive logic, premises provide a conclusive/deductive support to a conclusion. The deductive support which is called as entailment is a precisely characterized one. Deductive logic formulates a certain set of rules and methods to determine whether a conclusion is valid or not. Deductive rules determine whether a conclusion follows from its premises or not. While deductive logic informs with a precise set of rules to determine deductive support of the conclusion, its inductive counterpart awfully lacks methods and principles to determine the inductive support to hypothesis and generalization. Realm of inductive justification lacks a precise set of rules to determine valid inductive support of hypothesis. Project of confirmation aims to fill the void by precisely characterizing the relation of inductive support between hypothesis and evidence.

Analyzing the concept means analyzing what the confirmation relation is. That is, answering the question 'what is the confirmation relation'? In other words, we can say the attempt is to define the confirmation relation. Defining the confirmation relation means providing the defining characteristics of the confirmation relation. Providing the defining characteristics means providing certain characteristics or nature or mark of particular concept, which distinguish it from anything else.

2.1. Defining as distinguishing

Suppose we are defining a concept of justice regarding human actions then our definition must be able to provide certain characteristics of justice, which distinguish it from the action of cruelty, action of love, action of sympathy, action of help, action of dominance, action of hate, action of indifference, etc. That is, based on the definition of justice, we must be able to distinguish just actions from actions which are not just. Finally, we must able to distinguish an action of justice from the action of injustice.

In another sense, through definition, what we are providing is a certain kind of characteristics, which enable us to distinguish the cases of a concept from the non-cases of the concept. In a similar fashion, through the definition of confirmation, we must be able to distinguish the relation of confirmation from the other kinds of relation like deductive logical relation, explanatory relation, ignorance relation (non-committal relation), causal relation, irrelevance relation, emotive relation to hypothesis or evidence, part and whole relation, relation aroused from confusion/misunderstanding/ ambiguity.

That is having a definition of confirmation is having a certain set of characteristics which entail the distinction of confirmation from other relational concept which exists between propositions. That is, a definition of confirmation entails the distinction of the concept from other kinds of concepts, which are non- confirmational. Definition of confirmation necessarily involves elucidating the distinction between the relations which are confirmatory, and the relations, which are not confirmational.

Not only that the concept must be distinguished from other concepts, but also it must be able to provide an account of inner distinction of the concept, if there is any inner distinction of the concept. So far, we have used the notion of confirmation as an inductive justificatory relation which exists between observation and hypothesis. But a justification to a hypothesis can be positive and negative. That is, there are two kinds of confirmation relations that may obtain between a hypothesis and a piece of evidence, which are known as confirmation and disconfirmation. Confirmation can be considered as a positive confirmation relation that supports the truth of the hypothesis. While disconfirmation is a negative confirmation relation since it supports the falsity of the hypothesis. Our definition of confirmation must be able to distinguish not only the confirmation relation from other concepts but also it must be able to account for the inner distinction of confirmation. Through the definition of confirmation, we are providing defining/ distinguishing characteristics of the concept of confirmation, which account for the exterior and inner distinction of the concept. Characteristics of concepts are often called in our debates as conditions of confirmation, since the notion 'characteristics' is often associated with notion of property which are metaphysical in nature while the term 'condition' or 'criterion' are epistemological notions.

The second task of any definition of confirmation is to show that the distinction, which we have, between confirmation and non-confirmation is not an arbitrary distinction. That is, we must be able to show that distinction between confirmation and non-confirmation is based on a general principle, and not on any arbitrary positions. That is, the second task of the definition is to formulate a general principle, which determines certain relation between hypothesis and evidence as confirmatory and disconfirmatory. That is formulating a general principle, which governs our judgments of confirmation and disconfirmation.

However, the purpose of definition is solely limited to the task of distinguishing the concept from other concepts. I think definition of a concept must enable us to relate/compare the concept with other concepts. That is, through the definition we must be able to form a set of related concepts. A particular sufficient condition can distinguish the concept from other concepts in certain cases. Only a comprehensive sufficient condition (a condition that is both necessary and sufficient) can distinguish the concepts or its applications from other concepts in all cases. Therefore, it could be argued that a comprehensive sufficient condition provides an adequate definition of confirmation.

Comprehensive sufficient condition is either itself a necessary condition or it consists of all necessary conditions. For example, in the definition of knowledge, a comprehensive sufficient condition of knowledge is a combination of all three necessary conditions of knowledge: truth, belief and justification. [1]

However, the question is whether 'strict necessary conditions' (conditions which are only necessary, not sufficient or comprehensively sufficient) are inevitable or necessary constituents of a definition of concept. I hold the position that the definition which does not include 'strict necessary conditions' are inadequate even though the definition consists of comprehensive sufficient conditions. Though a comprehensive sufficient condition can help draw distinction among concepts, they cannot form a set of related concepts. A related set of concepts can be formed only through application of 'strict necessary conditions'. For example in the case of confirmation, there are various kinds of inner distinction like confirmation/disconfirmation and absolute/relevant confirmation (which will be discussed later). All these inner distinctions form a related set of concepts. In addition, concepts like deductive logical relation, explanation and confirmation can be put together to form a related set of concepts. We can see that some of the strict necessary conditions like equivalence condition are satisfied by all these related concepts. Specification of strict necessary conditions at least would be desirable nature/characteristics of a definition of confirmation since they help us to relate a concept with other concepts.

Our task is to define the confirmation relation thus to formulate a principle of confirmation which Hempel called 'logic of confirmation'. Theories of confirmation aim to formulate certain conditions and rules on the basis of which one can determine whether the

[1] *S* knows that *p* iff (1) *p* is true, (2) *S* believes that *p*, (3) *S* is justified in believing that *p*

given observation is confirming or disconfirming evidence to a hypothesis. Theories of confirmation characterize the relation through formulating certain rules for distinguishing the given set of evidence as relevant evidence and irrelevant evidence. And it also aims to distinguish the relevant evidence as confirming and disconfirming evidence. In other sense, the explicit definitions of concept of confirmation and disconfirmation and the explication of the basic principles behind the definitions constitute the project of confirmation.

But there can be three kinds of definition: qualitative, comparative and quantitative definition of confirmation. Hempel defines them as follows:

> (A) To give precise definitions of the two non-quantitative relational concepts of confirmation and of disconfirmation; i.e. to define the meaning of the phrases " E confirms H" and " E disconfirms H ". (When E neither confirms nor disconfirms H, we shall say that E is neutral, or irrelevant, with respect to H.)
> (B) (1) To lay down criteria defining a metrical concept “degree of confirmation of H with respect to E ", whose values are real numbers; or, failing this, (2) To lay down criteria defining two relational concepts, “more highly confirmed than " and " equally well confirmed with ", which make possible a non-metrical comparison of hypotheses (each with a body of evidence assigned to it) with respect to the extent of their confirmation. (Hempel, “Studies in the logic of confirmation”, 13-14)

And focus of this study is the qualitative definition of confirmation. Our study focuses on the basis, which determines a particular proposition or data as evidence of hypothesis. The conditions/ principle, which determine(s) the confirmation relation. In other words, defining the notion of confirmation or answering the question what is a confirmation relation. As we have pointed out, the question of confirmation is an attempt to characterize the inductive relation between hypothesis and evidence. The central task for any theory of confirmation is the formulation of principle(s) which account for the notion of inductive support and characterize this non-demonstrative relation in a precise manner.Inductive relation and its characterization and its foundation are intensely debated topics in philosophy in various aspects. To formulate our question of 'Definition of confirmation' in its preciseness, at least we need to distinguish the definition task from the major debates of induction.

3. Distinction from Hume's Problem of Induction

It is commonly held that any philosophical debate on induction meets its ends, its dead lock with Hume's challenge to induction: commonly known as 'Hume's problem of induction'. Hume's argument is that the principle of induction cannot be justified. There is no rational basis to hold induction as a valid method of inference. His specific point is that any attempt to justify induction itself pre-supposes the principle of induction. Thus justificatory argument becomes a circular argument. Hume characterises the principle of induction as 'uniformity nature', i.e. future events would be similar to past events. Hume shows that any attempt to justify the principle of 'uniformity of nature' itself assumes the principle itself. One could easily challenge Hume on the point that principle of 'uniformity of nature' is an inadequate characterization of the principle of induction. Hume's challenge transcends time since any adequate characterization of the principle of induction cannot be justified without assuming the validity of induction. And one could seriously doubt that how the project of definition of confirmation would have significance in view of Hume's problem of induction. Therefore, discussion of confirmation needs to be distinguished from discussion of problem of induction.

Salmon clarifies the distinction as 'inductive logic proper' and methodology of induction. "Previous failure to make the distinction led to all sort of mischief. There was serious doubt about the very possibility of an exact inductive logic, due in part to a confusion of inductive logic, due in part to confusion of inductive logic with methodology"

(Salmon, Carnap's Inductive Logic, 727). Here Salmon comments about Carnap’s theory of confirmation, inductive logic but his comments can be understood in general about a theory or definition of confirmation.

Salmon’s point is that Hume’s problem of induction is only a methodological problem of confirmation/induction and it is not a challenge to confirmation proper’ or definition of confirmation. Hume’s challenge is to point out how a principle of induction/confirmation which is provided through a theory/definition of confirmation can be justified. Critics are in right in stating yet we do not have a way or argument to justify the principle of confirmation as a valid principle of inference. However, the challenge of justifying a method or its principle is distinct from the challenge of characterizing or defining the method.

4. Confirmation/Induction

Our project is earmarked as definition of confirmation, rather than definition of induction. Method of induction or inductive inference is an ambiguous expression or a broader category, which consists of various kinds of construal of inductive relation while confirmation is a very specific kind of characterization of inductive relation. Even Hume himself formulated the challenge by analyzing how we proceed from certain observation to hypothesis or inductive conclusion. He was analyzing the principle of induction by analyzing our ways of formulating our inductive conclusion. Though his focus of analysis was inductive relation per se, he presented the inductive relation as a procedure or a means to formulate a theory or hypothesis on the basis of observation.

Logical Positivists/Empiricists clarify that logic/principle of inductive relation is distinct from the way or procedure of inductive inference. Thus they formulated the famous distinction between the Context of Discovery and the Context of justification. In the context of discovery what we are concerned with is how we formulate a hypothesis from certain observations by using an inductive relation or its principle. It seeks to characterize rules or procedures to be satisfied to derive a warranted hypothesis or generalization on the basis of a certain set of data. Its enquiries are directed towards the valid procedure for the formulation of hypothesis. However, the context of justification is surely concerned with how we justify a hypothesis on the basis of observation employing an inductive relation or its principle, irrespective of the way it is formulated. The context of the discovery of hypothesis is substantially different from the context of the justification of hypothesis. With this distinction, testing the hypothesis against a certain set of observation became the sole concern of the justification project. And the project of confirmation is distinct from the inductive inference on the basis of this distinction. While the former is the project of context of justification, the latter does not mention such distinction. Hempel sums up the discussion as follows:

> Another issue customarily connected with study of scientific method is the quest for “rules of induction”. Generally speaking, such rules would enable us to infer, from a given set of data, that hypothesis or generalization which accounts best for all particular data in the given set. But this construal of problem involves a misconception: while the process of invention by which scientific discoveries are made is as a rule psychologically guided and stimulated by antecedent knowledge of specific facts, its results are not logically determined by them; the way in which scientific hypothesis or theories are discovered cannot be mirrored in a set of general rules of inductive inference. (12-13)

As Hempel pointed, it is difficult to see that validity (warranted) is based on the way it is formulated. Only because one uses a crazy way or his/her fascination to formulate a hypothesis it is illegitimate to count the hypothesis as unwarranted. The legitimacy of hypothesis depends upon how it is related to certain evidence, not how it is formed derived from the evidence. Confirmation relation is an inductive relation characterized in the context of justification. In another sense confirmation relation is an inductive relation characterized as justification relation between evidence and hypothesis. The question of confirmation is that on what basis we can say that a hypothesis is justified by observation. What is the principle, which justifies a hypothesis on basis of observation?

5. Clarification of the term 'Hypothesis'

Prior to the discussion of confirmation/justification of hypothesis, the term 'hypothesis' has to be clarified to a certain extent. The primary distinction can be made of hypothesis is the distinction of 'universal hypothesis' and 'existential hypothesis'. Universal hypothesis holds that a relation between two entities/ objects is a universal one. For example, All ravens are black, Light consist of particles, There is a number which is greater than any other number. On the other hand, existential hypothesis affirms the existence of an object, the existence of a property in a relation with an object, the existence a relation with an object, or the existence a relation among certain objects. For example, life exists in Mars, some ravens are black, five is greater than two. Our discussion of confirmation of hypothesis is not merely about the confirmation of universal hypothesis but also about the confirmation of existential hypothesis. It should be noted that existential hypothesis does not merely consist of statement 'some ravens are black' but also consists of any singular proposition like the 'next observed raven would be (is) black'. By the word 'hypothesis' we mean any statement whose truth or falsity is not ascertained.

Nevertheless, there are two important kinds of statements: 1. Formal statement and 2. Empirical statement. Formal statements are those statements whose truth or falsity do not depend upon experience, instead it follows from definitions or assumptions, which we hold. (For example, A raven is a raven, 2+3=5.) But the truth or falsity of empirical propositions solely can be established by experience or experimental findings. Formal statements like statements of mathematics and logic are only concerned about the formal relations between two entities, while the empirical statements are concerned about a relation, which can be ascertained by experience or on the basis of experience.

In language, the terms are introduced into statements or employed in statements either by definitions or by minimal characterization of the features the terms refers to. And these definitions some time can be understood only in relation with certain axiomatic system or in relation with stock of knowledge which is not even necessarily considered as a finite set of statements. Only definitions which are using certain rules and which are related to the finite set of propositions or assumptions or axioms are considered as explicit/ rigorous definitions.

Obviously the term in our common language/statements are not employed by such rigorous definitions but they are definitional in nature since they satisfy one of the minimal conditions of definition that is distinguishing one term from another. In that sense, the terms like raven, atoms, protons are definitional in nature, though may not have rigorously defined as terms like greater or less than relations, truth-functional connectives. Certain definitions itself provide the possibility of various relations between two terms. In a certain sense, we can say that the richness of an axiomatic system can be understood in terms of the vast number of possibility of relations between terms solely provided by the definitions and axioms. For example, an axiomatic system, which consists only of the definition of equality, is less rich than the system, which consists of the definition of both equality and various inequalities.

The relation between two terms, which is on the basis of the definition of two terms, is called a formal relation. Certainly, the relation itself can be another definition. Since the definition of relation is parasitic or the definition of terms. A relation, which follows from definition of terms and by other definitions, can be called formal relations. Since it follows from the definition, truth or falsity of formal relation is determined by the definitions and assumptions of system itself. Statements about such formal relations or formal content are called formal statements. On the other hand, empirical statements are about the relation, which does not follow the definitions. In other words, the truth or falsity of empirical relation/ content cannot be established within a defined system. For example truth 'all ravens are black' cannot be established within a system like language which defines the terms raven and black. Truth of such statements needs a support from outside the system, which we call as 'reality'. The way we grasp/ assimilate the support for a statement from outside the system is called as experience. Experience is cognitive activity of agent, which relates the reality and a particular system. The point is that truth/falsity of an empirical statement is based on our experience. Experiential findings establish/support the truth/falsity of empirical statements but for the formal statements, it is not the case. In other sense, the distinguishing mark between the formal statements and empirical statements is the way truth/falsity of statements is established. That is why Hempel's remarks in the opening lines of his essay "the defining characteristic of an empirical statement is its capability of being tested by a confrontation with experiential findings" (10).

However, while the truth of certain empirical statements can be conclusively established by experimental findings, for other kinds of empirical statements, no finite amount of experimental finding can establish truth/falsity of it. Such conclusive establishment of truth of empirical statement is called verification. Though only universal statements are unverifiable, certain existential statements are unverified in present circumstances.

That is, the relation between experiential findings and empirical statements is more comprehensive than the notion of verification. Since most empirical statements like scientific hypotheses cannot be verified, yet they are related to experimental findings to determine truth/falsity. And this comprehensive relation between empirical statements and experimental findings is called as confirmation. Though in most cases, experimental findings do not establish truth of statements, they confirm/disconfirm the truth of propositions.

So generally, we can say that support/ justification provided by the experiential findings to truth of hypothesis is the distinguishing mark of the empirical statements. The experiential findings, which provide test of empirical statements, are known as evidence. The empirical statements, which are being tested, is called hypothesis whose truth or falsity is not ascertained. In addition, confirmation is understood as a relation between hypothesis and evidence.

6. Explication: Methodological Principle

A relation or concept can be characterized in various ways. That is, one could come up with various kinds of definition of confirmation relation. Then the challenge is to choose or determine the adequate characterization or definition of confirmation. On what basis we can see that a particular definition or confirmation is an adequate one. So normative characterization or definition of concept always requires a clarification of the basis of the characterization. Analysis of method or concept always presupposes or necessarily consists of a ground or basis which validates a particular characterization or definition. Each attempt to define a concept consists of a method. To determine a particular definition as valid we need to validate our position (method) on some other basis. A methodological principle validate the our method.

The point is that a particular definition of confirmation needs to be validated on certain grounds. In other words, we need to specify the methodology of our definition task. And I am not going to say that in the task of definition of confirmation, all philosophers wholly adhere to a particular methodology. In their attempt to provide a characterization (definition of confirmation) philosophers adhere to certain methodology though they may differ on the question of methodology. Rudolf Carnap is one philosopher who explicitly comes up with a methodology what is called as explication.

> [E]xplication consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second. We call the given concept (or the term used for it) the explicandum, and the exact concept proposed to take the place of the first (or the term proposed for it) the explicatum. The explicandum may belong to everyday language or to a previous stage in the development of scientific language. The explicatum must be given by explicit rules for its use, for example, by a definition which incorporates it into a well-constructed system of scientific either logicomathematical or empirical concepts. (Carnap Logical foundations of probability, 3)

Following are the four criteria of explication:

1. “The explicatum should be similar to the explicandum.

2. The explicatum should be given an exact specification within a rule-governed system of scientific concepts.

3. The explicatum should be a fruitful concept, and in particular allow for the formulation of many universal statements.

4. The explicatum should be as simple as possible. (This condition Carnap makes subsidiary to the first three).” (Sarkar and Pfeifer, *The Philosophy of Science: An Encyclopedia*, 288*)*

7. Hempel’s Theory of Confirmation

The concept of confirmation is introduced into the philosophical debate by Hempel in his paper “Studies in the logic of confirmation”. Hempel’s principal contribution to the discussion of confirmation is his formulation of adequacy condition and his definition of confirmation is subservient to his adequacy conditions.

Hempel’s primary and much discussed contribution to the theory of confirmation is his formulation of condition of confirmation. Through the formulation of certain adequacy conditions he aims to characterize the notion of confirmation in a precise manner. Hempel attempts to clarify the nature of confirmation relation by deducing certain characteristics which are the consequences of the concept of confirmation. Rather than providing a replacement of imprecise concept his attempt was to deduce the consequence of the concept and characterize the consequence in a precise form. His key point of formultion of adequacy condition is the following: how the confirmation relation between H and E is related to another confirmation relation between H_1 and E where H and H_1 are logically related. In a certain sense he assumes a class of propositions (H_1-H_n) which are logically related and then tries to work out how a confirmation relation between H_1 and E is related to the confirmation relation between H_i and E. He attempts to characterize a confirmation relation by forming a set of propositions which are logically related. Let us call the set as a logical space. He understood confirmation relation as a relation which can impact on certain logical relations between various propositions. And he assumes that through sufficient characterization of logical relation one can sufficiently characterize the confirmation too.

Following are Hempel's adequacy conditions:

1. Entailment Condition

"Any sentence, which is entailed by an observation report, is confirmed by it." (Hempel, "Studies in the Logic of Confirmation", 35). That is, If E logically implies H, then E confirms H.

2. Consequence condition

"If an observation report confirms every one of a class of K of sentences, then it also confirms any sentence which is a logical consequence of K." (35)

Hempel works out two corollaries for the consequence condition: equivalence condition and special consequence condition.

2.1. Equivalence condition

"If an observation report confirms a hypothesis H, then it also confirms every hypothesis, which is logically equivalent with H." (35) i.e. If E confirms H and H $\equiv$ H_1 then E confirms H_1

2.2. Special consequence condition:

"If an observation report confirms a hypothesis H, then it also confirms every consequence of H." (35) i.e. If E confirms H and H $\vdash$ H_1 then E confirms H_1

3. Consistency Condition

"...Every logically consistent observation report is logically compatible with the class of all the- hypotheses which it confirms."(37)

It has following two corollaries:

3.1. "Unless an observation report is self-contradictory, it does not confirm any hypothesis with which it is not logically compatible."(37)

3.2. "Unless an observation report is self-contradictory, it does not confirm any hypotheses, which contradict each other."(37)

Hempel, the pioneer of the philosophical debates of the confirmation, is a strong advocate of the need of correspondence between certain elementary forms of validity (of deductive logic) and the principle of confirmation. Hempel holds that correspondence between certain forms of validity and principle of confirmation is a necessary though not a sufficient condition to be a valid principle. On that basis, he works out certain adequacy conditions of confirmation, which he holds as necessary.

In the discussion of equivalence condition, we can see that the requirement of explication for a minimal level of formalization is necessary to check the validity of principle of confirmation. Equivalence condition which is highly intuitive can be precisely characterised only in a formal framework. Hempel's consequence condition (especially the Equivalence condition) is a paradigm example of how a meticulous analysis can bring out certain valid features of deductive reasoning which is necessary for any kind of valid reasoning. This feature of valid reasoning is adopted into the confirmation project too. In the line of deductive logic, confirmation too is construed as a relation between sentences.

Adoption of the minimal deductive logical framework helps us to adopt minimal valid constraints of deductive logic.

Other than the intuitive support of equivalence condition, Hempel's insistence for equivalence condition might have stemmed from the idea that different properties of propositions like truth and confirmation satisfy certain minimal common valid constraints. One of the valid constraints in the deductive logic regarding the assignment of truth is the following: if a sentence is true then its logically equivalent sentence is also true. In a similar fashion, Hempel adopted equivalence condition as a constraint for 'confirmation' and it turns out to be a highly intuitive condition.

The special consequence condition too is suggested in a similar line. Another constraint in the deductive logic regarding the truth is that if H is true, then all its consequences are true. But the adoption of this indisputable constraint of deductive logic faces severe criticism in the context of confirmation. Another condition which Hempel adopts by following valid feature of deductive logic is consistency condition. Here too we can see that this condition is a result of an adoption of a deductive constraint. In deductive logic if two propositions can be true then they are consistent. In a similar line, consistency condition argues that if two propositions are confirmed (by the same evidence) then they are consistent. The primary condition which Hempel held is the Entailment condition. Any sentence which is entailed by an observation report is confirmed by it which is a direct application of deductive entailment (any sentence which is entailed by a true statement is also true). Certainly, his claim of equivalence condition as an adequacy condition of confirmation is not challenged by major theoretical opponents, though most of them discarded the consistency condition and the special consequence condition and considered the entailment condition as a trivial one.

7.1. Hempel's Definition of Confirmation

Hempel aims to develop a definition of confirmation, which satisfies all three conditions of confirmation. Hempel's theory of confirmation, at best, is only a precise characterization of instance of confirmation. Hempel considers his account of confirmation as providing only a sufficient condition and not an adequate account of confirmation. Hempel's idea of definition starts from the entailment condition. The notion of confirmation as characterised by the entailment condition (If E entails H then E confirms H) is a formally valid one. Hempel points out that the challenge is to formulate a definition of confirmation which will not only satisfy the formal adequacy conditions of confirmation but which is also materially adequate. By the phrase 'materially adequate' he means "provide a close approximation to the conceptions of confirmation which is implicit in scientific procedure and methodological discussion"(Hempel, "Studies in the Logic of Confirmation", 39)

Suppose one is able to provide all possible evidence of a hypothesis then undoubtedly we can say that hypothesis is confirmed if and only if the evidence entails the hypothesis. But having all possible evidence for a universal hypothesis is a logical impossibility. More than that, the supposition of possible evidence is simply inadequate in the case of confirmation. What makes a confirmation relation interesting is its ampliative nature. But Hempel attempts to define confirmation relation by creating an analogical situation. For that Hempel introduced the notion of development of hypothesis. Development of hypothesis (DoH) is a form of hypothesis if the available evidence is the only possible evidence of hypothesis. In other words "the development of H for C states what H would assert if there existed exclusively those objects which are elements of C….." (41) Hempel considers that at least for a language with a simple logical structure, hypothesis is confirmed if it holds true for a class of individuals which are mentioned in the observation report. Thus Hempel defines confirmation as follows:

"(1) An observation report B directly confirms a hypothesis H if B entails the development of H for the class of those objects which are mentioned in B" (41).

As per this definition, the hypothesis (x) (Rx ⊃ Bx) is confirmed by Ra . Ba, since (Ra.Ba) entails Ra ⊃ Ba, which is the DoH of (x) (Rx ⊃Bx) for the class of objects (a & b). But the hypothesis that the next observed Raven would be black (Rb.Bb) is not confirmed by (Ra.Ba) since the development of the hypothesis, i.e, 'Rb.Bb' itself [2]is not entailed by (Ra.Ba). Thus Hempel introduces second corollary for his definition to meet the cases of such confirmation.

(2)"An observation report B confirms a hypothesis H if H is entailed by a class of sentences each of which is directly confirmed by B" (41)

As per this definition, Ra .Ba confirms Rb . Bb by virtue of confirmations of (x) (Rx ⊃Bx). Hypothesis Rb.Bb is confirmed by Ra.Ba because Ra.Ba directly confirms (x) (Rx ⊃Bx) and (x) (Rx ⊃Bx) entails Rb.Bb.

Hempel further shows that how his definition of confirmation satisfies all adequacy conditions. It is easy to show that his definition satisfies the equivalence condition in the case of simple hypothesis like (x) (Rx ⊃Bx). As per the Equivalence Condition (E.C) and Nicod's Criterion (N.C) all the four following propositions, Ra . Ba, ∼Ra . ∼Ba, ∼ Ra and Ba

[2] According to Hempel, "…the development of a hypothesis which contains no quantifiers… is defined as hypothesis itself, no matter what the reference class of individual is" (41)

are confirming evidence for the hypothesis. It is easy to see that all four evidence entail Ra ⊃Ba which is the DoH of the hypothesis.

Often Hempel's article makes it clear that he is quite aware of the inadequacy of the definition to capture the confirmation of hypothesis by observation which is formulated in a complex language. Hempel chooses to confine his definition to simple cases of confirmation because he prefers to have a definition which satisfies the adequacy conditions than the one which are materially adequate but violates the conditions. In the last section of this chapter, we will provide a detailed analysis of pros and corns of theories of confirmations. Hempel's theory satisfactorily explains the confirmation of empirical hypothesis, but it cannot even address the confirmation of theoretical hypothesis. "Thus, Hempel's account is silent on how statements drawn from such sciences as theoretical physics—for example, all protons contain three quarks—can be confirmed by evidence gained by observation and experiment." (Earman and Salmon, "The Confirmation of Scientific Hypotheses", 52). In the case of theoretical hypothesis there cannot be any DoH which is entailed by an observation report.

Hempel discusses the plausibility of the Converse Consequence Condition (C.C.C.) since it provides greater scope to any confirmation theory. Following is the converse consequence condition: 'If E confirms H then E confirms any propositions which logically entail H'. However, Hempel rejects the condition as invalid.

> The adoption of the new condition (converse consequence condition)… would have the consequence that any observation report B would confirm any hypothesis H whatsoever. Thus, e.g; If B is the report "a is raven" and H is Hookes law, then according to (8.1) (entailment condition), B confirms the sentence a is raven and hence B would, according to the converse consequence condition, confirm the

> stronger sentence "a is raven and Hooke's law holds." (Hempel, "Studies in the Logic of Confirmation", 36)

It is important to see that Hempel rejects the C.C.C clearly knowing the advantage of the C.C.C.:

> "but is it not true after all that very often observational data which confirms a hypothesis H are also considered as confirming a stronger hypothesis? Is it not true, for example, that those experimental findings which confirms Galileo's law or Kepler's laws are considered also as confirming Newton's law of gravitation? This is indeed the case, but does not justify the acceptance of the converse consequence condition as a general rule of the logic of confirmation." (36)

8. H-D Model

The Hypothetico-deductive model (H-D Model) is introduced primarily to overcome the lacuna of Hempel's theory of confirmation which fails to address the confirmation of theoretical hypothesis. According to the H-D model, the deductive consequence of a hypothesis is the confirming evidence of the hypothesis. Deductive consequences are largely considered as parts of the content of propositions. H-D model defines confirmation as follows:

1. If $H \vdash E$ and E is true then E confirms H.

2. If $H \vdash E$ and if E is not true (that is $\sim E$ is true) then E disconfirms H.

3. If $H \nvdash E$ and if E is true then E is neutral to H.

Unlike Hempel's theory of confirmation, the H-D model of confirmation is able to capture the confirmation of theoretical hypothesis. The H-D model construes observation findings or experimental results as deductive consequences of the hypothesis, auxiliary hypothesis and initial conditions. Earman and Salmon provides the following example to show that evidence of hypothesis are the deductive consequences of the hypothesis.

Premise 1 (Hypothesis): At constant temperature, the pressure of a gas is inversely proportional to its volume (Boyle's law).

Premise 2 (Initial Condition): The initial volume of the gas is 1 cubic ft.

Premise 3 (Initial Condition): The initial pressure is 1 atm.

Premise 4 (Initial Condition): The pressure is increased to 2 atm.

Premise 4 (Initial Condition): The temperature remains constant.

Conclusion: (Evidence): The volume decreases to *1/2* cubic ft.

In the similar way, the description of the phenomena of reflection, diffraction, rectilinear motion (unlike sound) can be considered as evidence of particle theory of light since the theory entails the evidence along with certain auxiliary hypothesis. Earman and Salmon gives the following example of disconfirmation as defined by the H-D model:

Premise 1 (Hypothesis): H: Light consists of corpuscles that travel in straight lines.[3]

Premise 2 (Initial Condition): I: A circular object is brightly illuminated.

Conclusion: (Evidence): The object casts a uniform circular shadow.

[3] Except when they pass from one medium (e.g., air) to another medium (e.g., glass or water).

“Surprisingly, when the experiment was performed, it turned out that the shadow had a bright spot in its center. Thus, the result of the test was negative; the observational prediction was false”(Earman and Salmon, “The Confirmation of Scientific Hypotheses”, 47).

Hempel’s theory of confirmation focuses only on qualitative statements. That is, a statement about a qualitative features of an object or the relation between different qualitative features of the object. Certainly Hempel’s theory is only a qualitative theory of confirmation. But a qualitative judgment of confirmation can be made about a relation between a hypothesis or evidence where either of them or both of them are quantitative. Often the evidence in sciences consists of quantitative features of object. That is, measurements of various properties like temperature, volume, pressure etc., are the relevant evidence for various relations between these properties. Unlike Hempel’s theory, H-D model captures the confirmation hypotheses which are quantitative statements. The above mentioned example of Boyles’s law exhibits the claim.

Capturing the confirmation by quantitative statements is quite important from philosophical aspect because, the challenge of under determination is largely constrained by the hypotheses or evidence which are quantitative statements. Qualitative evidence about optical phenomena continues to confirm both particle theory and wave theory of light till the formulation of evidence of optical phenomena in quantitative terms. Though particle theory and wave theory of light were equally good in explaining the most of the evidence about qualitative features of optical phenomena, particle theory of light fails to account for the evidence about quantitative nature of various optical phenomena. That is, by capturing the confirmation of theoretical hypothesis and by capturing the confirmation by quantitative evidence, the H-D model made significant steps in defining the confirmation relation. Though the H-D model succeeded in capturing the confirmation by quantitative evidence, it failed to capture the confirmation of a certain class of quantitative hypothesis, which is largely known as statistical hypothesis. This is because in the case of statistical hypothesis, evidence are not entailed by the hypothesis.

But the difficulties related to confirmation of statistical hypothesis do not seriously challenge the H-D model since confirmation theories' primary focus have been on confirmation of qualitative statements (universal or existential statements). But the possibility of non-deductive evidence seriously undermines even the H-D model. And it shows that the H-D model cannot be a necessary condition of confirmation. None of the criticisms, however, shows that true observational consequences are not confirming evidence of the hypothesis that is the H-D model cannot be a sufficient condition. But the paradox of Irrelevant Conjunction (I.C) challenges even the modest claim of the H-D model as a sufficient condition.

It is interesting to note that the H-D model of definition of confirmation is closely related to Hempel's theory of confirmation. Both employ relation of entailment to capture the notion of confirmation. Unlike Hempel's theory, the H-D model succeeded in capturing the confirmation of at least some theoretical hypothesis. But it crumbled under the paradox of irrelevant conjunction which states that if E (H-D) confirms H then E also (H-D) confirms H.X, where X is any irrelevant hypothesis. Interestingly the pitfall of H-D model was foreseen by Hempel in his discussion of Converse Consequence Condition and it seems that Hempel narrowed the scope of his definition to avoid the pitfall.

9. Inference to the Best Explanation

In 1965, Gilbert Harman came up with a striking position, known as the Inference to the Best Explanation (IBE) which claims that our inferential practices are guided by explanatory considerations. However, till Peter Lipton's path breaking work, the idea of 'Inference to the Best Explanation' remained an intuitive idea without much precise characterization of the notion of 'explanation' and more importantly of 'Best explanation'. Lipton's was an ingenious attempt to define the notion 'best explanation' and argues that how the best explanation leads to inferential practices. The key idea of the IBE is that explanatory considerations are the guide to inferential practices. That is, we are inferring to what we consider as the best explanation of the phenomena or evidence.

The key claim is that the explanatory relation between hypothesis and evidence is prior to the confirmation relation. That is explanatory relation is a necessary prerequisite for the confirmatory relation. The general point of the IBE is that confirmation relation can be defined on the basis of an explanatory relation. That is, we would be able to determine whether a hypothesis is confirmed by evidence E by analyzing whether it provides the best explanation of E. We have seen that when a hypothesis provides an explanation of an observation report or a particular fact, the particular fact or observation report can be considered as a confirming evidence to the hypothesis.

One of the major claims of the IBE is that explanatory considerations are a guide to inference. That is, the scientists infer from a set of available evidence to a hypothesis on the basis of explanatory considerations. Lipton's claim of 'explanatory considerations are guide to inference' can be understood in various ways. One way is that explanatory considerations often help or guide to formulate a hypothesis. As we have discussed, the project of confirmation is not about a set of rules or procedures, which characterise the way in which hypothesis or theories are formulated. In that sense, confirmation is not about an inductive relation, which leads scientists from evidence to hypothesis. Lipton's claim too is not that explanatory considerations works as guide to formulate a hypothesis. Instead his claim is that explanatory considerations are key factors in determining whether given evidence is a confirming one or not. So how can we understand Lipton's claim that explanatory considerations are a guide to inference in the context of justification or in the context of testing of hypothesis.

IBE's context of justification or context of testing is crucially different from the context of justification which other theories have. In the case of the H-D model, and Hempel's model, we are determining the confirming nature of evidence solely by analyzing the relation between evidence and particular hypothesis. If a particular relation holds then it is confirming. Otherwise it is not. In IBE, the confirming nature of evidence is not decided by analyzing the relation between a particular hypothesis and its evidence. Confirmation relation between hypothesis and evidence can be analyzed only by bringing the competing hypotheses into the picture. The question of confirmation in the IBE is not that whether the hypothesis provides an explanation to evidence. Instead the question is that whether the hypothesis provides the best explanation to the evidence, the best explanation among the explanations provided by various competing hypotheses. When Lipton says that explanatory considerations are guide to inference he means explanatory considerations are guide to determine the hypothesis which provides the best explanation from the set of competing hypotheses. In a sense, explanatory considerations are guide to choose the best explanations from various kinds of explanations. From evidence, by using explanatory considerations we infer the hypothesis which provides the best explanation. So, in IBE the confirmation relation is not explicated by analyzing the relation between a particular hypothesis and evidence, but by analyzing a relation between evidence and a set of hypotheses which are competing.

In IBE's account, competing hypotheses play a significant role in determining whether evidence confirms a hypothesis or not. So the important question is how the competing hypotheses are defined in IBE's framework. We might be tempted to think a set of competing hypotheses in IBE as the same as the set of competing hypotheses in practice. IBE defines the competing hypothesis in a very broad sense. All hypotheses which provides potential explanations of evidence are included in the set of competing hypothesis. In a sense, all possible explanations are potential explanations.

Now the challenge is to define the notion of the 'best explanation'. On what basis we can determine that a particular explanation is the best. How can we define the notion of the best explanation? One way of defining the best explanation is to choose the likeliest one from all the possible explanations. That is the best explanation is the explanation which is likely to be true. Such an interpretation defeats IBE, because the very purpose of IBE is to claim/show that the hypothesis which provides the best explanation is true or likely. And the claim is that likeliness or truth can be captured or determined by explanatory considerations.

So Lipton argues that the best explanation is not determined on the basis of likeliness but on the basis of loveliness. Loveliest explanation is an explanation which provides maximum understanding. Suppose we asked the question that 'why a particular raven is black' or 'why all observed ravens are black'. In a sense, the hypothesis 'all ravens are black' explain the evidence 'A is a black raven'. Since 'all observed ravens are black' it is likely that 'all ravens are black'. But this is an explanation which is likely to be true, but it is not a lovely explanation. It is because it does not provide any kind of understanding to the question 'why it is black'. Suppose we provide an explanation by citing climatic conditions or genetic structure. It provides us some kind of understanding about the phenomenon or fact. We have defined Inference to the Best Explanation (IBE) as Inference to the Loveliest Potential Explanation (ILPE) but the basic question remains unanswered 'what is an explanation' or 'what is the loveliest explanation'.

9.1. Causal Model

For Lipton, the adequate model of explanation is the causal model. According to the causal model of explanation, to explain a phenomenon means to give information about its causes. For example: why there is fire because there is electrical short circuit. Why there is long stride because a tall man walked through here. Why there is a change in the orbit than we predicted? It is because of the gravitational force of another planet. Why the light from other galaxies are shifted to red spectrum because other galaxies are moving away from the Milky Way. Why there is bright spot in the shadow of circular object because of the wave-nature of light. Why the volume of gas is decreased because the pressure of gas increased. In all such cases we explain a phenomenon by citing its causes. Giving information about the cause of phenomena is the explanation which we provide in all these cases. Often it is not about a citing a single cause of the phenomena. According to the causal model, explaining means giving information about its causal history. That is, giving information about the mechanism which is linking cause and effect and the phenomenon which is being explained is construed as the effect. The main objection to the causal model is that it is too weak or too permissive. That is providing the cause of the phenomena often does not provide a good explanation because often the causal histories are long and wide. Most information about the causal history of a phenomenon is explanatorily irrelevant. So explaining cannot be simply giving all information regarding causal history.

That is, though some causes explain the phenomena, most of the causes are explanatorily irrelevant. So the challenge is to select the causes which are explanatorily relevant or to account for the causal selectivity of our practice. Lipton thinks that by exploring the structure of why question and how why question itself focuses on certain aspects of phenomena, we can have an account of causal selectivity of our explanatory practices. Lipton's point is that what gets explained is not simply a phenomenon ('why this') but a contrastive phenomenon ('why this rather than that'). That is, our why question implicitly or explicitly consists of the contrastive nature. Our phenomena may be 'Leaves turn yellow'. But when we are asking the question it can be specified in many various way. 'Why leaves turn yellow rather than blue'. Why leaves turn yellow in November rather than January. Why leaves turn yellow rather than not being changed. Lipton says that

> When I asked my, then, 3-year old son why he threw his food on the floor, he told me that he was full. This may explain why he threw it on the floor rather than eating it, but I wanted to know why he threw it rather than leaving it on his plate. An explanation of why I went to see Jumpers rather than Candide will probably not explain why I went to see Jumpers rather than staying at home, an explanation of why Able rather than Baker got the philosophy job may not explain why Able rather than Charles got the job, and an explanation of why the mercury in a thermometer rose rather than fell may not explain why it rose rather than breaking the glass.......... A fact is often not specific enough: we also need to specify a foil. Since the causes that explain a fact relative to one foil will not generally explain it relative to another, the contrastive question provides a further restriction on explanatory causes. (Lipton, *Inference to the Best Explanation*, 33-4)

Lipton's point is that what we explain is a contrastive phenomenon. He analyzes the nature of contrastive phenomenon: A contrastive phenomenon consists of facts and foil. 'Why the fact rather then the foil? And Lipton points out that in such contrastive why question, we assume that fact occurred but foil did not. Once we explicated the nature of contrastive question, now the challenge is to explain the causal selectivity on the basis of the nature of contrastive question.

For that Lipton formulates a difference condition on the basis of Mill's method of difference. While Mill's method of difference concerns about the discovery of causes, Lipton's difference condition concerns about finding an explanatory cause from the given causal history. Method of difference and Method of agreement are two central methods for Mill's account of causal inference. " According to the Method of Difference, when we find that there is only one prior difference between a situation where the effect occurs and an otherwise similar situation where it does not, we infer that the antecedent that is only present in the case of the effect is a cause." (18)

In the similar lines of Mill's method, Lipton formulates difference condition to account for causal selectivity. And Lipton points out that to explain why P rather than Q, we must cite the difference between the causal histories of P and Q. That is to explain why P rather than Q, we must cite the cause of P and the absence of a corresponding event of the cited cause of P in the causal history of not Q. Lipton's famous example is the following: Why paresis is contracted to Jones and not to Smith. In this case, the syphilis which is present in causal history of Jones and which is absent in the causal history of Smith, is the explanatory cause. He defines the explanatory cause as follows: "Explanatory cause is the one which is present in causal history of P and its corresponding event is absent in causal history of ~Q. In this case the corresponding event is Smith's syphilis.

9.2. Definition of Confirmation and Disconfirmation

Lipton's point is that by using the difference condition we can determine the cause, which is explanatorily relevant. Moreover, the hypothesis, which provides explanatorily relevant cause, is the best explanation of the contrastive phenomenon. In another sense, the hypothesis, which provides the explanatory cause, is confirmed by the contrastive phenomena. According to IBE, confirmation relation is determined not only by the relation between a particular hypothesis and evidence but also by the relations between the evidence and the individual members of set of competing hypotheses. We defined competing hypothesis as hypothesis which provides potential explanations of the evidence or fact. In that sense, all hypotheses which provides potential explanations are relevant hypothesis. So evidence is neutral to all other hypothesis which are not part of the set of potential explanations. According to IBE, Evidence E neutrally confirms (neutral to) H if it does not provide not a potential explanation of evidence. The hypothesis/theory which provides potential explanation of evidence are either confirmed or disconfirmed by phenomenon or contrastive phenomenon. The hypothesis, which provides explanatory cause of the contrastive phenomenon, is confirmed by the contrastive phenomena. Subsequently all other hypotheses which are incompatible/inconsistent with the confirmed hypothesis are disconfirmed by the contrastive phenomena. And in his work, 'Inference to the Best Explanation' he attempts to sharpen this definition in various ways.

9.4. Critical Analysis of IBE

Undoubtedly, Lipton's difference condition is an adequate criterion to determine the best explanation from a set of competing hypotheses/potential explanations if the competing hypotheses adhere to the same causal history of the phenomenon but differ only in determining the explanatory cause of the agreed causal history. As his example suggested, in many of the cases of confirmation/ explanation, there may not have any dispute/disagreement on the causal history but only on explanatory cause. However, in most of the important cases of confirmation, the various potential explanations provide various causal conditions.

Consider the contrastive phenomenon 'why there is advancement of perihelion of Mercury rather than the Earth'. Newton explains the phenomenon by providing a cause of gravitational force of unknown planet which he calls Vulcan. In Newton's explanation, this cause was present in the causal history of advancement of perihelion of Mercury and absent in the causal history of non- advancement of perihelion of Earth. In that sense, Newton's theory satisfies the difference condition. But in Einstein's explanation, he was providing an entirely different set of causal conditions or different causal history. Similarly in explaining the phenomenon of 'Combustion' Lavoisier explains it by citing the presence of Oxygen and absence of Oxygen in foil. But Priestley explains it by providing the presence of Phlogiston in the fact and absence of it in the foil.

So to select the best explanation from a vast set of potential explanations, first one needs to determine which causal history is the actual one. We can say that there are two kinds of potential explanations. In the first set of potential explanations, all hypotheses hold the same causal history but differ only in choosing the explanatory causal conditions. To find out the best explanation from the first set of potential explanations, difference condition is an adequate criterion. However, the second kind of potential explanations differ not merely on the explanatory causal conditions instead, they differ on the causal history of phenomenon itself. Then the question is how can we choose the actual/loveliest/likely causal history from the various causal histories provided by various competing hypotheses. Lipton somewhat succeeded in choosing a loveliest explanation from the various potential explanations which talk about the same causal history. But he left unaddressed the issue of choosing the best explanation from various hypotheses which provide different causal histories.

10. Probabilistic Approach to Confirmation

Undoubtedly explanation relation captures a strong intuition of confirmation. At least in an intuitive sense or ordinary language sense, explanation can work as a basis to explain/understand the confirmation relation. In ordinary language discourse, notion of explanation is still a powerful tool to comprehend the nature of confirmation. Preciseness lacks in what is considered as the most-comprehensive theory of confirmation. But the precise characterization of explanatory account of confirmation, i.e. IBE, fails in many fronts. The key idea of the IBE is that the hypothesis which provides the best explanation of phenomena is considered as a confirmed hypothesis. The IBE account defines best explanation using the difference condition. But IBE's definition fails to comprehend the notion of best explanation in comprehensiveness. I think certain other account of explanation like 'unification model' might be more comprehensive than the best explanation as defined by difference condition. But the notion of 'unification' is not a precisely formulated concept. Deductive logic accounts of confirmation (H-D model and

Hempel's Model) are precisely formulated account but they lose the comprehensive nature of confirmation.

The task of defining confirmation faces the challenge of formulating a principle which is precisely defined and comprehensive. Here we can say that often we consider the precise characterization of the principle as of primary importance, since we often fail to determine whether a principle/account is comprehensive if it is an imprecisely formulated one. The probabilistic approach to confirmation gains significance in this context.

According to the probabilistic approach, confirmation can be defined on the basis of probability relation. Generally, we can say that probability is a relation which relates certain events or sets to real valued numbers. Formal definition considers this relation as a function. In probability calculus, primarily we ascribe real value numbers to certain proposition or events. Significant point is that the ascriptions of these real-valued numbers to events or propositions are rule-governed one. Ascription of values/ drawing the relation between domain (set of events or propositions) and codomain (set of real numbers) is necessarily followed by the rules of probability calculus which is known as 'Axioms and definition of probability'.

On the basis of the rules of probability, various kinds of relations are defined on probability space. And these rules of probability define various relations between proposition in the probability space in a precise manner. Rules specify how various relations like union (disjunction), intersection (conjunction) and complement (negation) can be defined in the probability space. Moreover, probability calculus introduces new relation (between sets or proposition) called as conditional relation. Probabilistic approach to confirmation is significant not merely because it is rule-governed but also because it defines a wide-range of relation between propositions. e.g. union, intersection and complement. Relational concepts like conditional relation and probabilistic independence and their precise definitions are extremely useful in explicating various intuitive relation between propositions. While rule-governed nature of probability ensures the preciseness (exactness) of the notion of probability, abundance of various relations within the rule-governed system adds to the comprehensive nature or richness of probability relation.

One way to characterize the notion of confirmation is the following: assume a confirmation relation between two propositions and introduce various relations which proposition H and E have with other propositions. And analyse how the confirmation of H by E affects/determine the confirmation of other related proposition by E. Hempel's conditions are clearly an attempt in this way: he introduces the relation of logical equivalence and consistency and assumes confirmation relation between H and E and determines the confirmation of related proposition by the same evidence E. Hempel succeeds in determining the confirmation of related propositions but he fails to determine how the relations itself are determined by the assumed confirmation relation between H and E. That is he could not determine how a relation (like conjunction, disjunction) itself is affected by the confirmation of a conjunct/disjunct. The H-D model was able to determine the confirmation of conjunction when one of its conjuncts is confirmed. We can say that the paradox of irrelevant conjunction is the price which the H-D model pays for defining the confirmation of conjunction on the basis of the confirmation of only one of its conjunct. In a similar fashion, the H-D model includes the relation of 'logical independence' also. According to the H-D model, if a proposition, H is confirmed by E then the proposition H_1 which is logically independent of H is neither confirmed nor disconfirmed by E.

For the H-D model, relevant evidence (both confirming and disconfirming) are logically dependent propositions. But the logical dependence is too narrow a relation to comprehend the confirmation relation. But probabilistic calculus introduces a more comprehensive relation called probabilistic dependence/independence.

10.1. Bayesian Confirmation Theory

The Bayesian confirmation theorist attempts to capture the notion of confirmation on the basis of probabilistic independence. If a proposition, E is probabilistically dependent of H, then E is relevant (confirming/disconfirming evidence) to the hypothesis. According to the Bayesian confirmation theory (BCT) probabilistic dependence is a key notion in defining the notion of confirmation as probabilistic independence is defined on the basis of conditional probability. P(E|F) means probability of E given that F is true. Probabilistic independence is defined as follows: If $P(E|F) = P(E)$ then E and F are independent. According to the BCT, relevant evidence and hypothesis are probabilistically dependent. If $P(H|E) \neq P(H)$ then E is a relevant evidence for H. On this basis, the BCT defines the confirming and disconfirming evidence.

Carnap's preliminary point is that the notion of confirmation is used in an ambiguous way in our ordinary language discourse. Different senses are attached to the term 'confirmed'. So any attempt to define the concept must discern the different senses attached to the term 'confirmation'. Carnap's point is that, in one sense, the notion of confirmation is employed in a strong sense. In the stronger sense, on the basis of evidence, a hypothesis is considered as an acceptable part of knowledge system. That is support of the evidence should be strong enough to consider H as a firmly grounded one. He terms this notion of confirmation as 'Absolute confirmation'. But he points out that in absolute sense of confirmation, evidence need not be even relevant to the hypothesis. Notion of evidence that we are concerned about in our account of inductive analysis is a relevant one. And also our attempt is to explore the minimalist account of justification. Otherwise a characterization of justification relation would not be a comprehensive one. Thus Carnap introduces the notion of confirmation which is called as relevant concept of confirmation which captures our intuitive notion of relevance relation between propositions and which is comprehensive enough to account for a minimal conception of justification.

Carnap defines relevant notion of confirmation as follows:

E confirms H iff $P(H|E) > P(H)$

E disconfirms H iff $P(H|E) < P(H)$

E neither confirms nor disconfirms H (i.e. neutral to H) iff $P(H|E) = P(H)$

The relevant notion of confirmation explicates the principle as the relation between two probabilities of a proposition called the 'posterior probability' and the 'prior probability' of hypothesis. That is, the mathematical relations of inequality and equality between the probabilities work as the basis of determining the nature of confirmation relation. According to the BCT, the equality/inequality is a sufficient and necessary condition of confirmation. Confirmation relation is construed as a mathematical relation between two different probabilities. In that sense, it constitutes a minimal level algorithm for the determination of confirmation relation given that an agent can determine both probabilities.

But it is interesting to note that posterior and prior probabilities of propositions are not independent probabilities. Value of posterior (prior) can be determined by using the other. If our construal is not extended to inter-dependence of posterior and prior probabilities, it sounds like that an agent is determining both probabilities independently and drawing a mathematical relation between them. But such construal is not a sufficient algorithm for an agent to determine the confirmation relation as it does not sufficiently determine the posterior on the basis of the prior probability. From the definition of conditional probability[4], it is clear that both probabilities (prior and posterior probabilities) are dependent. It does not sufficiently characterize how they are probabilistically dependent. The relation between probabilities is a mathematical relation. A sufficient characterization requires a condition which is necessary and sufficient to hold this mathematical relation.

Sufficient characterization between prior and posterior probability would reveal the nature of confirmation relation in a more precise sense. A rich characterization of the relation between prior and posterior probability is quite important since it helps one to draw a correspondence to confirmation relation which might be inexplicit in various forms. The concept of confirmation can be characterized in a formal language, but in practices of confirmation, the confirmation takes various forms, often the complex forms, depending upon the compulsions of the context. And often this characterization cannot be readily translated as a mathematical relation between the prior and the posterior probability. It is because often these practicing forms of confirmation may not even refer to the prior and the posterior probability of the hypothesis. For example, take instance confirmation or confirmation by deductive consequence, these forms are in certain sense are the practical forms too. These practical forms do not refer to the probability of hypothesis. In other words, characterisation which identifies only probabilities of hypothesis as the constituent

[4] $P(H|E) = \frac{P(H.E)}{P(E)}$

of confirmation relation would be limited in its scope. The same limitation would also occur in having a fruitful engagement with other theories of confirmation.

So what are the factors or constituents which determine equality/inequality relation between prior and posterior probabilities? In other words, what is the algorithm to determine equality/inequality relation between the prior and the posterior probability if only the prior probability is given. A characterization of a relation between the posterior and the prior probability which do not answer these questions would remain uninteresting as it does not explicate the nature of confirmation.

A surprising and exciting discovery for probabilistic confirmation is that a theorem of probability calculus precisely defines how prior probability of a hypothesis is related to its posterior probability. It specifies what are the conditions to be satisfied to hold greater than relation/less than relation between the prior and the posterior probabilities of hypothesis. It tells us what are the necessary constituents which play a role in the determination of equality or inequality relation between the prior and the posterior. Thus it contributes to the scope and fruitfulness of the characterization of confirmation. This theorem is known as the Bayes theorem. In simplest form it is as follows:

$$\mathrm{P(H|E)} = \frac{\mathrm{P(E|H)\times P(H)}}{P(E)}$$

A practically useful version of the theorem is as follows since it explicates the role of competing hypothesis.

$$\mathrm{P(H|E)} = \frac{\mathrm{P(E|H)\times P(H)}}{\mathrm{P(E|H)\times P(H)+\ P(E|{\sim}H)\times P({\sim}H)}}$$

According to Bayes' theorem, the posterior probability of a hypothesis is based on three factors.

1. P(H): Prior probability of the hypothesis
2. P(E|H) and P(E|∼H): Likelihoods of hypothesis
3. P(E): Probability of Evidence

The prior probability of hypothesis means probability of hypothesis prior to testing of H by E. Before the testing of hypothesis what is the probability of hypothesis. Likelihood of hypothesis is probability of evidence given the hypothesis is true. In other words, the probability a hypothesis assigns to the evidence. Likelihood is the value which represents how strongly a hypothesis explains or predicts the evidence.

One of the main short comings of the H-D model is that it overlooks the role of competing (alternative) hypothesis in the case of confirmation of hypothesis. In the H-D model, deductive consequences are the evidence of hypothesis but infinite number of competing hypotheses or alternative hypotheses could also entail the evidence. Thus we can say E confirms not only H but also most of its competing hypotheses. This contradicts with our intuitive notion of confirmation. The point is that in the case of confirmation of H by E alternative hypotheses' relation to E plays a crucial role and the Bayes' theorem sufficiently characterises how a set of alternative hypotheses figures in the determination of posterior probability of a hypothesis, which is the corner stone of confirmation.

Probabilistic approach to confirmation holds that confirmation relation is based on the mathematical relation between posterior and prior probability of hypothesis. The theory of probabilistic approach to confirmation which holds that Bayes theorem sufficiently defines the relation or provides necessary and sufficient condition for the relation between posterior and prior probability of hypothesis is known as the Bayesian confirmation theory (BCT). The BCT holds that Bayes' theorem provides an algorithm to determine the posterior probability of hypothesis.

The formal schema of confirmation characterized by the Bayes' theorem is based on three factors. The first one is prior probability of hypothesis. Prior probability of hypothesis means probability of hypothesis prior to testing of hypothesis. But we need to specify that what are the considerations which determine prior probability of hypothesis. In relation with a particular confirmation by E_n, we can argue that confirmation by prior evidence, $E1_{\dots}$ $E_{n-1,}$ (evidence which are discovered prior to the testing by E) constitute the prior probability of hypothesis.

But the question remains what constitutes the prior probability in the case of confirmation by E_1. It suggests that rather than prior confirmations, there are other factors too which constitute the prior probability of hypothesis. Salmon argues that plausibility judgments which scientists make about a hypothesis constitute the prior probability.

Salmon's point is that in judging the probability of hypothesis, scientists make use of various formal considerations like 'internal consistency of hypothesis' or material considerations (considerations which pertain to the content of hypothesis) like simplicity, fruitfulness and scope of the hypothesis. An account of probability of hypothesis or confirmation of hypothesis which excludes such plausibility considerations bound to be an inadequate one as strongly argued by Kuhn ("Objectivity, Value Judgment, and Theory Choice", 320–39). Bayes' theorem represents these considerations by introducing the factor 'prior probability of hypotheses'

10.1.1 Likelihood

P(E|H) is called as likelihood of hypothesis. How likely the evidence is if we know that the hypothesis is true. If H is assumed as true then what is the probability of E?. The value of P(E|H) is determined one in most of the cases.Likelihood of hypothesis can be easily determined on three cases

1. When H entails the evidence E, P(E|H)=1 for example: P(Ra.Ba|(x)(Rx ⊃ Bx))=1
2. When H and E are logically incompatible. P(E|H)=0 for example: E: Light travels through empty space and H: Wave theory of Light
3. When H is a statistical hypothesis.

In certain other cases, we may employ our past experience to determine likelihood.

But still there are many other cases where determination of likelihood is neither simple nor straight forward. We will address such concern in the interpretation of probability. Especially the subjective interpretation of probability, allows different judgment about likelihood which are equally rational one. The point is that an agent would be able to determine the likelihoods of hypothesis by having a judgment evidence-hypothesis relation.

10.1.2. Probability of Evidence

Third factor is the probability of evidence P(E). As we have seen a probability theorem straightaway defines the probability of evidence, on the basis of likelihoods and prior probability of evidence.

$$\mathrm{P(E)} = P(E|H) \times P(H) + P(E|{\sim}H) \times P({\sim}H)$$

But the Bayesians who hold subjective interpretation do not consider the recourse to the probability theorem is necessary in determining the probability of evidence.

11. Philosophical Interpretations of Probability

The key point of the BCT is that confirmation relation between hypothesis and evidence can be defined/determined on the basis of probability relation, especially on the basis of conditional probability relation. And Bayes' theorem provides an explicit formulation of the conditional probability which can be directly related to the practices of confirmation.

One significant question could be raised is that are the axioms and the rules of probability sufficient to determine the probability/conditional probability of a proposition? This question would be vivid if we draw a contrast with deductive logic. In deductive logic, if two propositions are assumed as logically dependent, then on the basis of truth-value of one proposition, we would be able to determine the truth-value of other proposition, since the proposition is either inconsistent with the first proposition or a deductive consequence of the first proposition. But in the case of probability, if two propositions are probabilistically dependent, probability of one proposition, does not determine the probability of the other proposition. In deductive logic, the truth value of one proposition given the truth value of the other proposition is determined on the basis of rules of logic. However, rules of probabilities do not determine the probability value of a proposition but only constrain the values. Following are the basic rules or axioms of probability.

1. $0 \leq \mathrm{P(H)} \geq 1$
2. $\mathrm{P(T)} = 1,$ where T is a tautology
3. $\mathrm{P(H.H_1)} = \mathrm{P(H)} + \mathrm{P(H_1)}$ if H and H_1 are mutually exclusive propositions.

In relation with the determination of probability value, the first axiom only says that probability value of any proposition is a real number between 0 and 1 including 0 and 1. It constrains the value which can be ascribed to a proposition, but does not determine. Certainly it is highly strong constraint than it appears, especially when we determine the probability value of a proposition on the basis of another proposition. The second axiom determines the probability value of a proposition but it is applicable only to tautology. And the third axiom determines the probability value of a conjunction given that the probability values of both conjuncts are known. Certainly it constrains the probability value of conjuncts also. For example, if two propositions are mutually exclusive and we assign probability value which is between 0 and 1 and if the probability value of the conjunction exceeds 1, certainly we can say that assignment of probability values to the conjunct(s) is incorrect or inadmissible. In a similar way, various definitions of probability (like definition of independence, and definition of condition probability) along with the axioms of probability provide strong constraints to the assignment of probability value but does not determine. In a sense, the axioms of probability do not determine/define the notion/concept of probability sufficiently, though they are the necessary conditions of probability. The remaining question is that how can we sufficiently define the concept of probability. Certain other rules or methods must be employed for sufficient determination of probability. But it seems that these complementary methods/conditions can be explicated only by analyzing or characterizing the concept of probability.

Insufficiency of probability axioms in defining the concept of probability opens up the debate of interpretation of probability which attempts to answer the question ‘what is probability' or which attempts to explore the meaning of probability. Following are the major five theories of probability, which attempt to define the concept of probability.

1. Classical interpretation of probability
2. Frequency interpretation of probability
3. Propensity interpretation of probability
4. Logical interpretation of probability
5. Subjective interpretation of probability

There are two necessary conditions which have to be satisfied by any adequate interpretation of probability.

1. Interpretation satisfies the axioms and derived rules of probability calculus.
2. It provides a method/rule to determine/ascertain probability value of a proposition or event.

Classical theory of probability is what we assume as the interpretation when we work on standard text book example of random experiments. Standard random experiments are tossing a coin, throwing a die, drawing a card from a deck of cards. In all such examples, we ascertain the value of probability of event by dividing the sample space into equally possible events or cases. Then we determine the value of an event by using the formula: $\frac{Number\ of\ favorable\ outcomes}{Number\ of\ all\ possible\ outcomes}$. But the formula would be directly applicable only in cases, where we can divide all possible outcomes into equally possible cases. The challenge would be in assigning the probability to unequal cases of outcome.

All interpretations of probability attempt to provide a principle/method to answer the challenge based on their characterization of probability. In this dissertation, I only discuss two interpretations of probability: logical and subjective interpretation since these two are the interpretations used in the discussion of confirmation. Now it is widely held that the concept of probability which is employed in characterizing the notion of confirmation is substantially different from the concept of probability used in physical sciences. Philosophers like Carnap and Frank Ramsey argue that there are two distinct concepts of probability and I focus only on two interpretations, since other interpretations of probabilities are not about the notion of probability in the context of confirmation. The point is that there are two distinct concepts of probability, therefore there must be two distinct kinds of definition/interpretation of probability. Patrick Maher argues the point of two distinct concepts as follows:

> "… suppose you have been told that a coin either has heads on both sides or else has tails on both sides and that it is about to be tossed. What is the probability that it will land heads? There are two natural answers: (i) 1/2; (ii) either 0 or 1 but I do not know which. These answers correspond to different meanings of the word "probability". The sense of the word "probability" in which (i) is the natural answer will here be called the inductive probability. The sense in which (ii) is the natural answer will be called physical probability. Physical probability depends on empirical facts in a way that inductive probability does not. We can see this from the preceding example; here the physical probability is unknown because it depends on the nature of the coin, which is unknown; by contrast, the inductive probability is known even though the nature of the coin is unknown, showing that the inductive probability does not depend on the nature of the coin."(Maher, 2006, "Confirmation Theory", 4)

When we say that it is highly probable that you will fall if you walk on the slippery surface, we talk about the probability of occurrence of an event. Given certain kind of causal conditions, we can talk about the probability of the effect. In such circumstances, we are talking about the probability of the occurrence of an event. Physical probability (which is often called as chance) is the feature of world/feature of event. Inductive probability often called epistemic probability is not the feature of the world; it is the feature of the agent's judgment about the occurrence of the event. Consider tossing of a coin which is biased towards the occurrence of the head. In that case the physical probability of getting head is very high. But consider the judgment of the epistemic agent who does not have the knowledge or biased nature of the coin, agent's judgment about the probability of getting head would not be high as the physical probability. When we say that astronomical data makes it very probable that our universe had a beginning, clearly what we employ is an epistemic probability. There is no point in discussing the physical probability of the event 'beginning of universe' since it is either zero or one. The distinction between the concepts is vivid in the case of all past events.

The striking difference between the two concepts would appear in the discussion of probabilistic independence. Standard text books example for probabilistic independence is the tossing of a coin for the first time and the tossing of the same coin a second time. Certainly the outcome of the first toss is quite independent of the outcome of the second toss. It is because both outcomes have different causal conditions. In a similar way, an object a raven and black and another object and being raven and being black are probabilistically independent since the causal conditions of both events are different. So in terms of physical probability, $P(Rb.Bb|Ra.Ba) = P(Rb.Bb)$. But in our practices of confirmation, $P(Rb.Bb|Ra.Ba) > P(Rb.Bb)$. That is what is considered as probabilistically independent in physical probability are probabilistically dependent in epistemic probability.

In physical probability relevance/ probabilistic dependence is defined on the basis of causal conditions. If two events are not the product of same causal conditions then they are probabilistically independent. But in epistemic probability, probabilistic independence is defined on the basis of stock of knowledge of an agent. That is, it depends on agent's knowledge of causal conditions or agent's judgment about similarity of causal conditions or similarity of effects.

Probability involved in the practices of confirmation is inductive probability or epistemic probability.

> It is widely agreed that the concept of probability involved in confirmation is not physical probability. One reason is that physical probabilities seem not to exist in many contexts in which we talk about confirmation. For example, we often take evidence as confirming a scientific theory but it does not seem that there is a physical probability of a particular scientific theory being true. (The theory is either true or false; there is no long run frequency with which it is true, nor does the evidence have a propensity to make the theory true.) Another reason is that physical probabilities depend on the facts in a way that confirmation relations do not. Inductive probability does not have either of these shortcomings and so it is natural to identify the concept of probability involved in confirmation with inductive probability. (Maher, 2006, "Confirmation Theory", 4)

11.1 Logical Theory of Probability

Subjective and logical theories of probability are dominant theories which attempt to define/interpret the concept of epistemic probability. Keynes who propounded the logical theory of probability construes the probability relation as the logical relation. Logical relation or entailment is a notion which is worked out and well defined within the framework of deductive logic. Deductive logical relation in a sense is a part and whole relationship between propositions. According to Deductive logic, logical relations can be established only when entailing proposition implicitly or explicitly contain the entailed proposition. But how such relation can be adopted to construe a relation which is ampliative in nature? For example, deductive logic can work out a logical relation between the following two propositions. Case 1: All ravens are black and "a is a black raven". But for deductive logic there is no logical relation between 'All observed ravens are black' 'and 'a' is a black raven'. (Case II)

Keynes' point is that if a necessary connection can be worked out in the first case, it is intuitively appealing to argue that at least there is a weak necessary connection between two propositions in Case II also. As per deductive logic there is no logical relation in Case II. But yet we can see that there is a compulsion to derive the conclusion in Case II also. The position of no logical relation cannot explain rational compulsion of agent to derive the conclusion in case II. Thus Keynes argues in both cases there in an entailment relation but in the first case it is complete entailment relation, while in the second case, it is partial. And conclusive support is only special cases of larger notion of support.

> Inasmuch as it is always assumed that we can sometimes judge directly that a conclusion *follows from* a premiss, it is no great extension of this assumption to suppose that we can sometimes recognise that a conclusion *partially follows from,* or stands in a relation of probability to a premiss (Keynes, 52)

It is not clear how the notion of partial entailment can be understood. I think Keynes point is that the truth of one proposition partially supports the truth of another proposition but the relation of partial support is a logical relation. But Keynes does not specify how the relation of partial entailment or partial deduction can be shown. Moreover, Keynes does not specify how the axioms and derived rules of probability play the role at least as a necessary condition in showing partial deduction. According to him, partial logical relation or partial entailment is not something which can be shown instead it is a matter of perception. Certainly he is not referring to sense perception, by perception he refers to is something like a logical intuition. As Keynes says: "We pass from a knowledge of the proposition *a* to a knowledge about the proposition *b* by perceiving a logical relation between them. With this logical relation we have direct acquaintance" (Keynes, 13).

But in Keynes' approach we are able to form (inductive) argument because we have direct acquaintance with logical relation which is partial entailment. The approach of knowing a probability relation on the basis of logical intuitions seems as contradicting the mathematical method of probability calculus. One of the conditions to be satisfied by the interpretation of probability is the interpretation of probability axioms. Axioms of probability are considered as assumptions or obvious truths which help to draw a probability relation among the sentences with the help of definitions. If the probability relations are directly known logical intuitions it reduces the significance of the axioms. Certainly Keynes attempts to explain the significance/use of axioms. But he employs them as additional tools besides the logical intuitions. While logical intuitions can make explicit the simple probability relations, axioms are needed to recognize the complex logical relationship of probability. Here axioms are extensions of logical intuitions. Another major question is how on the basis of logical intuition we can assign the numerical value to probability. A second condition to be satisfied by interpretation of probability calculus is that it must provide a method to measure the numerical value of probability since probability calculus considers probability as a number in the interval [0,1]. To measure numerical partial entailment, Keynes adopts the principle called 'Principle of Indifference'. Keynes' point is if we can divide a set or case into mutually exclusive and exhaustive set of equiprobable cases, then probability of each set/case is $\frac{1.}{\text{Total number of sets}}$.

11.1.1 Principle of Indifference

But the question is on what basis we can determine that given cases are equi proabable. In Keynes' system, 'Principle of Indifference' is the basis of such judgement. "The name is original to him but the principle itself, he says, was introduced by J. Bernoulli under the name of the Principle of Non-sufficient Reason" (Gillies, *Philosophical Theories of Probability,* 35). Keynes gives the following preliminary statement of the principle: "The Principle of Indifference asserts that if there is no known reason for predicating of our subject one rather than another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an equal probability." Consequently Keynes holds that if cases cannot be divided into set of mutually exclusive and exhaustive and equi-probable sets, then the probability value of the set cannot be measured. That is the probability of an outcome of tossing of coin can be measured only if it is a fair coin, and if the coin is biased, probability of the outcome cannot be divided into set of equi-probable subsets. That is, the adoption of the principle of indifference considers a large number of cases where we intuitively apply probability judgment, as non-measurable probabilities and consequently cases in which probability calculus remains as inapplicable. It is also unclear that on what basis we can divide a set into mutually exclusive and exhaustive and equi-probable subsets. Moreover, the principle of validity and its applicability of indifference are challenged by the formulation of number of paradoxes.

The first of the paradoxes is called the book paradox. Consider a book in a specified place in a library. Let us suppose that we have never visited the library or seen a copy of the book. So we have no idea what the colour of its cover is. In these circumstances it could be argued that we have no more reason to suppose that the cover is red than that it is not red. Thus, using the Principle of Indifference, we have P(red) = 1/2. Similarly, however, P(blue), P(green) and P(yellow) are all 1/2, which contradicts the principle of the probability calculus that the sum of mutually exclusive possibilities must be less than or equal to 1"

Keynes responds to the paradoxes by adding a restraint that the principle of indifference is applicable only in cases where divided subsets are finite in number and indivisible. But the notion of 'indivisible' is highly unclear. The colour 'Red' certainly can be divided into further subsets like 'light Red', 'Dark Red' etc; crucial problem of Keynes modification is that Keynes' modification renders the principle of indifference inapplicable in the set which is a continuous one. But probabilities with continuous parameters occupy a central part in probability calculus. Let us now turn to a more problematic case

> There is, however, a profound difficulty connected with the principle of indifference; its use can lead to outright inconsistency. The problem is that it can be applied in different ways to the same situation, yielding incompatible values for a particular probability. Again, consider an example, namely, the case of Joe, the sloppy bartender. When a customer orders a 3:1 martini (3 parts of gin to 1 part of dry vermouth), Joe may mix anything from a 2:1 to a 4:1 martini, and there is no further information to tell us where in that range the mix may lie. According to the principle of indifference, then, we may say that there is a fifty-fifty chance that the mix will be between 2:1 and 3:1, and an equal chance that it will be between 3:1 and 4:1. Fair enough. But there is another way to look at the same situation. A 2:1 martini contains 1/3 vermouth, and a 4:1 martini contains 1/5 vermouth. Since we have no further information about the proportion of vermouth we can apply the principle of indifference once more. Since 1/3 = 20/60 and 1/5 = 12/60, we can say that there is a fifty-

> fifty chance that the proportion of vermouth is between 20/60 and 16/60 and an equal chance that it is between 16/60 and 12/60. So far, so good?
>
> Unfortunately, no. We have just contradicted ourselves. A 3:1 martini contains 25 percent vermouth, which is equal to 15/60, not 16/60. The principle of indifference has told us both that there is a fifty-fifty chance that the proportion of vermouth is between 20/60 and 16/60, and also that there is a fifty-fifty chance that it is between 20/60 and 15/60. (Earman and Salmon, The Confirmation of Scientific Hypotheses, 75-76).

The first and foremost difficulty of Keynes' theory is that the notion of partial entailment remains as an intuition. It is not developed into a theoretical basis: a basis on which partial entailment can be shown. The second problem is that the principle of indifference is not adequately developed so as to apply in various cases; especially the notion of indivisible set is not a precisely characterised one. Though Keynes introduced an highly intuitive idea of logical interpretation in every respects, it remain an imprecise characterization.

11.1.2 Carnap's Inductive Logic

The first and significant attempt to construct precise characterization of logical theory of probability owes to Rudolf Carnap which is known as inductive logic. He defines the inductive relation on the basis of language which is defined by terms and connectives of deductive logic since he works out inductive logic as a clear expansion of deductive logic. He defines both inductive and deductive relations on this language.

Earman and Salmon illustrate Carnap's inductive logic on the basis of an example which is defined on first-order logical language L. L contains only one monadic property F and only three individual contains a, b, c. Since there is only one monadic property F every individual (a,b,c) either has or lacks the property F. In this language, there are only 8 possible states of the world that can be described.

1. Fa . Fb . Fc
2. Fa . Fb .~Fc
3. Fa. ~Fb . Fc
4. ~Fa. Fb . Fc
5. Fa . ~Fb .~Fc
6. ~ Fa . Fb .~Fc
7. ~ Fa . ~Fb . Fc
8. ~ Fa . ~Fb .~Fc

Thus there are eight possible states of the world. And each possible state of the world is called a state description. Key point is that any consistent statement can be expressed in this language by these state descriptions.

1. (x)(Fx) : Fa . Fb . Fc
2. (∃x)(Fx): (Fa . Fb . Fc) v (Fa. ~Fb . Fc) v (Fa . Fb .~Fc) v (~Fa. Fb . Fc) v (Fa . ~Fb .~Fc) v (~ Fa . Fb .~Fc) v (~ Fa . ~Fb . Fc) v (~ Fa . ~Fb .~Fc)
3. Fa : (Fa . Fb . Fc) v (Fa. ~Fb . Fc) v (Fa . Fb .~Fc) v (Fa . ~Fb .~Fc)
4. Fa.Fb : (Fa . Fb . Fc) v (Fa . Fb .~Fc)

“The state descriptions in any such disjunction constitute the *range* of that statement. A contradictory statement is equivalent to the denial of all eight of the state descriptions. Its range is empty.”(86)

11.1.2.1. Defining Deductive relations

1. Entailment

"If one statement entails another, the range of the first is included within the range of the second. This means that every possible state of the universe in which the first is true is a possible state of the universe in which the second is true" (86).

E.g. Consider two propositions, (∃x)(Fx) and Fc. Here Fc.⊢ (∃x)(Fx). That is range of Fc (state descriptions (1,3,4, 7) is completely included in the range of (∃x)(Fx) (state descriptions 1-7)

2. Logically equivalent statements:

"If two statements have identical ranges, they are logically equivalent, and each one entails the other."(86)

3. Logical Inconsistency

"If two statements are logically incompatible with one another, their ranges do not overlap at all—that is, there is no possible state of the universe in which they can both be true." (86)

11.1.2.2. Defining Inductive Relations

Probability of a statement: $\frac{\text{Range of the propostion}}{\text{Total nuber of state descriptions}}$

Suppose H: (x)(Fx) and E: Fa, E supports H though it does not entail H. Carnap defines the degree of confirmation/support as follows: $\frac{\text{Range of (E.H)}}{\text{Range of E}}$

This is Carnap's first measure function called $m^{+.}$ On the basis of the measure function $m^{+.}$:

$P((x)(Fx)|Fa)) = ¼$

$P((x)(Fx)|Fa.Fb)) = ½$

"It looks reasonable to say that our hypothesis had a probability of 1/8 on the basis of no evidence, a probability of 1/4 on the basis of the first bit of evidence, and a probability of 1/2 on the two pieces of evidence." (Salmon and Earman, "The Confirmation of Scientific Hypotheses", 86) It appears that Carnap's measure function m^{+} captures the conditional probability between a universal statement and particular statements. But it fails to capture the conditional probability relation between two particular propositions. Based on Carnap's probability function, one object has property F (e.g. Fa) does not affect the probability of Fc.

Range of Fc: 4

P(Fc): 4/8 = ½

$P(Fc|Fa)$ =Range of (Fa.Fc)/Range of Fa = ½

That is, the probability function m^{+} fails to capture the learning from experience.

> When we examined the hypothesis (x)Fx … we appeared to be achieving genuine confirmation, but that was not happening at all. The hypothesis *(x)Fx* simply states that *a, b,* and *c* all have property *F.* When we find out by observing the first ball that it is red, we have simply reduced the predictive content of *h.* At first it predicted the color of three balls; after we examine the first ball it predicts the color of only two balls. After we observe the second ball, the hypothesis predicts the color of only one ball. If we were to examine the third ball and find it to be red, our hypothesis would have no predictive content at all. Instead of confirming our hypothesis we were actually simply reducing its predictive import. (Salmon and Earman, "The Confirmation of Scientific Hypotheses", 87)

By facing this difficulty, Carnap proposed another probability function m*. In m^+ all state descriptions are considered as equiprobable but in m* he introduces the notion of structural descriptions and considered structural description as equiprobable instead of state descriptions. Carnap considers the state descriptions 2, 3, 4 belonging to one structural description since all three state descriptions only say that two entities have property F and one entity lacks it. The state description 3, (S.D. 3): Fa. ~Fb . Fc can be obtained from the state description 2, (S.D.2): Fa . Fb .~Fc by replacing constant c by b and b by c. Similarly the state description 4, (S.D.4): ~Fa. Fb. Fc can be obtained from (S.D.2) by replacing c by a and a by c. In these state-descriptions, one S.D can be obtained from another by some permutation of individual constants. "A structure description in L is a disjunction of state descriptions each of which is a permutation of others" (Fitelson, "Inductive Logic", 388)

Though there are eight state descriptions, in language L, there are only four structural descriptions:

1. All are F: (Fa.Fb.Fc)
2. Two are F and 1 is ~ F: (Fa . Fb .~Fc) v (Fa. ~Fb . Fc) v (~Fa. Fb . Fc)
3. 1 F and Two ~F: (Fa . ~Fb .~Fc) v (~ Fa . Fb .~Fc) v (~ Fa . ~Fb . Fc)
4. No F: ~ Fa . ~Fb .~Fc

And Carnap assigns equal weights to each structural descriptions and then "assign equal weights to state descriptions within each structure descriptions" (Salmon, 88). m*(Fa. Fb. ~Fc)=1/12 since the proposition belongs to the second structural description and its m* is ¼ and the second structural description consists of three state descriptions which have equal-weights.

And confirmation C* is defined as follows:

$C^*(H|E) = m^*(H.E)/m^*(E)$.

Now consider, the probability of Fc conditional on the truth of Fa

$m^*(Fc) = 1/2$. $C^*(Fc/Fa) = 2/3$. It clearly indicates that m* and c* allow learning from experience.

Later, Carnap formulated more complicated probability functions to address continuum cases. But critics points out that such complicated frameworks ran into various technical difficulties. And various philosophers like Jaakko Hintikka provide various alterations to the Carnapian systems. But rather than the technical details and its troubles, what marks an end to the Carnapian project is its deeper philosophical problems. Carnapian system is undoubtedly a substantially improved version of logical interpretation than Keynes. It is improved mainly on three respects: it defines both inductive and deductive relation on the same realm, and both the relations are defined by the same notion: overlapping of ranges. Thus it provides an insight to the claim that inductive/ probability relation is a logical relation. It defines the probability relation precisely which is completely left to person's intuitive capacity in Keynes' system. For Keynes, principle of indifference applies to indivisible units. But Carnap's system specifies what are the indivisible units of the language and it also specifies how principle of indifference can be applied to indivisible units of language, that is state descriptions. Application of principle of indifference is not straight forward as Keynes thinks. Carnap specifies that first it must be applied to structural description, a category prior to the state descriptions.

11.1.2.3. Critical Analysis of Carnap

In spite of Carnap's all these incredible achievements, critics argue that, Carnap's project does not come near to his purpose. Carnap's aim is to show that the probability relation between two propositions is a logical relation. That is, the probability relation is a necessary connection and thus an objective one. The basis of Carnap's measure function is assigning prior probabilities to the statements of language. Carnap applies principle of indifference to assign prior probabilities to the statements. On the basis of principle of indifference, primarily Carnap chooses state descriptions as equiprobable statements but later chooses structural descriptions as equiprobable. Critics point is that what is the logical necessity in choosing the structural/ state descriptions as equiprobable. The principle of indifference states that if there is no known reason to choose one proposition over another, then both propositions are equally probable. Clearly the principle is based upon an agent's knowledge or all agents' collective knowledge. And the assignment of equal probability is based on the absence of a sufficient reason to choose one over another. How an agent's knowledge system can be the basis to show that there is a necessary connection between two propositions? On the basis of E, an agent may legitimately state that $P(H|E) = P(H_1|E)$. But how can we hold that $P(H|E)= P(H_1|E)$ is a necessary truth? The point is that the principle of indifference, at best is only an epistemic principle, not a logical principle.

> In what sense are Carnap's theories of logical probability (especially his later ones) logical? His early theories (based on the measure functions m^+ and m^*) applied something like the principle of indifference to the state and/or structure descriptions of the formal language L in order to determine the logical probabilities $P(\cdot|\cdot)$. In this sense, these early theories assume that certain sentences of L are equiprobable a priori. Why is such an assumption logical ? Or, more to the

> point, how is logic supposed to tell one which statements are equiprobable a priori" (Fitelson, "Inductive Logic", 390).

Since there is no logical basis to the 'principle of indifference', the only constraint for a probability function is its axioms and satisfaction of axioms is sufficient to be admissible interpretation of probability calculus. Then there can be infinitely many functions which satisfy probability axioms. Earman and Salmon work out a following probability function.

Sl.no	*State Descriptions*	*Weight*	*Structure Description*	*Weight*
1	*Fa.Fb.Fc*	*1/20*	*All F*	*1/20*
2	*Fa.Fb.~Fc*	*3/20*	*2 F, 1 ~F*	*9/20*
3	*Fa.~Fb.Fc*	*3/20*		
4	*~Fa.Fb.Fc*	*3/20*		
5	*Fa.~Fb.~Fc*	*3/20*	*1 F, 2~F*	*9/20*
6	*~Fa. Fb.~Fc*	*3/20*		
7	*~Fa.~Fb.Fc*	*3/20*		
8	*~Fa.~Fb.~Fc*	*1/20*	*No F*	*1/20*

> This method of weighting, which may be designated $m^{\triangle}$, yields a confirmation function $C^{\triangle}$, which is a sort of counterinductive method. Whereas m* places higher weights on the first and last state descriptions, which are state descriptions for universes with a great deal of uniformity (either every object has the property, or none has it), $m^{\triangle}$ places lower weights on descriptions of uniform universes. Like *c*,* $c^{\triangle}$ allows for ''learning from experience,'' but it is a funny kind of anti-inductive ' 'learning.'' Before we reject $m^{\triangle}$ out of hand, however, we should ask ourselves if we have any a priori guarantee that our universe is uniform. Can we select a suitable confirmation function without being totally arbitrary about it? This is the basic problem with the logical interpretation of probability. (Salmon and Earman, "The Confirmation of Scientific Hypotheses", 89)

Though there are philosophers who still argue for the possibility of construing probability as logical relation, it is widely held that apart from a deductive logical relation, there is no logical relation between two propositions. That is, logical probabilities do not exist.

11.2. Subjective Interpretation of Probability

Much before the Carnapian formulation of logical probability, the very idea of probability as a logical relation is severely criticised by Frank Ramsey ("Truth and Probability"). The critical analysis paves way to a new interpretation of probability called the subjective interpretation, formulated independently by Ramsey and De Finneti. Keynes' position is that in inductive inference two propositions are related by an objective logical relation. According to Keynes, we form a judgment of probability relation or rational degree of belief by analyzing/perceiving the objective relation between the two propositions.

Ramsey points out that such a construal is quite incompatible with the way we form our probability relations or the way we judge the probability relations. We form probability relations between two propositions not persuaded by an objective relation but persuaded by an argument. In view of a better argument, we may change our opinion/judgment about probability relation between the two propositions. But that does not mean that either of our judgment about probability relation is irrational. In an inductive inferential relation, different agents can reach two different conclusions, or can make two different judgments about probability relation between the given premises and conclusion, yet both can be held as rational. It is because in the judgment of probability relation or in the judgment of inductive inferential relations, we are not guided by the objective relations between propositions; instead we are guided by the argument. Ramsey's fundamental point is that there does not exist anything like an objective logical probability relation; instead the probability relations are construed on the basis of our argument about the relation between the propositions.

As per Keynes' construal, logical relation is the basic defining factor which determines the probability value. Rules of probability are logically posterior to the defined probability relation in the sense that, its application presumes an agent's ability to determine the probability relation or probability value. But Ramsey claims that often in practice we do have little agreement about the basis or factor or argument which relates two propositions. Ramsey says that "it is as if everyone knew the laws of geometry, but no one could tell whether any given object were round or square"(Ramsey, "Truth and Probability", 64) Neither do we have an agreement about the basis or factors which relates two propositions, nor do we have agreement on the probability value of a proposition conditional upon another proposition.

According to Ramsey, the impression of a partial logical relation comes from the cases which are so similar to deductive cases. Such an impression is false, since we cannot perceive such relation in many cases, especially in the case of simple propositions. "…we take the simplest possible pairs of propositions such as 'this is red' and 'that is blue' or 'this is red' and ' that is red' whose logical relation should surely be easiest to see, no one, I think pretends to be sure what is the probability relation which connects them" (64)

Ramsey's point is that we do not estimate the probability of a proposition 'this is red' by contemplating its relation with another 'that is red'. Instead, we determines the probability by considering/ analyzing our actual or hypothetical belief. Ramsey's point is that we assign probability to a proposition by analyzing how strongly we believe in the proposition. So probability of a proposition is the degree of (partial) belief on the proposition. But certainly irrational degrees of beliefs are counted as the probability. In subjective interpretation, rationality of our degree of belief can be judged only in relation with other degrees of beliefs. And rational degree of beliefs are not any set of degrees of beliefs of an agent instead they are a coherent set of degrees of beliefs. And set of degrees of belief that does not violate the axioms and derived rules of probability constitute the coherent set of degrees of beliefs. In other words, coherent or rational beliefs can be judged only on the basis of set of beliefs and the criterion of coherence/rationality is the calculus of probability. That is, according to the subjective interpretation, an agent's set of degrees of belief of a propositions that does not violate the axioms of probability constitute the probability of propositions. Since the probability axioms are the constraints of degrees of beliefs subjective interpretation satisfies the primary condition that it satisfies the rules of probability. But it needs to be analyzed how it satisfies the second condition of ascertainability of probability. It is more important in subjective interpretation of probability since it is commonly held that psychological entities like belief cannot be measured.

Ramsey argues that how far we are ready to act on a belief is the measure of belief. "…the degree of a belief is a causal property of it, which we can express vaguely as the extent to which we are prepared to act on it" (71). Thus he introduced the betting method to measure our partial belief. To measure the degree of belief of an agent A, we must set up a hypothetical betting situation, in which A is prepared to bet on proposition 'P'. Suppose P is that the next observed raven is red'. In betting, (s)he is choosing a real number in correspondence with his/ her level of conviction on the truth of the proposition. That number is called 'betting quotient' (q). Then the bookie chooses the stake 'S'. If 'P' is true, then bookies pay the stake 'S' to the agent 'A' (bettor) and A returns qS to the bookie. For example, if the stake is Rs 10/- and 'A' chooses the betting quotient .1 in correspondence with his/her level of conviction, on the proposition P and 'P' is true. Bookie payS Rs10/- to A in exchange for Rs 1/-, i.e. qS(.1×10). And if 'P' is false 'A' loses qS to the bookie. That is, 'A' pays Rs 1/- to the bookie.

Betting is one kind of action. But unlike other actions, in betting method, the extent to which we are ready to act can be determined. Betting compels an agent to choose a value even if (s)he cannot determine the numerical value of partial belief. The compulsion to choose a real number provides an advantage to the betting method, since the structure of the introspection method does not have any means to pinpoint a numerical value in correspondence with degree of belief on 'P'. But in the betting method, the extent to which we are ready to act on the truth of 'P', is expressed as the betting quotient. And the betting quotient 'q' is considered as the agent's degree of belief.

But the question is, what is the basis to consider that the betting quotient, q, the numerical expression of the agent's preparedness to act on 'P' in betting, corresponds with the agents' actual degree of belief on 'P'. Because in a betting scenario, agent's purpose or goal is not to introspect his/her degree of belief,[5] his/her aim is to win the game: that is make profit out of the game. In such circumstances, an agent can choose a betting quotient which is completely different from his/her actual degree of belief. Suppose an agent has a high level of conviction on the truth of 'P', so (s) he must choose a value near .8 or .9 to match with actual degree of belief. But if (s)he chooses the value in correspondence his/her level of conviction, then his/her pay off would be minimal in case where 'P' is true since pay off S-qs.

[5] If it is the case, then betting method is equal to introspection method

So to have maximal gain, an agent would choose 'q' which is very small in cases where (s)he has high level of conviction. That means, there would be a large disparity between the betting quotient and the degree of belief. To prevent such a possibility, subjective interpretation introduces the following constraint: the stake which is chosen by the bookie can be positive or negative, and an agent 'A' does not know whether the stake will be positive or negative while choosing the betting quotient.

Stake is positive means bookie asked the agent to book on the truth of P and his stake is negative means agent is required to book on the falsity of P (i.e. on the truth of ~P). Suppose in a case where the agent has a high level of conviction on P, and the agent chooses the betting quotient which is very small like .2. But then it means that the agent has .8 degree of belief on the truth of ~P (which (s)he himself/herself considers as very unlikely). And if the bookie asked the agent to bet on '~P' and if ~P is false (that is, P is true, as expected by the agent) (s)he would lose qS, which is comparatively huge since betting quotient q is very high.

A better way of understanding the point is the following: stake is positive means the agent who choses the betting quotient and (s)he pays qS in the event that P is true and bookie is paying the stake. But if stake is negative, the roles are swapped; agent has to provide the stake to bookie in return of qS in the event that P is true.

> If Mr…(A) knew that *S* would be positive, it would be in his interest to choose *q* as low as possible. If he knew *S* would be negative, it would be in his interest to choose *q* as high as possible. In neither case would *q* correspond to his true degree of belief. However, if he does not know whether *S* is going to be positive or negative, he has to adjust *q* to his actual belief (Gillies, *Philosophical Theories of Probability*, 55)

To a certain extent, the subjective theory succeeds in showing that an agent's betting quotient on 'P' is approximately close to agent's degree of belief on 'P'. Further, the subjectivists show that any series of betting quotient which violates the probability axioms is incoherent or irrational. The subjectivist's point is that a rational person would not venture into a betting (or choose a series of betting quotient) where (s)he is bound to lose. A person wins or loses in a betting depending upon his/her luck factor, but choosing a betting quotient on which loss is certain, is an irrational or incoherent one. That is a rational agent wants to avoid the possibility of certain loss. That is, the possibility of a Dutch book being made against him/her by a clever bookie. "A Dutch book is a set of bets such that, no matter what the outcome of the event on which the bets are made, the subject loses" (Salmon and Earman, "The Confirmation of Scientific Hypotheses", 82) And the subjectivists insist that a person is subject to a Dutch book if and only if that person holds a set of degree of conviction that violates the probability axioms. In the following tables, I illustrate the possibility of a Dutch book in case of violation of probability axioms. Following is the Standard pay-off Table:

Truth Value of P	Pay off
P is True P is False	S-qS -qS

<u>Violation of the First Axiom</u> : $0 \leq P(H) \leq 1$

First Case of Violation: $q(P) < 0$.

Suppose stake is Rs 10/, $q = -(.2)$. In this case, the bookie chooses the stake as negative.

P's Truth Value	Pay off	Net Result
T	-10-(-.2×-10)	-12
F	-(-(2)×-10	-2

In a similar way we can illustrate the case when q (P) $>$ 1

<u>Violation of the Second Axiom</u>: P(T) $=$ 1

Suppose Betting quotient, q(P)=.5, where 'P' is tautology. Then bookie chooses the stake as negative.

P	Pay Off	Net Result
T	-10-(-10×.5)	-5

Since 'P' is a tautology, there is no possibility of 'P' being false.

<u>Violation of the Third Axiom</u>: P(E∪F) $=$ P(E) $+$ P(F) where E and F are mutually exclusive events. Suppose A chooses q(P_1) $=$ .3, q(P_2) $=$ .4 and q($P_1 \cup P_2$)$=$ .5, and stake is Rs. 10. Then bookie choose stake positive for P_1 and P_2 and stake negative for ($P_1 \cup P_2$)

P_1	P_2	$P_1 \cup P_2$	Payoff			Net Payoff
			P_1	P_2	$P_1 \cup P_2$	
T	F	T	10 - 3	-4	-10 - (-5)	7 - 4 – 5 = -2
F	T	T	-3	10 - 4	-10 - (-5)	-3 + 6 – 5 = -2
F	F	F	-3	-4	- (-5)	-3 - 4 + 5 = -2

Since both P_1 and P_2 are mutually exclusive events, there is no possibility of both being true. In a similar way, we can work out a case when $q(P_1 \cup P_2) > q(P_1) + q(P_2)$.The Dutch book argument which is also known as the Ramsey – De Finetti theorem shows that a set of degree of beliefs (betting quotients) is coherent if and only if they satisfy probability axioms. In all these possible cases, the net pay off is negative. That means a clever bookie can ensure that agent faces certain loss in case of violation of probability axioms. In present discussions, subjective interpretation of probability is considered as more adequate interpretation than the logical interpretation and the credit mainly owes to ingenious Dutch Book arguments. However, the subjective interpretation is criticized on various fronts. Many hold the position that probability axioms as constraints of degree of beliefs are too weak constraints to have a useful notion or a workable notion of probability.

12. Evaluation of Theories of Confirmation

In this last section of the introductory chapter, I evaluate the four theories of confirmation on the basis of certain adequacy conditions which I have formulated in order to encapsulate the discussions of theories of confirmation. Adequacy conditions are developed from the discussion of pros and cons of various theories of confirmation. And it is also an attempt to characterize the various restraints, the structure of explication imposes upon a theorization.

Following are the adequacy conditions of confirmation:

1. Principle of confirmation should be able to capture various characteristics of confirmation which is prevalent in the practices of confirmation.

2. It should not have any consequence which is incompatible with the practices and intuitive idea of confirmation. (That is, it should be free from paradoxes of confirmation)

3. It must have one and only one interpretation in the realm of confirmation.

4. It should be able to distinguish between correctly and incorrectly justified hypotheses.

12.1. First Condition

First condition is clearly the primary objective of any explicatory project of confirmation. Following are the various characteristics of confirmations which are discussed in relation with theories of confirmation.

1. Instance Confirmation: Confirmation of hypothesis by its instances.

2. Confirmation of theoretical hypothesis

3. Confirmation by consequence of hypothesis

4. Confirmation by non-deductive evidence

5. Confirmation by explanatory evidence

6. Distinction between verification and confirmation, and falsification and disconfirmation

7. Stronger confirmation by surprising evidence.

8. Stronger confirmation by variety of evidence

9. Confirmation of quantitative hypothesis (statistical hypothesis)

Hempel's theory of confirmation is largely considered as inadequate since it fails to capture most of the characteristics of confirmation which is prevalent in the practice. It captures only one characteristic of confirmation which is instance confirmation. Instance confirmation states that instances of a hypothesis are its confirming evidence. H-D model acts much better in capturing the various characteristics of confirmation. It certainly captures the confirmation of theoretical hypotheses, instance confirmation and confirmation by true consequences. However, H-D model presumes that only the consequences of a hypothesis are evidence of the hypothesis. Thus, it fails to capture the confirmation by non-deductive evidence. For example: Drought is evidence to a hypothesis that 'There would be famine'. Severe body pain given certain background information is evidence to the hypothesis that 'The person has fever'. 'Habit of Smoking' is evidence to the hypothesis that 'The person would be infected with cancer'. Witness testimony and recovery of weapon is evidence of crime doing. Defective eye sight of a prosecution-witness disconfirms the hypothesis that witness testimony is reliable one. 'Long strides on a murder scene' confirms the hypothesis that 'murderer is a tall person'. But all these wide-range evidence are non-deductive evidence since the evidence is not necessarily implied by the truth of the hypothesis or the evidence are not the consequences of the hypotheses. For example, tall person's walk does not necessarily imply that the strides are long one. But long strides, the evidence which we have, make the hypothesis plausible.

Surprisingly the H-D model also fails to distinguish between falsification and disconfirmation. H-D model defines disconfirming evidence as false logical consequences but the false consequences are not merely disconfirmatory but falsificatory to the hypothesis. In another sense, the H-D model conflates the two notions: disconfirmation and falsification. Disconfirmation as a distinct notion from falsification is not even addressed in the H-D model. Moreover, the H-D model renders as quite inapplicable in the cases of confirmation of statistical hypothesis.

So far we have discussed six qualitative characteristics of confirmation and the IBE model of confirmation captures all the six characteristics of confirmation. Though the IBE model of confirmation is considered as a highly developed theory among qualitative theories of confirmation, it fails to capture many of the characteristics of confirmation which are related to the quantitative nature of confirmation. Primary one is 'stronger confirmation by surprising evidence'. Salmon explain it as follows: there is

> ...a widely held intuition that the more surprising the prediction a theory can make, the greater is their evidential value when they come true. A classic example of surprising prediction that came true is the Poisson bright spot. If we ask someone who is completely naive about theories of light how probable it is that a bright spot appears in the center of the shadow of a brightly illuminated circular object (ball or disk), we would certainly anticipate the response that it is very improbable indeed. There is a good inductive basis for this answer. In our everyday lives we have all observed many shadows of opaque objects, and they do not contains bright spots at the centers." (Salmon, "Rationality and Objectivity in Science or Tom Kuhn meets Tom Bayes", 187)

Poisson bright spot was the consequence of wave theory of light and the truth of this unlikely evidence provides greater confirmational boost to wave theory than the particle theory of light. The second major characteristic which consists of quantitative and comparative aspect is the stronger confirmation by a variety of evidence. Earman states that "It is a truism of scientific methodology that variety of evidence can be as important or even more important than sheer amount of evidence." (Earman, Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory, 77)

And more importantly IBE fails to account for the confirmation of statistical hypotheses. For IBE chooses the explanatory cause is the one which is present in causal history of fact and absent in the causal history of foil. But in the case of statistical hypothesis, difference is not drawn by showing a presence and its corresponding absence. But the difference is degrees of presence. Salmon argues the point as follows:

> He (Lipton) associates Mill's method of difference with the modern scientific method of controlled experimentation. I find this identification seriously problematic because controlled experiments typically yield statistical results that must be subjected to statistical analysis. Mill's method of difference does not involve statistical considerations; it is an all or nothing affair. When for example, scientists try to determine whether a substance is carcinogenic, they will administer the drugs to one group of subjects (the experimental group) and withhold it from another group (the control group). If the drug is actually carcinogenic then a higher percentage in experimental group should develop cancer than in the control group. If such a difference is observed, however, the result must be subjected to appropriate statistical tests to determine the probability that such a result would occur by chance even if drug were totally noncarniongenic. (Salmon, "Explanation and Confirmation", 70)

Though there are many staunch critics of the Bayesian confirmation theory, they mostly criticise the fundamentals assumptions of the BCT and the interpretations of the BCT's principles. Even those critics agree that given the BCT frame work, it precisely captures almost all known characteristics of confirmation. That is the main plank on which the BCT is held as the most successful theory of confirmation.

Primarily the BCT precisely formulated all characteristics in its probabilistic frame work and shows how those results or characteristics are the consequence/derivation of the BCT definition of confirmation. Since vast number of literature (Earman, *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory and* Howson and Urbach, *Scientific reasoning: A Bayesian Approach*) elaborates how the various characteristics of confirmation are the consequence of the Bayesian definition of confirmation, I demonstrate only one characteristics of confirmation (confirmation by surprising evidence) and the BCT's relation to the characteristics. Primarily the BCT defines the notion of surprising evidence. Surprising evidence means the probability of the evidence, P(E|K) is small. The smaller the value of P(E|K) the more surprising E is. $P(E|K)$ comes in the denominator of Bayes theorem, a smaller value of denominator tends to increase the value of the fraction, i.e the posterior probability of hypothesis. If the posterior probability of the hypothesis increases given a certain value of prior probability, the degree of confirmation also increases. That is confirmation becomes stronger.

I reserve the discussion of second condition for the last, as it is the main focus of my dissertation.

12.2. Third Condition

I consider third condition as the most basic adequacy condition of any theoretical construal. Failure to satisfy this condition renders the theory as invalid irrespective of whether it satisfies all other conditions or not. It is the standard norm of any kind of theorisation. In the case of Hempel's theory of confirmation, entailment (of hypothesis or development of hypothesis) by evidence is the principle of confirmation. And in the case of the H-D model entailment (of evidence) by hypothesis is considered as the principle of confirmation. The notion of ‘Entailment’ is precisely characterized in deductive logic and it has one and only one interpretation in any realm of application. In the case of the IBE model, the explanatory relation between a hypothesis and its evidence is the principle of confirmation. Lipton’s model specifies the explanation as causal explanation. In the case of the BCT, principle of confirmation is interpreted as a probability relation or function. Probability is interpreted as degree of beliefs which are constrained by probability axioms. The Carnapian inductive fails to satisfy the primary condition 3 because it fails to show that probability is a logical relation between propositions.

12.3. Fourth Condition

I consider the fourth condition as the most important condition from normative perspective. In practice, it is possible that an agent mistook a hypothesis as confirmed/ justified on the basis of certain evidence. Later on the basis of re-examination or discussion, s(he) may rectify the judgment of confirmation. A valid principle of confirmation must be able to distinguish these correct and incorrect practices. In the case of the BCT, the principle of updation of subjective belief cannot distinguish between correctly and incorrectly confirmed hypotheses. Principle of the BCT is satisfied even in the case of incorrect confirmation because an agent’s degree of belief is increased even in the case of incorrect judgment. It is because even at the point of incorrect confirmation, agent believes that hypothesis is confirmed one. But except the BCT, all theories of confirmation can easily distinguish between correct and incorrect confirmation, though their principle has only limited application compared to the BCT. It is simply because these theories characterize a confirmation relation independent of the agent.

Before coming to the discussion of condition two, which is the main focus of my dissertation, I would like to sum up my evaluation of theories. Apart from three, (which is satisfied by all theories of confirmation which we consider), I consider that primary adequacy condition is the first condition. If a given theory of confirmation fails to capture the characterization of confirmation which is prevalent in the practice of confirmation, then it there is no much sense attached in saying that it is a theory of confirmation. The basic condition of explication is that explicatum must be similar to explicandum. Satisfaction of the primary condition helps the BCT to claim the position of a front-runner among various theories of confirmation. But it cannot be ignored that the BCT is still suffered with certain infirmities which are crucial. And I articulate those infirmities as non-satisfaction of condition four. And I should concede that condition four is a crucial condition. In spite of the non-satisfaction of condition four, the BCT remains a worthwhile analysis of confirmation, because the non-satisfaction is not necessarily related to the BCT's central mechanism, Bayes theorem. And the BCT's unbeatable advantage of satisfaction of the first condition comes directly from its employment of Bayes' theorem. The non-satisfaction of condition four is not related to Bayes' theorem but rather related to the interpretation of the principle of confirmation, that is interpretation of probability. And it is evident from the fact that objective interpretation like logical interpretation clearly satisfies the fourth condition. My argument is that non-satisfaction of condition four makes a strong case for the rejection of subjective interpretation of probability calculus itself but yet the criticisms do not make a sufficient case for the rejection of the BCT as such.

12.4 Second condition

As we have discussed, the strength of the BCT, lies on the point of its closeness with practical realm of confirmation. But the claim of its closeness to the practical realm is crucially challenged by various paradoxes of confirmation. The general point of paradox of confirmation is that certain consequences of a definition of confirmation are incompatible with the intuitive notion and practice of confirmation. In the following chapters of my thesis, I analyze how major theories of confirmation respond to the paradoxes of confirmation. In the next three chapters, I analyse three paradoxes of confirmation: Paradox of Irrelevant Conjunction, Hempel's Paradox and Problem of Old Evidence. And in the discussion I only briefly analyze the response of the H-D model, Hempel's theory and IBE and the major part of discussion is reserved for the BCT's analysis of paradoxes.

Chapter I

Resolving the Paradox of Irrelevant Conjunction:

A Bayesian Approach

Introduction

In this chapter my attempt is to analyse the paradox of irrelevant conjunction which stands as a challenge to various theories of confirmation and the attempts of these theories to resolve it. The original problem of Irrelevant Conjunction (I.C.) was formulated by Clark Glymour (*Theory and Evidence*, 82) as a paradox against the Hypothetico-Deductive model (H-D model). According to the H-D model, if H entails E, then if E is observed then E confirms H. The paradox of irrelevant conjunction states that if E confirms H then E confirms the conjunction of H and any statement (X) since H.X entails E.

> For instance, intuitively, the return of Halley's Comet in 1758 (*E*) confirmed Newton's theory (*H*) of universal gravitation (relative to the background evidence (*K*) available at the time). But, according to the H–D account of confirmation, this implies that the return of Halley's Comet also confirms the conjunction of *H* and (say) Coulomb's Law (or any other proposition(s) one would like to conjoin to *H*). And, no matter how many irrelevancies are conjoined to *H, E* will continue to confirm the conjunction, according to the H–D account of confirmation. (Fitelson, "Putting the Irrelevance Back....", 612)

The Bayesian confirmation theory (BCT), which seems to be free from many of the difficulties of the H-D model, also could not resolve the paradox convincingly, even though it made an improvement towards solving the problem of irrelevant conjunction (I.C.).

In the first section of my chapter, I analyse how proponents of the H-D model revise their model in their attempt to resolve the paradox. And in the second section, I discuss the traditional Bayesian approaches to the paradox and its inadequacies. The core of the chapter starts from the third section where I characterize the modern Bayesian formulation of the paradox and the need for an adequate formulation of the notion of irrelevance. In the fourth section of the chapter, I focus on the debates of defining the notion(s) of irrelevancies / irrelevant conjunct. On the basis of their definition of irrelevance, various philosophers have worked out certain results in a probabilistic framework, which they have claimed has softened the impact of the paradox. In the fifth section, I elucidate these debates. Maher in his paper, "Bayesianism and Irrelevant Conjunction" states that what is needed is a clarification regarding the paradoxicality of confirmation of irrelevant conjunction (518-519). I pursue Maher's intuition about the notion of paradoxicality and conclude that an impression of paradoxicality is closely related to the context of confirmation in general and competing hypotheses in particular. That is the debate set out in the sixth section. And in the final section, I state that to capture such an approach, a new Bayesian analysis of confirmation has to be introduced which formally includes competing hypotheses into its framework of confirmation and I sketch an outline of a new Bayesian approach.

I.1. H-D Model and the Paradox

Before analysing the paradox in a Bayesian framework, an analysis of the attempted solutions suggested by the H-D model would be helpful in identifying the significance of a Bayesian analysis. What we have described above is only a crude form of an H-D model. In its simplest form the H-D model is as follows:

(H-D 1): E confirms H if H $\vdash$E, and E is true.

It is modified as follows because of Duhemian concern that a single hypothesis rarely entails any evidence.

(H-D 2): E confirms H with respect to Theory (T) if (1) E is true, (2) H & T is consistent, and (3) H& T $\vdash$ E.

Now the problem is that if T entails E then, according to the H-D model, E will confirm any hypothesis with respect to T. To avert such undesirable consequences the H-D model was reformulated as follows:

(H-D 3): E confirms H with respect to T if (1) E is true, (2) H & T is consistent, (3) H.T $\vdash$ E, and (4) T $\nvdash$ E.

All the three formulations of the H-D model yield the irrelevant conjunction paradox. This paradox has a devastating effect on the H-D model because it challenges even the minimally modest claim of the H-D model. The main criticism against the H-D model is that it cannot account for non-deductive evidence, so it cannot be a necessary condition for confirmation. But the H-D model had enjoyed the status of a sufficient condition for confirmation. The challenge of the paradox is devastating because it even attacks this modest claim of the H-D model. There are several brilliant attempts within the contexts of the H-D model to solve the paradox.

I.1.1. Merrill's Solutions and Glymour's Criticism

One of the significant efforts was the solution suggested by Garry Merrill ("Confirmation and Prediction"). Apparently, the problem is that irrelevant conjunct (X) is not used to derive the observational consequences. So Merrill added certain formal constraints to ensure that each conjunct of a hypothesis is necessary for the deduction of observation consequences.

Merrill's solution suggests that a hypothesis cannot be divided into conjuncts where one of the conjuncts entails evidence along with other auxiliaries. If the division is possible, and one of the conjuncts entails a piece of evidence, only then is that conjunct confirmed by that piece of evidence. Merrill reformulates the H- D method as follows:[6]

E confirms H if and only if

1. H & T is consistent
2. $H.T \vdash E$
3. $T \nvdash E$ and

there are no K, L and M such that the following conditions hold

4. $\vdash (H \equiv K.L)$
5. $K \nvdash H$,

---

[6] Merrill's formulation ("Confirmation and Prediction", 103) is slightly different from this formulation. It is the formulation of Glymour ("Hypothetico-Deductivism is Hopeless", 323) which is equivalent to the one proposed by Merrill.

6. $L \nvdash H$,

7. $M \nvdash H$

8. $\nvdash (L \equiv M)$

9. $(K.M.T) \vdash E$

10. $(K.M.T) \nvdash p.\sim p$

I think the basic idea can be read as follows: The first three condition does not pertain to solve the paradox as it only rephrases the general idea of the H-D model: evidence is the consequence of the hypothesis and if the auxiliary theory or background theory 'T' itself entails the evidence then it cannot be considered as a confirming evidence of the hypothesis in question.

Conditions (4-10) are the ones, which pertain to the solution of the paradox. Condition 4 says that it is a theorem that H is logically equivalent to K.L. That is, H can be considered as the conjunction of two propositions K and L. Conditions 5, 6, and 7 say that none of the propositions entails H alone. That is none of them is stronger than H. Condition 8 says that any of the conjunct is not equivalent to M which is basically the background information. Condition 9 states that while $H \nvdash E$, one of the conjuncts 'K' along with some proposition M (consider it as background information or as initial conditions) entails E. Merrill's point is that if a confirmation satisfies all these conditions then it is an invalid confirmation. So the basic idea of Merrill's is that conjuncts of hypothesis alone should not be able to entail the evidence. That is there should not be a conjunction such that at least one of the conjuncts entails the evidence.

In Merrill's idea E confirms H iff no conjuncts of hypothesis alone entail E. If E is an evidence of H then E should not be an evidence of any conjunct (part of the hypothesis). Though, it apparently seems that Merrill's idea blocks exactly the loophole of H-D model which allows the confirmation of irrelevant conjunction, analysis shows that it is too stringent to be a condition of confirmation. I think Merrill understood the problem in the following way: If a conjunct (C_1) entails evidence E, then the other conjuncts which are tacked to the C_1 would be deemed as irrelevant to the conjunct C_1 in relation with E. In this sense Merrill's reformulation construe the notion of irrelevant conjunct. And its consequence is that if E is an evidence of H then no part of/conjunct of H can be considered as confirmed by E. However, this is too stringent a notion of confirmation and obviously it violates the special consequence condition[7] too. Merrill views the paradox in a particular way: Irrelevant conjunct 'X' is problematic because one conjunct itself (H) is sufficient to deduce E. Next he quickly infers that if one conjunct is sufficient to deduce the evidence then the other conjuncts are irrelevant in relation with E. Merrill's intuition is right in some way because it is true that if one conjunct is irrelevant then the other conjunct by itself would entail E . That is, the deduction of E from a single conjunct or from a subset of conjuncts is a necessary feature of irrelevant conjunction. But it is not a sufficient condition to characterize the irrelevant conjunction.

Glymour notes that "Merrill's idea is a natural and plausible one: h will only be confirmed by e with respect to t if h cannot be divided into two strictly weaker sentences, k (p) and *l* (q) say, at least one of which is confirmed by e with respect to t" (Glymour, "Hypothetico-Deductivism is Hopeless", 323-324). In this paper Glymour shows that Merrill's

[7] While Merrill's solution states that If E confirms H then E confirms no consequence of H, the special consequence condition formulated by Hempel states that if E confirms H, then it also confirms every consequence of H.

solution also has the untoward result of reducing the scope of the method to the point where E confirms H only if E entails H.

Glymour's Criticism:

If E directly confirms H on T by Merrill's conditions (10 conditions) then E ⊢ H

Proof:

Suppose that H, E, satisfy Merrill's condition 1 (H & T is consistent), Merrill's condition 2 (H.T ⊢ E) and Merrill's condition 3 (T ⊬ E)and suppose that (*) ⊬ (H≡ (T⊃E) and let K be (T⊃E) and L be (T⊃E) ⊃ H and M be any tautology.

1. ⊢ (H≡ K.L):	definition of K.L : {(T⊃E) ⊃ H). (T⊃E) ≡ H} (M.P) Thus, Merrill's 4th condition holds:
2. K ⊬ H	:Because of the supposition *(⊬(H≡(T⊃E) and Merrill's condition 2. That is, if K ⊢H, i.e, (T⊃E) ⊢H then by Merrill's condition 2, (H≡(T⊃E). This contradicts the assumption *. Thus Merrill's 5th condition holds.
3. L⊬ H.	(Because of Merrill's condition 3. If L⊢ H then it contradicts Merrill's condition 3**.**

If $L \vdash H$, i.e. $(T \supset E) \supset H) \vdash H$ then

$\sim H \vdash \sim ((T \supset E) \supset H))$

Then $\sim H \vdash (T \supset E) . \sim H$

Then $\sim H \vdash (T \supset E)$

But by Merrill's condition 2, if H

$\vdash (T \supset E)$ then

$\vdash ((T \supset E) \equiv T \vdash E)$

It contradicts the premise 3.

Thus Merrill's 6th condition holds

4. $M \nvdash H$ because M is tautology, By Merrill's condition 2 and 3, H is not a tautology.
5. $\nvdash (M \equiv L)$ because M is tautology while L is not.
6. $K.M.T \vdash E$. : $\{T \supset E). T \vdash E$. (M.P)
7. $K.M.T \vdash p. \sim p$ (because it is not a contradiction).

If E directly confirms H on T by clause 1

8. $\vdash (H \equiv T \supset E)$:$H.T \vdash E$. by Merrill's condition 2.
9. $E \vdash (T \supset E)$ by premise 8.
10. $E \vdash H$.

And 8 contradicts the supposition (*) therefore k, l, m statements cannot be derived.

I.1.2. Waters' Solutions

Another major attempt to rescue the H-D model is a solution proposed by Kenneth Waters. Waters appreciates the line of thinking of Merrill that if division is possible and one part entails the hypothesis then certainly it is problematic. But Waters differs from Merrill on the point that how the division should be understood. Merrill's division was purely on logical terms. Merrill divided hypothesis into separate units on the basis of the rule of simplification and then assumes that one unit's confirmation is independent of the confirmation of other units whereas the units of hypothesis are conjuncts. Waters' point is that division of hypothesis contains many extra logical considerations. So, to understand the redundant part (irrelevant unit of hypothesis) we need to have a dividing rule substantially different from the rule of simplification. Rule of simplification does not create units which work as independent from other in terms of its truth. To make this point he employs Fresnel's theory of diffraction. He argues that, in logical terms, Fresnel's theory could have been broken up into two mutually exclusive sub theories: one is a theory of diffraction in Arago's laboratory and the other is about all other diffraction phenomena.[8] Argo's famous experimental test provides a greater support to theory of diffraction which is based on wave theory of light. Suppose we spilt Fresnel theory of diffraction (a theory of diffraction which is based on wave theory of light) into different conjuncts. One of the conjuncts is the theory of diffraction in the case of circular objects. And that conjunct itself is sufficient to entail the evidence: Argo spot. Then as per Merrill's condition, Argo's spot does not support the wave theory of diffraction in toto. It only supports the wave theory of diffraction in the case of

[8] Fresnel explains the diffraction phenomenon by using wave theory of light. His successful calculations of diffraction works as confirming evidence to the wave theory of light. But the greater confirmation of wave theory of light comes from Argo's famous experimental test. This experiment discovers what is called as Argo spot or Fresnel bright spot. It is a bright spot or point which appears at the centre of the shadow of a circular object. The normal assumption was that the shadow should be completely dark as the object obstructs the path of light. Fresnel theory of diffraction which is based on the wave theory of light offers an explanation not only to the Argo spot but to all other diffraction phenomenon. But there was substantial difference between Argo's spot (diffraction phenomenon when the object is circular) and other kinds of diffraction. The difference was that particle theory of light could not explain Argo's spot though it could explain other kinds of diffraction phenomenon at least in qualitative terms. Thus the diffraction of light in the case of circular object is considered as a special kind.

circular objects. But it is counter intuitive because Argo spot is widely held as greater confirming evidence of wave theory of diffraction than other optical phenomena like refraction and diffraction. But scientists believe that it provides support to the wave theory of all other diffraction phenomena as well. That is though it is true that Fresnel's theory can be divided into two sub theories by simplification, scientists seemed to have had good reason not to split the theory. Waters objects to Merrill's way of division because in Fresnel's theory truth or falsity of one conjunct has some impact on another conjunct's truth value. So Waters' general idea for reconstructing the H-D model is as follows:

> The general idea is to obtain separate packets of information whose truth or falsity will behave as individual units. This process yields "unit statements" of closely connected information, which are tested by using them to deduce predictions. A unit statement of a hypothesis is tested only if it played a necessary role in the deduction of the predicted evidence. Thus, according to this view, a hypothesis made up of a conjunction of unit statements is directly confirmed only if the following conditions are satisfied:
>
> a). H and T are consistent
>
> b). H, T $\vdash$ E.
>
> c). For every unit statement, S, of H: H-S, T$\nvdash$ E
>
> (Where H-S is the conjunction of all the conjuncts of H except for S). (Waters, 457)

Waters' approach is a step forward in the line of Merill's approach. The problem with Merrill's type of reconstruction is that it excludes extra-logical considerations that help set the proper scope of the tested hypotheses.

> This proposal [Waters' proposal] solves the irrelevant conjunction problem because it leaves no place to tack on a statement which is irrelevant to the evidence. Suppose e confirms H. If an irrelevant statement is added as a unit statement of H, then H will not be confirmed (because condition C will not be satisfied). This leaves the

> possibility of conjoining an irrelevant statement with a unit statement of H. (453)

The following example illustrates the point. Various tenets of Copernicanism along with Kepler's three laws of planetary motion can be called as ‘K.C’ theory: C & KI & K2 & K3. Consider evidence ‘O’ that concerns only relations of single planets. Waters’ point is that in relation with the evidence ‘O’, K3 is a unit independent from the other part C & KI & K2. It is because the third law (K3) is about a relation between different planets whereas evidence O' concerns only relations of single planets. That is, though ‘O’ could be a confirming evidence for C & KI & K2, it cannot be a confirming evidence for K3. Certainly this division is not merely based on simplification. But Waters’ point is that such a division can be envisaged. But the problem does not end there. Waters’ point is that evidence does not confirm a unit which is not necessary for the deduction of evidence. But the trouble is that hypothesis can be reformulated in various ways such as even the redundant part is necessary for the deduction of evidence. “Glymour indicates that there are alternative Axiomatizations, logically equivalent to K-C, such that the third law, and neither the first nor the second law, is necessary to deduce O'” (458). K-C' = (K3 ⊃ C). (C ⊃ O') & KI & K2 & K3.

It is a logically equivalent axiomatization of K1 but K3 is necessary to deduce the evidence ‘O’ whereas K1 and K2 are not necessary. Then the question would be that on what basis, we would decide that a a particular unit is not necessary for the deduction of evidence**.**

Water’s point is that such alternative axiomatization is not possible under the relevance logic:

> In R-logic, however, K-C' and K-C are not the same theory; K-C' is not logically equivalent to K-C. For neither (K3 ⊃ C) nor (C ⊃ O') can be derived from K-C (that is, from C & KI & K2 & K3). In

> effect, R-logic serves to select out the problematical alternative axiomizations. Glymour's objection rests on his appeal to classical logic in order to draw connections that the theory does not provide. (462)

But the trouble with Waters' solution is that, to defend his solution against criticism of Glymour, he ventures into the denouncements of classical logic and adopts relevance logic. And the basic trouble is that his solution rests upon the idea that hypothesis can be divided into certain basic units whose truth values are independent. But he fails to provide any adequate method for such a division.

I.1.3. Gemes' Reformulation of the H-D Model

Perhaps the most successful attempt to revive the H-D model is that of Ken Gemes. For the analysis of various solutions and the formulation of his analysis, Gemes defines the theory / hypothesis as follows: "...a theory is specified by a finite axioms set, that is, a finite set of statements such that every member of the theory is a consequence of that set" (Gemes, "Hypothetico-Deductivism, Content, and the Natural Axiomatization of Theories." 477). That means the hypothesis 'All ravens are black' can be expressed by the sole axiom (x) (Rx⊃ Bx) and as mentioned above all members of theory like (Ra.Ba) and (Rb.Bb) are consequence of the axiom.

He notes that some primary attempts to rule out paradoxical cases insist on the point that axioms which are not needed in the derivation of evidence are not confirmed by the evidence E. Consider the following example.

H: (x) (Px⊃ Qx)

H.X: (x) (Px⊃ Qx). (x) (Fx)

E: Pa.Qa.

In the above example it is clear that irrelevant conjunct X: '(x) (Fx)' is not needed for the derivation of E. Gemes notes that such a situation tempts to formulate a condition that evidence E, H-D confirms only the statements or conjuncts of hypothesis H which are necessary for the derivation of E. As a result, in the above example, E does not confirm the irrelevant conjunction H.X in toto, but it only confirms H. But the immediate challenge is that any hypothesis can be reformulated in such a way that even the intuitively redundant axioms are necessary for the derivation of E. For example the hypothesis H.X can be reformulated as follows: (x) (Fx) ⊃ (x) (Px⊃ Qx) . (x) (Fx).

(x) (Px⊃ Qx).(x) (Fx) is logically equivalent to ((x) (Fx) ⊃ (x) (Px⊃ Qx)) . (x) (Fx).

And to derive E (Pa.Qa) from ((x) (Fx) ⊃ (x) (Px⊃ Qx)). (x) (Fx), the irrelevant conjunct (x) (Fx) is needed. This shows that the above constraint does not prevent the confirmation of an irrelevant conjunction. But then the immediate response would be that E confirms the axioms (conjuncts) of T[9] only if those axioms (conjuncts) are needed in the derivation of E under any equivalent formulation of H. Certainly such a condition prevents the confirmation of irrelevant conjunction but it cannot be adopted as the condition of confirmation as it prevents the legitimate cases of confirmation also.

For example, consider following two cases:

H_1: (x) (Px⊃ Qx). (Pa⊃ Qa), H_2: (x) (x=a ⊃ (Px⊃ Qx). (x) (x≠a ⊃ (Px⊃ Qx) and E is Pa.Qa.

Here both H_1 and H_2 are the logically equivalent hypothesis of H ((x) (Px⊃ Qx)). However, to derive the evidence Pa.Qa from the H_1 the conjunct (x) (Px⊃ Qx) is not needed though we know that E is a confirming evidence of (x) (Px⊃ Qx). And in the case of H_2 also we can

[9] In this section of Gemes' Solution, I use T (theory) and H (hypothesis) as equivalent terms since Gemes too uses the terms as equivalent.

derive E without recourse to the sole axiom (conjunct) (x) (Px⊃ Qx). That means the axiom (x) (Px⊃ Qx) is not needed in the derfivation of E in a particular axiomatization. But intuively E is a confirming evidence of the axiom. Thus the added condition remains as inadequate. According to Gemes, the way out is excluding the unnatural Axiomatization of theory.

> Here then we have a suggestive solution to the tacking problem [tacking an irrelevant conjunct to a hypothesis]: Where T together with background evidence b entails e, e Hypothetico-deductively confirms axiom A of T iff for any natural axiomatization of T, A (or at least some logical equivalent of A) is needed in the derivation of e from b. The problem is that no one has yet come up with an adequate analysis of the notion of the natural axiomatization of a theory. (479)

Gemes' idea of a natural axiomatization is based on the theory of content. Traditionally every contingent consequence of a sentence/theory is counted. Gemes points out that there are some serious drawbacks to the traditional notion of content. An important drawback, according to the traditional theory, is that any two theories T and T_1 will have common content if ∼T ⊬T1. "According to T.C.1,(traditional content theory) for any (two) wffs α and β, as long as α and β are contingent and [∼α] does not entail β, α and β have at least one common content part, namely [(α V β)]" (Gemes 1994 "A New Theory of Content I: Basic Content."p. 597). And in the same way, intuitively H v X is not a content of H but traditional content theory states it as a content.

Ken Gemes notes that the major problematic element in the traditional theory of content is the rule of addition of classical logic. Arbitrary disjunction often causes trouble for the traditional theory of content. Thus, for p.q, the content is not only p and q but also p v r and q v s. And the recourse to relevant logic is also not much helpful, since most types of relevant logic allows the deduction of $(\alpha v \beta)$ from α. The more serious problem in adopting relevant logic is its abandonment of classical equivalence. That is, if relevant logic is adopted to define the theory of content, then we end up in a position where there are different content for classical equivalent hypotheses. And the abandonment of classical equivalence is often counter intuitive.

Thus Gemes seeks other ways to redefine the notion of content. Gemes attempts for a reformulation by focusing on the trouble that is mainly related with arbitrary disjunction, i.e. with the rule of addition. This is attempted by characterizing the nature of adding a disjunct. By adding a disjunct to a proposition we are basically weakening the proposition. We add a disjunct 'q' to a proposition 'p'; the resulting disjunction p v q is weaker than the original proposition. It is weaker in the sense that original proposition 'p' entails the disjunction p v q, but the disjunction does not entail 'p'. That is, disjunction is weaker than the original proposition. Gemes picks out this characteristic of disjunction to rule out arbitrary disjunctions from the realm of content. Gemes' point is that if α and σ are two consequences of β and α is stronger than σ and the atomic wff (well formed formula) which occurs in σ also occurs in α then σ cannot be considered as the content of β. His definition is as follows:

"$\alpha < \beta$[10] $=_{df}$ α and β are contingent, β ⊢ α, and there is no σ such that β ⊢σ, σ is stronger than α, and every atomic wff that occurs in σ occurs in α"(Gemes, Hypothetico-Deductivism, Content, and the Natural Axiomatization of Theories, 481). Further, Gemes elaborates how his new theory overcomes the infirmities of the traditional theory of content. But Gemes points out that despite these advantages it suffers from some

[10] $\alpha < \beta$ means α is a content part of β

drawbacks. Following are the unintuitive or problematic consequences of Gemes' theory of content:

1. p v q is not the content part of p . q since both p and q are the content part of (p.q)

2. (pvq) . (pvr) is not the content part of (pvq) . (pvr) . (qvr) though (pvq) and (pvr) are the content part of (pvq) . (pvr). (qvr).

Gemes considers these difficulties as minor drawbacks as his major task is to define the basic content of a proposition or hypothesis which is wff. His position is that p v q is not the basic content part of p . q while p and q are. (p v q) . (p v r) are not the basic content part of (p v q) . (p v r) . (q v r) while (p v q) and (p v r) are the content part. The content part like (p v q) . (p v r) is given to us by the application of disjunction or conjunction upon the basic content part. Therefore, one could provide a sophisticated account of content by adding suitable recursive clauses to the definition so as to include all content parts. For example, one could refine Gemes' definition by saying that conjunctions and disjunctions, by using basic content of the theory, are also the contents of theory. But it has to be emphasized that the content is not the disjunction derived from the basic content of theory, but the disjunction made by using basic contents of theory. That is, disjuncts must be only from the set of basic contents.

Nevertheless, Gemes points out that there is a serious drawback to the new theory of content as it violates equivalence condition. Equivalence condition of the theory of content is the following: "The set of content parts of a formula should be closed under the relationship of (classical) logical equivalence. After all, if two propositions are logically equivalent they have the same content and hence if one is a content part of a third statement then so is the other"(Gemes, A New Theory of Content I: Basic Content,604).

Following example shows how Gemes theory of content violates the equivalence condition: p and (pvq).p are logically equivalent propositions. p is the content part of p.q while its logical equivalent (pvq).p is not the content part of p since pvq is defeated by p.q. To preserve the equivalence condition Gemes reformulates the theory of content as follows: Here then is our final definition for basic content parts of wffs of L. BCPL1 (Basic Content Part of Language L: definition 1) gives a definition of the basic content parts of arbitrary wff of language L, hence we use '$<_b$' to indicate the relationship of being a basic content part:

"BCPL1: $\alpha <_b \beta =_{df}$ (i) both α and β are contingent, (ii) α is a consequence of β and (iii) for some μ , μ is equivalent to α and there is no σ such that σ stronger than μ, σ is a consequence of β and every atomic wff that occurs in σ occurs in μ" (605) .

Through this modification his point is that α is a content part of β even if there is a defeater 'μ' (that is there is a 'μ' which is the consequence of β and every atomic wff that occurs in α also occurs in , but μ is stronger than α) if defeater is logical equivalent to α. In a sense, all logical equivalent wffs of μ are considered as the content part of β, irrespective of the fact whether they are weaker than the equivalent wffs of μ .

Then Gemes makes an attempt to reformulate the H-D model of confirmation on the basis of his new theory of content. Gemes' focal point in the reformulation of the H-D model is that the relevant conjuncts of H (relevant axioms of H) are necessary for the deduction of E in all natural axiomatization of H. That is, E confirms H in toto only if all conjuncts of H are necessary for the deduction of E in all natural axiomatization of H. That is, to be confirmed in toto, H has to satisfy the following two conditions apart from the primary condition that H ⊢ E.

1. All conjuncts of H are necessary for the deduction of E from H.
2. Al conjuncts of H are necessary for the deduction of E in all natural axiomatization of H.

That is, if a conjunct X is not necessary for the deduction of E from H or its natural axiomatization then it is not confirmed by E. And Gemes' intuition is that irrelevant conjunct would not be necessary for the deduction of E at least in one natural axiomatization of H.

Next the formidable challenge is to define 'natural axiomatization'. Gemes' rough idea of natural axiomatization is as follows: if H and H_1 are logically equivalent hypotheses and all conjuncts of H_1 (all axioms of H_1) are the content of H and H_1, (content as defined by Gemes) then H_1 is a natural axiomatization of H. Consider the previous example which shows that even irrelevant conjuncts too are necessary for the deduction of E. Gemes shows through new theory of content that such axiomatizations are unnatural.

E: Pa.Qa.

H: (x) (Px ⊃ Qx). (x) (Fx)

H_1 : (x) (Fx).((x)(Fx) ⊃ (x)(Px ⊃ Qx))

Obviously (x) (Fx) in H is redundant for the deduction of E though it is necessary in H_1 which is logically equivalent to H. Gemes' point is that such reformulation(reaxiomatization) is unnatural as per his new theory of content because the conjunct of H_1 (the axioms of H_1) is not the content part of H or H_1. It is because (x) (Px ⊃ Qx) which is the consequence of H and

H_1 and is stronger than ~(x) (Fx) v (x) (Px ⊃ Qx) and yet contain only atomic wff occurring in ~(x) (Fx) v (x) (Px ⊃ Qx). Primarily Gemes defines the natural axiomatization as follows: "T' is a natural axiomatization of T iff T' is a finite set of wffs such that T' is logically equivalent to T and every member of T' is a content part of T'..." (Gemes, "Hypothetico-Deductivism, Content, and the Natural Axiomatization of Theories", 482)

However, the definition of natural axiomatization does not solve the problem of I.C. that does not solve the problem. Consider the following case: H_1: (Pa.Qa). (x) (Px ⊃ Qx). and H: (x) (Px ⊃ Qx) is equivalent to H and H_1 is a natural axiomatization of H as every conjuncts of H_1 is a content of H. Here we may be tempted to identify (Pa.Qa) as a redundant axiom and may add another constraint that 'natural axiomatization of H' should not contain any redundant axiom, but such a constraint is ambiguous and too weak. So Gemes provides an improved version of natural Axiomatization; "T* is a natural axiomatization of T iff: (i) T* is a finite set of wffs such that T* is logically equivalent to T; (ii) every member of T* is a content part of T*; and (iii) no content part of any member of T* is entailed by the set of the remaining members of T*."(483)

Then on the basis of the idea of natural axiomatization, Gemes states that E confirms the conjunct C_1 of H only if C_1 is necessary for the deduction of E in all natural axiomatizations of H. His precise definition of H-D model is as follows: "e hypothetico-deductively confirms axiom A of theory T relative to background evidence b iff (T&b) ⊢e and there is no natural Axiomatization, n(T) of T such that for some subset s of the axioms of n(T), (s&b) ⊢ e and A is not a content part of (s&b)" (483). Gemes argues that irrelevant conjunct is not necessary for the deduction of E in all natural axiomatization of H. Therefore, irrelevant conjunction H.X is not confirmed in toto by E though H ⊢E.

I.2. Traditional Formulation of the Paradox in Bayesian Confirmation Theory's Framework

Bayesian Confirmation Theory's (BCT's) claim is that the BCT is free from the original problem of Irrelevant Conjunction. There are several cases where evidence E confirms (according to the BCT) hypothesis H, given background information, K[11] and E does not confirm H.X where X is any statement. That is, unlike the H-D model of confirmation, in the Bayesian Confirmation theory (BCT), confirmation of H by E does not necessarily imply the confirmation of H.X by E, where X is any statement. It would be the case when X is a supporting statement.

Theorem 1

$P(H|E.K) > P(H|K) \nvdash P(H.X|E.K) > P(H.X|K)$ where X is any statement.

i.e there are instances where $P(H.X|E.K) \ngtr P(H.X|K)$ while $P(H|E.K) > P(H|K)$. Following are the some of the instances: 1. Where X is a contrary statement either to H or E. Or where $P(H|X.K) < P(H|K)$ or $P(E|X.K) < P(E|K)$.

Though the BCT is immune to the original problem of I.C., it is affected by a special case of Irrelevant Conjunction. In the case of deductive evidence, the confirmation of H implies the confirmation of H.X by E. That is, $H \vdash E$, (then it assumes that H is confirmed) implies the confirmation of H.X where X is any statement.

Theorem 2

In the case of deductive evidence, $(P(E|H.K)=1)$,

[11] Background information plays a crucial role in the Bayesian confirmation theory. "A key insight of Bayesian confirmation theory is that it represents confirmation as a three place relation between evidence E, hypothesis H, and background corpus K." (Fitelson and Hawthrone. "How Bayesian Confirmation Theory Handles the Paradox of the Ravens." 10).

$P(H.X|E.K) > P(H.X|K)$ where X is any statement provided that $Pr(E|K) < 1$.

From the above it is clear that if the $P(E|K)<1$ then $P(H.X|E.K) > P(H.X|K)$. So a Bayesian too faces the challenge of irrelevant conjunction in the case of deductive evidence. This is the traditional formulation of the paradox of irrelevant conjunction as a challenge to the BCT.

I.3. Traditional Bayesian Solution to the Paradox

The basic intuition of the traditional solution is that adding an irrelevant conjunct (X) to a confirmed hypothesis (H) would reduce the strength of the confirmation of conjunction (H.X). That is an added irrelevant conjunct plays a negative role in the degree of confirmation. They formulate the idea as follows: If X is an irrelevant statement, then $P(H|E.K) \geq P(H.X|E.K)$. Even in the case of deductive evidence, $P(H.X|E.K)$ would not exceed $P(H|E.K)$.

But Branden Fitelson points out that the above formula which the traditional view puts forward fails to capture its own intuition. (Fitelson, "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction", 612) The above formula is true not only for irrelevant conjunction but also for all relevant conjunction. The traditional solution only states the fact that no conjunction can be more probable than its conjuncts. The intuition of the traditional solution is that if X is an irrelevant conjunct to H in relation with E then E confirms H more than it does H.X. Traditional solution mixes up the notion of confirmation and probability. $P(H|E.K)> P(H.X|E.K)$ only states that H has higher probability than H.X given E.K. It does not say that E confirms H more than H.X.

I.3.1. John Earman's Solution to the Paradox

John Earman works out a formula which precisely captures the traditional intuition (Earman, *Bayes or Bust: A Critical Examination of Bayesian Confirmation Theory*, 64-65). Earman define the notion of 'confirms more' and tries to works out a characteristic of irrelevance conjunct on the basis of it. Earman defined the measure of confirmation as the difference between the posterior probability and the prior probability. Generally it can be expressed as c(H|E.K): degree of confirmation of H by E given K. In Earman's account E confirms H more than H.X if $P(H|E.K) - P(H|K) > P(H.X|E.K) - P(H.X.|K)$ i.e., if $d(H|E.K) > d(H.X|E.K)$. According to Earman, if 'X' is an irrelevant conjunct then the above theorem holds even when H entails E.

<u>Theorem 3</u>

$P(H|E.K) - P(H|K) > P(H.X|E.K) - P(H.X.|K)$ when H entails E

Shortcomings of Earman's account

1. Earman's account is measure sensitive.

Bayesians have developed different ways for measuring confirmation. Following are the major ones:

Difference measure (d): $d(H, E|K) = P(H|E.K) - P(H|K)$

Ratio Measure: (r) : $r(H,E|K) = P(H|E.K) / P(H|K)$

Likelihood measure : $l(H,E|K) = P(E|H.K) / P(E|\sim H.K)$

Earman's account of irrelevance does not work if one chooses the ratio measure instead of the difference measure. Since we do not have any sufficient argument to reject the ratio measure in favour of any of the other measures, one has to explain why Earman's account of irrelevance fails in the case of ratio measure.

2. Earman's formula is not a sufficient condition.

Fitelson criticizes the Earman formula as follows "A conjunct X could be (intuitively) relevant to H and E, but this would not prevent the conjunction H&X from being less strongly confirmed than H by E."(Fitelson, "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction", 615) That is, Earman's formula d(H.X|E.K)< d(H|E.K) is true even in a case where H.X is a relevant conjunction. That is the characteristic of being less confirmed than the original hypothesis is not sufficient to demarcate the irrelevant conjunct from a relevant one. Consider the following example where we can see that even relevant conjunctions satisfy the Earman formula. H: Wave theory of light E: Diffraction of light X: Light travels through empty space. It is obvious that d(H,E|K) would be high, since E has a strong bearing on H given K. Since both H and X are contraries given the back ground information that wave does not propagate in empty space d(H.X,E |K) would be zero. That is, d(H.X,E |K) < d(H,E |K). Contrary statement X is a relevant statement to the confirmation relation of H and E. That is, Earman's formula holds even for relevant conjunctions.

I.3.2. Roger Rosenkrantz's Solution

Roger Rosenkrantz states Earman's solution in a different way which sheds light on the notion of irrelevance more than any other attempts (Rosenkrantz, "Bayesian Confirmation:

Paradise Regained" 471). Rosenkrantz tries to extract the notion of irrelevance which is implicit in Earman's solution. Rosenkrantz works out the following result:

Theorem 4

d(H.X|E.K)= P(X|H.K)× d(H|E.K) when H entails E.

According to Rosenkrantz, if X is irrelevant then d (H.X|E.K) < d(H|E.K) and it is possible only if P(X|H.K) $<$ 1. So, if X is irrelevant conjunct to H then P (X|H.K) $<$ 1. As Fitelson points out it is the first attempt to define irrelevance from the Bayesian perspective. And all other attempts assume X as irrelevant and subsequently try to show a quantitative relation between irrelevant conjunction and hypothesis. In spite of these advantages, his account too suffers with certain drawbacks.

1. It is more measure sensitive than Earman's account as Rosenkrantz's solution does not work in the case of either r or *l*.
2. Here the degree of conditional probability (if less than 1) determines the irrelevance. It too cannot be a sufficient condition since for most of the confirmed hypothesis P(H |E) $<$ 1 but that does not mean that H and E are irrelevant.
3. Moreover, Rosenkrantz only attempts to define confirmational irrelevance of X only in relation with H. But Fitelson argue that evidence E too plays a role in determining the confirmational irrelevance of X.

I.4. Modern Formulation of the Paradox

In the Hypothetico –Deductive model, the paradox arises when X is assumed to be any statement. That is, X can be an irrelevant statement and X can be a contrary statement. As per the traditional Bayesian formulation of the paradox, BCT faces the same challenge in the case of deductive evidence. Even in the case of deductive evidence, if X is a contrary statement to H, P(E |H.X.K) would be indeterminate.

P(E |H.X.K) = .$\frac{P(E.H.X.K)}{P(H.X.K)}$.[12] Both numerator and denominator would be zero If X is contrary to H. Such a result is not much significant in the sense that traditional Bayesian were also interested in the case where X is an irrelevant statement. But there is a significant result for redefinition of X as an irrelevant statement and not as any statement. In this traditional formulation, paradox arises only in the case of deductive evidence. The modern formulation of I.C. rules out the cases where X is a statement contrary to H. It considers the cases where X is not any statement but X is a statement which is irrelevant to the confirmation relation between H and E. In the case of such an X, paradox arises not only in the case of deductive evidence, but in all cases of confirmation.

<u>Theorem 5</u>

"If E confirms H and X is confirmationally irrelevant to H, E and H.E (relative to K) then E also confirms H.X (relative to K). According to Fitelson, this is the problem of irrelevant conjunction that Bayesian confirmation theorists should be worrying about." (“Putting the Irrelevance Back into the Problem of Irrelevant Conjunction”, 617).

---

[12] As per definition of conditional probability, P(E|F) = $\frac{P(E.F)}{P(F)}$.

I.5. Criterion of Irrelevant Conjunct

1.5.1. Fitelson's criterion

Before proceeding further it is necessary to define the notion of irrelevant conjunct. Fitelson defines irrelevance in terms of probabilistic independence. "A is confirmationally irrelevant (relative to *K*) just when A & B are probabilistically independent (given *K*) - i.e. when P(A.B|K) = P(A|K) × P(B|K)" (617). Then Fitelson argues that 'X' is confirmationally irrelevant to H iff X is irrelevant to H, E and all logical combinations of H and E.

X is confirmationally irrelevant to H means

1. P(H.X |K) = P(H |K) × P(X|K)
2. P(E.X |K) = P(X |K) × P(E|K)
3. P(X .(H.E)|K) = P(X |K) × P(H.E)|K)
4. P(X .(HvE)|K) = P(X |K) × P(HvE)|K)
5. P(X .(H⊃E)|K) = P(X |K) × P(H⊃E)|K)
6. P(X .(H≡E)|K) = P(X |K) × P(H≡E)|K)

I.5.2. Fitelson – Hawthorne Criterion

Hawthorne and Fitelson find Fitelson's original criterion suggested above as overly strong. And they modified Fitelson's original criterion as follows: "X is an irrelevant conjunct to H given K with respect to E just in case P(E |H.X.K) = P(E |H.K) that is when E is independent of X given H together with K" (Fitelson and Hawthorne, "Discussion: Re-solving Irrelevant Conjunction with Probabilistic Independence", 509). They explain the intuition of the solution as follows: “in deductive case the intuition about the irrelevance of X flows from the idea that a hypothesis is tested by what it says about evidence. Adding X to H (given K) says nothing more about E than H (given K) already says” (508-509).

Fitelson’s original solution is abandoned because it is found as too stringent. But it seems that the criticism is unfair because though they claim that original solution is too stringent, they do not mean that the original solution cannot be a necessary condition of irrelevance. They used the notion of being stringent in a different way. Their claim is that the modified solution is much more an intuitively direct version of the idea that X is an irrelevant conjunct. That is, the original solution is stringent because it is an indirect version of our intuitive idea of irrelevant conjunct. I am claiming that the modified solution though intuitively seems as a necessary condition of an irrelevant conjunct, it is not a sufficient condition to demarcate irrelevant conjunct from a relevant conjunct as many of the relevant conjunctions too satisfy the modified solution proposed by Fitelson and Hawthorne.

I.5.3. Modified Version of Rosenkrantz’s Criterion

I claim that the following formula is an adequate definition of irrelevant conjunct. ‘X’ is confirmationally irrelevant to H given K with respect to E if and only if $P(X|H.E.K) = P(X|H.K)$. This is a modified version of Rosenkrantz’s formula. Rosenkrantz primarily defended Earman’s solution, provided that x is irrelevant to the confirmation relation between H and E, given K, $d(H.X|E.K) < d(H|E.K)$. And Rosenkrantz tries to extract out the notion of irrelevance which is implicit in Earman’s solution. Rosenkrantz works out the following result: $d(H.X|E.K) = P(X|H.K) \times d(H|E.K)$.

If X is irrelevant, then d(H.X|E.K) $<$ d(H|E.K) and it is possible only if P(X|H.K) < 1. So, if X is an irrelevant conjunct to H, then P(X|H.K) $<$ 1 Rosenkrantz's criterion of irrelevance is obviously flawed since there are several examples in which the statements and X satisfy the formula ' P(X|H.K) $<$ 1' but are relevant to each other. But a close analysis would reveal that Rosenkrantz's criterion of irrelevance not only merely concerns about the absolute irrelevance, but also tries to account for the process of being irrelevant and it is very vivid in Rosenkrantz's writing. He writes as follows: "Indeed the degree of compatibility of X with H should control the rate of depreciation and this is what (the theorem) says" (Rosenkrantz, "Why Glymour is a Bayesian", 92). If P (X|H.K) $=$ 1 then we can consider that X is absolutely relevant to H unless X is a tautology and any decrease from 1 would be its process of becoming irrelevant to H. Certainly P(X|H.K) cannot be less than p(X) unless they are relevant to each other. So, if P(X|H.K) $=$ 1 is the criterion for absolute relevance then P(X|H.K) $=$ P(X|K) is the criterion for absolute irrelevance. But the major problem of this criterion is that it does not incorporate the role of E in determining the irrelevance of X. By incorporating the role of E we can formulate the definition as follows: P (X|H.E.K) $=$ P(X|H.K)

I.5.4. Testing the Criterion of Irrelevant Conjunct

I claim that P (X|H.E.K) $=$ P(X|H.K) should be preferred over the Fitelson - Hawthorne condition as only the former can be defended as a sufficient condition of irrelevance. Intuitively it seems to me that both formulas are necessary conditions for irrelevance. And what we need to check is that whether these are individually sufficient conditions of irrelevance. For that we have to check whether relevant conjunctions satisfy the above formula. For checking each formula we are invoking different categories of evidence like confirming, disconfirming, neutral, conclusive, and non-conclusive.

1. In the case of disconfirming evidence (which is conclusive)

E.g. H: All ravens are black, E: A set of raven is white, X: Genetic structure of raven makes it black.

$P(E|H.X.K) = P(E|H.K) = 0$. H.X is a relevant conjunction, but still Fitelson and Hawthorne's definition for irrelevance is satisfied. But $P(X|H.E.K) \neq P(X|H.K)$. That is the relevant conjunction does not satisfy the modified version of Rosenkrantz's formula.

2. In the case of deductive evidence.

E.g. H: All ravens are black, E: A set of raven is black, X: Genetic structure of raven makes it black.

$P(E|H.X.K) = P(E|H.K) = 1$. In this example also, H.X is a relevant conjunction, but still Fitelson and Hawthorne's condition for irrelevance is satisfied but $P(X|H.E.K) \neq P(X|H.K)$.

I.6. New Formulation of the Paradox Based on the Modified Rosenkrantz's Condition

<u>Theorem 6</u>

If E confirms H and X is irrelevant to the confirmation relation between H and E (i.e. $P(X|H.E) = P(X|H)$ then E also confirms H.X (relative to K).

That is, if X is irrelevant to the confirmation relation between H and E and

$P(H|E.K) > P(H|K)$ then $P(H.X|E.K) > P(H.X|K)$

<u>Proof</u>

1. $P(H.X|E) = P(X|H.E) \times P(H|E)$
2. $\quad\quad = P(X|H) \times P(H|E)$: Irrelevance of X.
3. $P(X|H) \times P(H|E) > P(H) \times P(X|H)$: in the case where E confirms H
4. $P(H.X|E) > P(H.X)$: Definition of Conditional Probability

   That is H.X is confirmed by E.

I.7. Bayesian Solution to the Paradox: Softening the Impact of Paradox

I.7.1. Fitelson's Account

One of the major steps of Bayesian reasoning in resolving the paradox is the softening theorem. Fitelson claims that though irrelevant conjunction is represented as confirmed, BCT has the richness or tools to soften the impact of that irrelevant conjunction. (Fitelson "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction", 617). And it is an elaborative version of Earman's solution to the paradox. Fitelson's softening theorem is as follows:

<u>Theorem 7</u>

"If E confirms H, and X is confirmationally irrelevant to H, E and H.E (relative to K) then (provided that $P(X|K) < 1$) : $c(H, E|K) > c(H.X, E|K)$ where 'c' (confirmation measure) may be either difference measure 'd' or the likelihood ratio measure '*l*' of degree of confirmation. (But, c may not be the ratio measure r, since in cases of irrelevant conjunction we will have $r(H, E|K) = r(H\&X, E|K)$." (617)

Fitelson's claim is that the theorem softens the impact of the paradox. That is adding an irrelevant conjunct makes a conjunction only less confirmatory than its original hypothesis. As more and more irrelevant conjuncts are added the degree to which E confirms the conjunction will tend to decrease. The important question would be how it softens the impact of paradox. If we compare the Bayesian result with other theories of confirmation, certainly the theorem provides an improved situation. In the H-D model, there is no difference between the confirmation of H and H.X. But the BCT can distinguish between the confirmation of H and H.X as both are not equally confirmed hypotheses.

I.7.2. Ratio Measure

Though, Fitelson's softening theorem is an improvement, it is too little to claim that it softens the impact of paradox. As per the theorem, due to the irrelevant conjunct, the confirmatory strength of H.X becomes less than the confirmatory strength of H. That is irrelevant conjunct is visualised as one which is able to change the confirmatory relation. Such a scenario is counterproductive to the claim of softening attitude because a conjunct which is intuitively ineffective in affecting the confirmation relation is represented as one reducing the confirmatory strength. That is through softening theorem Bayesians are only asserting the point of paradox. The general idea of the claim of softening might be the following one.

1. Theorem pinpoints the role of irrelevant conjunct in a confirmation relation.
2. Characterizing the role of irrelevant conjunct as opposed to the role of relevant conjunct.

Certainly the softening theorem satisfies the first condition. It is able to pinpoint he role of irrelevant conjunct whereas it is not the case with H-D model but fails to characterize the role as opposed to the role of relevant conjunct. Decreasing confirmatory strength is one of the roles of relevant conjunct. If a relevant conjunct is added to a confirmatory relation certainly it would affect the confirmatory relation in a positive or negative manner. To modify the softening theorem, it is significant to explore the role of irrelevant conjunct. The role of a relevant conjunct is being remained as incapable of making any changes in the confirmatory relation. Changing the confirmatory relation precisely means changing the degree of confirmation. That is irrelevant conjunct should be represented as one which is unable to increase or decrease the degree of confirmation. That is, confirmatory strength of the irrelevant conjunction must remain same as of the original hypothesis. One can see that that is exactly what happens in the case of *r* measure: r(H.X,E) = r(H,E).

But ironically Fitelson rejects it as incapable of exhibiting the softening impact of the BCT, precisely because of this characteristic of r measure. My claim is Fitelson's conception of softening of paradox is inappropriate. Proponents of the above theorem may criticize that the new softening theorem makes H and H.X equally confirmed. Indeed the confirmation of H.X is a matter of worry but the matter does not worsen if the degree of confirmation of H and H.X are the same. Another kind of criticism would be that the new softening theorem is based on the ratio measure of confirmation and the ratio measure is a less preferred one when compare with the *l* measure. I am not claiming that the ratio measure surmounts all difficulties but the *l* measure also has many counter intuitive features though it is the most preferred one in the contemporary discussion. Following are the some of counter intuitive results of *l* measure:

1. l measure fails to capture the degree of confirmation in the case of conclusive confirmation. (where E entails H):

H: At least one coin turns out to be head (In the case of tossing of two coins)

E: Both coins turn up to be head

It is intuitively clear that H is conclusively confirmed, i.e. P (H|E) = 1. But the measure $l\left\{\frac{P(E|H)}{P(E|\sim H)}\right\}$ fails to capture such an intuitive notion as $\frac{P(E|H)}{P(E|\sim H)} = \frac{1/3}{0}$, which is indefinable.

2. l measure fails to capture the degree of confirmation in the case of conclusive disconfirmation:

   H: At least one coin lands head. E: both coins turn up to be tail.

   Here H is conclusively disconfirmed. $\frac{P(E|H)}{P(E|\sim H)} = \frac{0}{1/3}$ = 0.

   Since P (H) ≠ 0, degree of disconfirmation cannot be zero. Moreover, *l* measure is defined as log $\left[\frac{P(E|H)}{P(E|\sim H)}\right]$ and Log (0) is indeterminate.

My point is that the characteristic which is exhibited through the r measure is the stepping stone in solving the paradox. Of course, there could be many other characteristics of irrelevant conjunct which demarcate it from a relevant conjunct. So it appears that there is something more compelling in the softening theorem which Fitelson propounded but it needs to be more precise. It says H.X is less confirmed than H. Intuitively we can say that even the confirmatory strength of H.X is due to H. So adding X basically reduces the strength otherwise H would have. So we can say that by adding X to H, X makes the confirmation of H less significant. All confirmations are not significant confirmations. That's why we tend to ignore certain piece of information as crucial evidence though it raises the probability of H. To be significant evidence we intuitively demand that there should be a substantial difference between the prior probability of H and the posterior probability of H. I think the Bayesians (Earman, Rosenkrantz, and Fitelson) were trying to bring the notion of significant confirmation by claiming that H.X is less confirmed. By saying that adding an irrelevant conjunct makes confirmation less significant, I mean that adding an irrelevant conjunct makes the prior probability and the posterior probability closer than before. That is, when a large number of irrelevant conjuncts are added, it becomes too difficult to distinguish between the prior probability and the posterior probability. It would be difficult for one to see any confirmation relation between the irrelevant conjunction and its evidence. Such an idea cannot be represented by saying it has reduced the confirmatory strength. It is because reducing of the confirmation strength does not necessarily mean confirmation has become insignificant. Addition of countering or contrary position too reduces the degree of confirmation of conjunction in comparison with hypothesis H, but countering positions are not insignificant statements. Consider that degree of confirmation of H as determined by difference measure, d(H,E), is 0.6 and the degree of confirmation of conjunction, d(H.X,E) is −0.4; here also the confirmatory strength is reduced drastically, but we cannot say that the added conjunct spoils the confirmation or that it makes the confirmation insignificant. It is because though the degree is reduced it widens the gap between the posterior probability and the prior probability. That is it made the confirmation more clear and vivid. Though some relevant conjunct (both positive and negative) can make confirmation obscure, it would not be the necessary result of the relevant conjunct. But the added irrelevant conjunct can never widen the gap between the posterior and the prior probability; it can only make the confirmation obscure. That is, an irrelevant conjunct can never change the direction of confirmation. That is, if H is positively relevant to E, the added irrelevant conjunct can never make H.X negatively relevant to E. X can only make the degree of confirmation near to zero. The Bayesian softening theorem can therefore be modified as follows:

If E confirms H, and X is confirmationally irrelevant to H, E and H.E (relative to K) then (provided that $P(X|K) < 1$) then the absolute value of the degree of confirmation of H.X by E would be less than the absolute value of the degree of confirmation of H by E. My previous criticism to the original theorem would not be applicable here. Relevant conjunct as a category does not have the above characteristic.

I consider the nature of irrelevant conjunct exhibited through ratio measure vital in solving the paradox. $r(H.X,E) = r(H,E)$ reveals that irrelevant conjuncts are incapable of affecting a confirmatory relation. They remain as ineffective. But then the major problem is though r measure exhibits that, in an irrelevant conjunction, irrelevant conjuncts do not play any role, the irrelevant conjunction as a whole is projected as confirmed one. BCT is successful in showing that irrelevant conjuncts are inactive in the confirmatory relation. Till then it is precisely a success. But an irrelevant conjunction carries the same confirmatory power as that of the original hypothesis and as a consequence of that irrelevant conjunction would be considered as confirmed.

I.8. Solution to the Paradox

Though Bayesians made an improvement by defining what an irrelevant conjunct is and formulating a theorem which softens the impact of the paradox, still they are unable to resolve the paradox as irrelevant conjunction is still considered as confirmed which is counter intuitive. Bayesians like Maher points out that attempts to resolve the paradox ought to clarify why confirmation of Irrelevant Conjunction (I.C.) is paradoxical.

I.8.1. Clarifying the Paradoxicality

The Bayesian approach says that confirmation of I.C. is paradoxical. But it does not explore why it is paradoxical or to what extent it is paradoxical. One source of the paradoxical impression might be the similarity between being true and being confirmed. In deductive logic H.X can be true if and only if H is true and X is true. Such a conception may suggest that confirmation of H.X by E minimally means that H is confirmed and X is confirmed. Confirmation is understood as H being more probable than it was. It can be seen that H.X is more probable does not mean that both conjuncts are more probable than its priors. If one conjuncts' probability is increased it is sufficient to say that the conjunction is more probable. It can be shown mathematically and it is not counter intuitive too. Even if X is relevant, in the case of increased probability of H.X, both H and X need not be more probable than its priors. Then the significant question to be asked is whether increase of probability adequately represents confirmation. That is whether the confirmation of a conjunction in the intuitive level (explicandum) demands that both its conjuncts should be confirmed. It is not clear whether explicandum demands it. I think it is inadequate to demand that confirmation of H.X means E confirms H and X separately. It would not be the case always as in the case of many of relevant conjunctions auxiliary hypothesis would be about reliability of instruments which produced the evidence and evidence hardly confirms such auxiliary hypothesis. Certainly there could be further argument on this. But I am pursuing an idea that the paradoxical impression should be explained by some other factor than the demand that each conjunct should be confirmed.

I think the paradoxical impression might be the notion that what is in irrelevant conjunction to be tested apart from the testing of one of its conjuncts. That is we are often identifying the irrelevant conjunction with non-testable hypothesis. Such a notion is strengthened by the 'standard' examples of I.C. discussed in the literature. Strevens gives the following example for irrelevant conjunction: "All ravens are black and the Pope is infallible" (Strevens, Notes on *Bayesian* Confirmation Theory, 79). Then it seems that its confirmation by the observation of a black raven is an odd one. But the oddness of the confirmation does not lie on the point that a particular evidence E confirms H.X, but even if it is the confirmation by any evidence (i.e E or E1 or E2...) the confirmation sounds as odd. That means irrelevant conjunction is considered as non-confirmable or non-testable by a simple particular statement expressing the evidence. So something which is non-testable is considered as testable or positively tested is the reason for the paradoxical impression.

But what is the basis of considering I.C. as non-testable? I think our conception of testability of a conjunction is the one which is prevalent in the testing of relevant conjunction. If we adopt that conception as a criterion for testability then what we are claiming at the end is that I.C. is non-testable because it is I.C. So we need to thoroughly check that whether our conception of testability which is borrowed from the realm of relevant conjunction is adequate. In the realm of relevant conjunctions actually what we are testing is the truth of a single proposition. In the testing of a conjunction we are actually testing a new single proposition which can be derived from the conjunction by using all its conjuncts or some of the conjuncts itself which are confirmed by other conjuncts. So in conjunction, what is being tested is one single proposition. That single proposition could be the new one which is derived from conjunction or one of the conjuncts itself confirmed by all other conjuncts. I think the following formula can be considered as the criterion of testability of relevant conjunctions. There is at least one D such that if $H.X \vDash D$, $P(D|H.E) \neq$

$P(D|E)$ and $P(D|X.E) \neq P(D|E)$ then the conjunction is a testable one.[13] And as per the above criterion I.C. is not testable.

I think there is no adequate reason to consider the above criterion or any other similar criterion which is prevalent in the realm of relevant conjunction as a general criterion of testability. Notion of testability which is prevalent in the relevant relation is inadequate to adopt as a general criterion of testability because, in such cases, testability of a proposition 'X' in relation with E is not independently defined from relevance relations (Probabilistic- conditional independence). In the above definition, the concept of testability is defined as equivalent to the concept of relevance.

One can conclude that concept of testability and the concept of relevance are equivalent concepts or equivalent relations. But to claim the notion of equivalence between the both concepts, definitions of each concepts should be justified on independent reasons. While the above definition considers relevance and testability as equivalent relations without providing independent reasons for each definitions. In other words, the claim of equivalence between two relations is a circular one. Moreover, it is hard to the claim that the definition of testability is adequate even in the realm of relevance relation because even in the case of relevant conjunctions, consequence of the conjunction need not be relevant to each conjunct. Suppose, in relevant conjunction one conjunct is auxiliary hypothesis, in most cases, consequence of conjunction would not be relevant to the consequence in isolation.

[13] H.X: Conjunction, E: Evidence, and D: consequence of H.X

I.8.2. Maher's Solution

Our common notion of testability is crucially challenged by the Maher's example. Maher's position is that there is nothing paradoxical in that I.C is confirmed. Therefore the assumption (*) that 'If *X* is irrelevant to *H* and *E* [given *K*] then *E* does not confirm *H.X* [given *K*]' is false. Maher argues as follows:

> Here is a numerical example: Suppose *K* tells us that a coin is fair, let *H* be that the coin lands heads on the first toss, and let *X* be that it lands heads on the second toss. Thus *X* is irrelevant to *H* given *K*. This is a case where inductive probabilities have precise quantitative values; the inductive probability of *H.X* is 1/4 given just *K* and is 1/2 given *H* and *K*, so *H* r-confirms *H.X* given *K*, contrary to *. The observation that * is false solves the problem of irrelevant conjunction without either of the drawbacks of previous Bayesian solutions. (Maher, "Bayesianism and Irrelevant Conjunction", 519)

Though Maher claims that he has solved the paradox, his solution does not evoke much debate or interest. I think his account lacks sufficient argument to claim that *(the intuitive notion) is false. Maher only shows that for a particular explicatum (incremental confirmation) it is quite adequate to consider that the irrelevant conjunction is confirmed. But he does not show that for the explicandum, the confirmation of I.C. is adequate.

Maher's solution might be an inconclusive one but has a powerful insight to crack the paradox. I think the only way we can maintain the I.C.'s confirmation is paradoxical, is by holding that the I.C. is non-testable. What is interesting in Maher's account is that his example of I.C. challenges the notion that I.C. is non-testable. But still we are uncomfortable with the confirmation of I.C. because there are many other examples (like Strevens) of I.C., whose confirmation arouses a sense of oddness. Some may wonder whether there is any real testing in the case of Maher's example. I think we need to explore more about what is non-testability. Maher's account is significant because it asks us to come up with a weaker notion of testability.

Certainly background information (hypotheses which are certain) is considered as non-testable in a unanimous way. This can be used to construe what is minimal in saying that something is non-testable. In our account, we do not represent the hypothesis to be tested as H.K but only as H|K. because K is not being tested. K as background information is known as certain. So there is no point in saying that K is tested along with H. But one can show that P (H.K|E) = P (H|K.E)

$$P(H.K|E) = P(K|E) \times P(H|K.E)$$

$$P(H.K|E) = P(H|K.E) \qquad : (P(K|E) = 1)$$

Though P (H.K|E) = P (H|K.E) we tend not to say that what is being tested is K also. Here what prevents us from holding such a position? The reason why we do not use H.K is that K does not offer anything to be tested. In its crude form the minimal criterion to be testable is that hypothesis offers anything new to be tested. How can this 'new' be defined. Usually it is defined in relation with background information. But H.K obviously provides something new to be tested in comparison with this background information K. But comparing with another competing hypothesis H or H|K, the hypothesis H.K does not provide anything new to be tested. So the simpler hypothesis H should be preferred for testing than H.K. Testability is decided on the basis of anything new which is offered/ provided by hypothesis in relation with background information and competing hypotheses.

I.8.3. Confirmability/Paradoxicality Depends on Context

My point is that there are cases in which we can say that conjunction of two probabilistically independent statements are tested. In a sense conjunction of any two statements can be tested. My claim is that in general neither I.C. is non- testable nor I.C. is testable (as Maher claims). That is in certain particular contexts, I.C. can be considered as testable, where as in other context it remains as non-testable. Here I am specifying/representing context in terms of competing hypotheses. In the context of certain competing hypotheses, if I.C. cannot offer anything to be tested other than what is offered by its competitors then the testing of I.C. is trivial. So I am keeping a middle path between Maher's position which states that I.C. is testable and traditional Bayesian position that I.C. is not testable. My position is that I.C. is testable but only in certain contexts[14].

Through the following paradoxes, I try to argue that why we should take side neither with Maher nor with the traditional Bayesians. I think the following reformulation of the paradox shows that the traditional Bayesian assumption that I.C. cannot be confirmed is false.

Paradox 1:

1. An irrelevant conjunction is not supposed to be confirmed
2. I.C. cannot be confirmed entails that H cannot be confirmed
3. But H can be confirmed.

[14] Such an approach includes the context of discovery into logic of confirmation but in very a minimal sense.

I think the premises 2 and 3 are firmer than the first premise. We can solve the paradox only by rejecting the first premise (the traditional Bayesian position)

I claim that Maher's position is also incorrect.

Paradox 2:

1. If all possible (not merely actual) evidence confirms/disconfirms H and H_1 to the same degree then $H \equiv H_1$
2. All evidence confirms H and H.X to the same degree.[15]
3. But $H \not\equiv H.X$

Maher says that there is nothing unintuitive about the confirmation of H.X where X is an irrelevant conjunct. If the confirmation of H.X is accepted as correct then one is bound to accept its corollary that all evidence confirms H and H.X to the same degree. So the second premise is the consequence of Maher's position. Though I do not have any argument for first premise, I consider it as right. I think the only way to solve the paradox is to accept that second premise is wrong that is Maher's position is wrong. That the confirmation of a hypothesis depends upon the context of confirmation (competing hypothesis) cannot be captured by the traditional Bayesian idea of confirmation. So the notion of confirmation has to be redefined.

[15] This is true at least as per ratio measure. And unlike Fitelson, Maher does not take a position which rejects one measure in favour of other. Therefore, the implication of ratio measure cannot be ignored in Maher' position.

I.9. A New Approach to Confirmation

I think the discussion of the paradox leads to a more general position regarding the confirmation. The claim is that confirmation of H can be determined only in relation with a best competing hypothesis of H.

Carnap distinguished incremental confirmation to three categories.

1. Classificatory incremental confirmation: C(H,E) > 0
2. Comparative incremental confirmation: C(H,E) > C(H_1, E)
3. Quantitative incremental confirmation : C(H,E) = w

I think Carnap's (and the traditional Bayesian's) account of classificatory confirmation is inadequate. Classificatory account itself should include comparative aspect. So I am defining the categories as follows:

1. Classificatory confirmation of H by E: C(H,E) > C(H_c, E)

   i.e. $\left\{ \frac{P(H|E.K)}{P(H|K)} > \frac{P(Hc|E.K)}{P(Hc|K)} \right\}$ (where H_c is a best competitor)

2. Comparative incremental confirmation: E confirms H more than H_1

   $\left\{ \frac{P(H|E.K)}{P(H|K)} - \frac{P(Hc|E.K)}{P(Hc|K)} \right\} > \left\{ \frac{P(H1|E.K)}{P(H1|K)} - \frac{P(H1c|E.K)}{P(H1c|K)} \right\}$ (where H_c is a best competitor of H and H_{1c} is a best competing hypothesis of H_1)

3. Quantitative incremental confirmation: C(H,E) − C(H_c, E) = w .

   i.e $\left\{ \frac{P(H|E.K)}{P(H|K)} - \frac{P(Hc|E.K)}{P(Hc|K)} = w \right\}$

The concern of the new approach can be addressed even within the framework of the traditional Bayesian framework also. In his paper "Logical versus Historical Theories of Confirmation" Musgrave raised a similar concern in his attempt to solve the raven paradox. His suggestion is that background information should be understood as background theory and he suggests that background theory for a new theory must be the best available theory actually present in the field ("Logical versus Historical Theories of Confirmation" 17-18). I think that though such constraints on background information are intuitively appealing, they are not compatible with the subjective Bayesian framework. Moreover, according to Musgrave, for all contrary hypotheses, best competitor is a single particular theory. But I think best competing theory varies from theory to theory and the challenge is to formalize the notion of best competitor.

My definition of the best competitor is as follows:

H_1 is the best competitor to H regarding evidence E iff

then $\mid P(E|H.K) - P(E|H1.K) \mid < \mid P(E|H.K) - P(E|Hi.K) \mid$ [16]; (where 'i' is any competing hypotheses in the field of inquiry).

The best competitor is the notion which is defined in relation with particular evidence. And the best competitor for H may vary in the context of varying evidence. So, the best competitor is a notion which is defined in a context which is posterior to evidence.

And in the new Bayesian frame work, paradox of irrelevant conjunction does not arise at all.

[16] Absolute value of $P(E|H.K) - P(E|H1.K)$

And as per the definition of the best competing hypothesis, the best competitor of H.X is H. So H.X would be considered as confirmed only if $\left\{ \frac{P(H.X|E.K)}{P(H.X|K)} > \frac{P(H|E.K)}{P(H|K)} \right\}$. It can be shown that it is not the case and L.H.S and R.H.S are equal. Therefore as per new definition, E is a neutral evidence to irrelevant conjunction H.X.

<u>Theorem 8</u>

If X is irrelevant to the confirmation relation between H and E (i.e, P(X|H.E.K) = P(X|H.K) then H.X is not confirmed by E (in new Bayesian confirmation) in relation with its best competitor, H, i.e.

$$\frac{P(H.X|E.K)}{P(H.X|K)} = \frac{P(H|E.K)}{P(H|K)}$$

<u>Proof:</u>

Assume that K is tautologous.

$$\frac{P(H.X|E)}{P(H.X)} = \frac{P(H|E) \times P(X|H.E)}{P(X|H) \times P(H)}$$

$$= \frac{P(H|E) \times P(X|H)}{P(X|H) \times P(H)}$$: Since X is confirmationally irrelevant to H and E.

$$\frac{P(H.X|E)}{P(H.X)} = \frac{P(H|E)}{P(H)}$$

This proof is based on modified Rosenkrantz's condition of irrelevance, but the same result can be proved by using Fitelson - Hawthorne condition (P(E|H.E) = P(E|H))

I.9.1. *l* Measure and the New Bayesian Approach

Introduction of competing hypotheses into the realm of confirmation is not a new one in the Bayesian confirmation debate. The most defended measure of confirmation in the Bayesian literature is the l measure: $\frac{P(E|H.K)}{P(E|{\sim}H.K)}$.

$$\frac{P(E|H.K)}{P(E|{\sim}H.K)} = \frac{P(H.|E.K)/\,P(H.|K)}{P({\sim}H.|E.K)/\,P({\sim}H.|K)}$$

That means, according to the l measure, confirmation of H is not merely the contrast between its prior probability and posterior probability. According to the l measure only if the ratio of posterior and prior probability of H is higher than the ratio of posterior and prior of its all competitors, then H is confirmed. The new approach basically only amends the l measure, H's ratio need not be higher than the ratio of all competing hypothesis but only higher than the ratio (ratio of posterior and prior probability) of its best competitor. Amendment is required as the l measure fails to resolve the paradox and also it has many counter intuitive features.

And it seems that the new Bayesian approach dismisses and clarifies many of the criticisms which the traditional Bayesians faced. According to the traditional Bayesian approach, confirmation means that the posterior probability is higher than the prior probability. Its counter examples are mainly in two categories

1. The posterior probability of H is higher than its prior probability but intuitively there is no confirmation.

For example "When Mark Spitz goes swimming he increases the probability that he will drown but the fact that he is swimming is not evidence that he will drown" (Achinstein 1983

152). I think in the analysis of this example it becomes clear that notion of confirmation of a hypothesis is too ambiguous without referring to its competitor. Consider H: He will drown and H_1: He will not drown. I don't think that H's confirmation is counter intuitive in relation to H_1. We often feel it is counter intuitive because we consider H_2 as its competing hypothesis. H_2: He may drown. As swimming in a pool provides only less chance for drowning (for a person who knows swimming), H_2 is more confirmed (in traditional sense) than H. So in the new Bayesian approach, in relation with H_2, H is not confirmed. $\left\{ \frac{P(H|E.K)}{P(H.|K)} < \frac{P(H2|E.K)}{P(H2|K)} \right\}$

2. Second kind of counter example is that posterior probability is not increased than its prior but still H is confirmed.

> To see this, consider an agent who is wondering whether deer live in a nearby wood. He comes across a pile of deer droppings, and his confidence in the deer hypothesis increases to near 1. Shortly thereafter, he finds a shed antler. Since his confidence in the deer hypothesis is already so high, this new evidence does not have any significant impact on it. Now, subsequent to the agent's finding the second piece of evidence, let us ask whether the evidence about the droppings and the antler provide rational support for our agent's belief about deer. Intuitively, the droppings and the antler provide equally strong rational support for the agent's deer belief. There is, I think, no sense in which the droppings currently provide stronger support or better evidence than does the antler. But only the droppings are historically associated with a significant increase in the agent's probability for deer. The historical approach thus makes contemporary evidential support depend in an unintuitive way on the order in which evidence was discovered. Intuitively, synchronic support should depend on the agent's present epistemic state, not on such historical accidents. (Christensen, "Measuring Confirmation", 444 - 445)

Consider H: A deer lives in a nearby wood and H_1 : A deer does not live in a nearby wood

E_1: A pile of deer droppings and E_2: Shed antler.

According to the new Bayesian approach E_2 would be evidence as

$\frac{P(H|E1.E2.K)}{P(H|E1.K)} > \frac{P(H1|E1.E2K)}{P(H1|E1.K)}$. L.H.S $\approx$ 1 as prior and posterior probabilities of H is almost same. But $P(H1|E1.E2K) < P(H1|E1.K)$ as E_2 further disconfirms H_1 and disconfirmation of H_1 by E_1 is not conclusive also. Therefore R.H.S would be much less than 1.

The prominent criticism against the new approach would be that in this account we cannot say that H is confirmed or disconfirmed. The claim can be made always only in relation to a competing hypothesis, which is counter intuitive. My claim is that in the practicing realm when we made the claim the H is confirmed there is an implicit reference to the best competing theory. The exciting point is that the new account comes closer to Lipton's account of confirmation. In Lipton's account, what is being explained or confirmed is not merely 'Why this' but 'Why this rather than that'. (Inference to the Best Explanation, 33-35) Lipton's account is also charged with relativism but he has some good replies to it. (Inference to the Best Explanation, 72)

I.9.2. Measure of Confirmation in the new Bayesian Approach

The following three measures of confirmation shows that I.C. is not confirmed where H has a competing hypothesis. But I prefer new difference measure (comparative difference measure) because the other two measures yield some counter intuitive results.

1. Comparative difference measure: $\frac{P(H.|E.K)}{P(H.|K)} - \frac{P(H1|E.K)}{P(H1.|K)}$ i.e. $\frac{\{P(E|H.K) - P(E|H1.K)\}}{P(E)}$

2. Comparative ratio measure: $\frac{\frac{P(H.|E.K)}{P(H.|K)}}{\frac{P(H1|E.K)}{P(H1.|K)}}$ i.e. $\left\{\frac{P(E|H.K)}{P(E|H1.K)}\right\}$

3. Comparative rate of Increase : $\dfrac{\frac{P(H.|E.K)}{P(H.|K)} - 1}{\frac{P(H1|E.K)}{P(H1.|K)} - 1}$

The counter intuitive results are the following ones:

1. The measures (comparative ratio and comparative rate of increase) fail to capture the confirmation by deductive evidence in certain contexts. E.g. H: All ravens are black H_1: No ravens are black, E: A black raven. $\frac{P(E|H.K)}{P(E|H1.K)} = \frac{1}{0}$ which is indeterminate and the measure rate of increase yields a negative value which is also counter intuitive.

2. In the case of conclusive disconfirmation comparative ratio measure says that confirmatory strength is 0. In the case of comparative ratio measure, it is unclear as to what does it mean to say that confirmatory strength is 0. In the case of conclusive disconfirmation, posterior probability and confirmatory strength are same (i.e. zero) which is counter intuitive. But the remarkable point is that two major measures of traditional confirmation (r and *l*) also faced the second difficulty.

Conclusion

All attempts of the H-D model's proponents and the major attempts of contemporary Bayesians primarily relied on the presumption that confirmation of I.C. is counter intuitive. Maher challenges such a position and suggests that any attempt to resolve the paradox ought to clarify why such confirmation is paradoxical. I think there is no good reason to hold that confirmation of I.C. is absolutely paradoxical or not paradoxical. The impression of the paradoxicality depends upon the context of testing. It seems that the context of testing can be explicated through the notion of 'best competing hypothesis'. Inquiries on this direction are compelling because it seems that such an approach will do well in solving various other problems of the BCT.

Chapter II

Hempel's Paradox: An Analysis

Introduction

Carl G. Hempel formulated the paradox as a derivation of the paradoxical / false conclusion from two plausible conditions of confirmation. That is, Jean Nicod's Criterion and Equivalence condition entail that 'A non- black non-raven confirms the hypothesis all ravens are black'.

- Nicod's Condition (N.C): For any object 'a' and any pair of predicates F and G, the proposition that 'a' has both F and G confirms the proposition that every F has G.

(Fa·Ga) confirms (x)(Fx ⊃ Gx), for any individual term 'a' and any pair of predicates 'F' and 'G'.

- Equivalence Condition (E.C): Evidence which confirms/disconfirms hypothesis also confirms/disconfirms its logically equivalent hypotheses.

From (N.C) and (E.C), we can deduce the following, conclusion:

- Conclusion (Con): The proposition that a is both non-black and a non-raven,

(∼Ba · ∼Ra), confirms the proposition that every raven is black, (x)(Rx ⊃ Bx).

In the first section of my chapter I analyse plausibility and acceptance of the premises from which the paradoxical conclusion is derived. From second section onwards I discuss various suggested solutions to the paradox. And in the final section, I discuss the solution to the paradox which I consider as a synthesis of Hempel's and Bayesian approaches.

II.1. Plausibility of the Premises and the Paradoxicality of the Conclusion

The paradoxical conclusion is unavoidable because it is the consequence of principles which are hard to reject. According to Hempel, though Nicod's Criterion cannot be maintained as a necessary condition of confirmation, it can be maintained as a sufficient condition of confirmation. That is, instances confirm the hypothesis although it is not necessary that all confirming evidence are instances. For example, a black raven is an instance of the hypothesis 'All ravens are Black'. But an instance is not the sole category of evidence; there are other kinds of evidence which are not instances of the hypothesis; for example, the black feather of a raven is a confirming evidence of the hypothesis, but it is not an instance. Hempel maintains that E.C. is an adequacy condition of confirmation because confirmation of a hypothesis by evidence should not be dependent upon the way it is formulated. Hempel's defence of E.C. as necessary condition of confirmation is as follows: "Otherwise, the question as to- whether certain data confirm a given hypothesis would have to be-answered by saying 'That depends on which of the different equivalent formulations of the hypothesis is considered '-which appears absurd" ("Studies in the Logic of Confirmation", 21). Hempel's point might be that hypothesis is all about a logical relation between simple statements or between the predicates no matter how the logical relation is expressed or formulated. Apart from the logical nature of the proposition/hypothesis, other factors such as emphasis of the proposition, apparent notion of what the hypothesis talks about are insignificant regarding confirmation. Logical nature means how simple statements or predicates are related to each other. Usually it is either implication or conjunction. So, in principle, equivalence condition cannot be denied. But the question can be asked whether classical logic is able to capture the logical relation correctly.[17] But if those concerns are set aside, equivalence condition can be held as an adequacy condition of confirmation.

The paradox is that these principles entail that a sentence expressing the proposition 'a non- black non- raven' confirms the sentence expressing the proposition that 'all ravens are black'. As per Nicod's Criterion (N.C), an object which is non- black non-raven confirms the hypothesis that all non-black things are non-raven. The sentence 'All non-black things are non-raven' is logically equivalent to the hypothesis expressed by the sentence 'all ravens are black'. As per the equivalence condition the confirming evidence of a hypothesis should also confirm its logically equivalent description and vice- versa. That is, a non-black non- raven confirms the hypothesis 'all ravens are black'. This conclusion is found to be paradoxical. A non- black non-raven is an object which is neither black nor raven, e.g. a red pencil, or a white shoe. It is counter intuitive to assume that a red pencil confirms the hypothesis that 'all ravens are black'.

There are many reasons why such conclusion seems absurd or counterintuitive because, the same evidence 'red pencil' can confirm (as per E.C & N.C) all contrary hypotheses of H, except one, that is, all ravens are red. And the oddness of the conclusion would be vivid in a quantitative frame work: it seems that red pencil confirms all contrary hypotheses and original hypothesis to the same degree. A statement which confirms almost all contrary hypotheses can hardly be considered as evidence. Intuitively, we assign a stronger relation between a hypothesis and its evidence. And another reason is that the

[17] There are inquiries in that direction: Ken Gemes "Carnap-Confirmation, Content-Cutting, and Real Confirmation".

contra positive instances do not disconfirm any of the contrary hypotheses. Our intuitive notion of confirmation demands that evidence should disconfirm at least some of the contrary hypotheses. That is two plausible principles entailing a counter-intuitive statement is what Hempel's paradox is all about.

II.2. Attempted Solutions of the Paradox

The attempts to solve the paradox can be largely divided into two categories.

A. The first category of approach is the attempts to refute the premises which are used for the derivation of the paradoxical conclusion.

B. The second category of approach is the attempt to show that the paradoxical nature of the conclusion is only an impression and thus attempts to dissolve the paradox.

There are mainly three strands in the first category of solution.

A1: Identifying the premises or assumptions other than the N.C and the E.C which are involved in the derivation of the conclusion and exploring the inadequacies of those premises.

A2: Attempts to refute the Equivalence Condition.

A3: Attempts to refute Nicod's Criterion or modify it such that it would not allow the derivation of the paradoxical conclusion.

II.3. Exploring Hidden Premises

Hempel's primary attempts to dissolve the paradox involve identifying the other assumptions/premises, which are involved in the derivation of the paradoxical result. Hempel examines several possible solutions from the category of solutions, A1, but at the end rejects them as inadequate. Though Hempel dismisses the validity of alternative formulations, such approaches found resonance in later discussions of the paradox.

In Hempel's paper "Studies in the Logic of Confirmation"[18] itself, he examines the possibility of a hidden premise, which helps to derive the paradoxical conclusion. Such a move was mainly motivated by the strong plausibility of N.C and the E.C. Our intuition pulls us back from rejecting N.C and the E.C; then the obvious way out appears to the possibility of a hidden factor. Thus, he identified a third premise which, according to him, is necessary for the derivation of paradox. The instances of the equivalent hypothesis that are expected to confirm the original one is the basis of the paradox. What would be logically equivalent is determined by how the hypothesis is originally expressed. Thus, the third premise is about the form of expression of hypothesis. "The meaning of the empirical hypothesis can be adequately expressed by means of sentences of universal conditional form." (Hempel, "Studies in the Logic of Confirmation", 23) The third premise is that 'All Ravens are black' is adequately expressed by (x) (Rx⊃ Bx). And Hempel examines the validity of the third premise. He assumes that one major criticism would be that such an expression does not confer existential import to the hypothesis. And he says that:

> Universal conditional sentence, in the sense of modern logic, has no existential import; thus, the sentence '(x) (Mermaid(x) ⊃Green(x)) does not imply the existence of mermaids ; it merely asserts that any object is either not a mermaid at all, or a green mermaid; and it is true simply because of the fact that there are no mermaids. General laws and hypotheses in science, however-so it might be argued, are meant to have existential import; and one might attempt to express the latter by supplementing the customary universal conditional by an existential clause. Thus, the hypothesis that

[18] It is the first paper which formulates a theory of confirmation in a systematized manner. The task of the paper is to formulate certain adequacy conditions of confirmation. The paradox is formulated in that paper itself in a bid to clarify the conditions of confirmation. Since it is the paradox which is formulated in the first paper of theory of confirmation, undoubtedly it is the first paradox of confirmation.

> all ravens are black would be expressed by means of the sentence S1: ' (x) (Raven(x) ⊃ Black(x)). (Ex)Raven(x); and the hypothesis that no non-black things are ravens by S2: '(x)(∼Black(x) ⊃ ∼ Raven(x)) . (Ex)∼ Black(x)'. Clearly, these sentences are not equivalent. (24)

But Hempel defended the third premise and opposed the alternative formulation discussed above on the following grounds:

1. Many of the logical inferences which are generally accepted would be invalidated by the above formulation.
2. Contrapositive hypothesis is logically equivalent even according to common understanding and usage. Adding of existential cause nullifies it.

Moreover he strongly defends the third premise on the basis that customary formulation of general hypothesis in empirical science clearly does not contain an existential clause.

> Finally, many universal hypotheses cannot be said to imply an existential clause at all. Thus, it may happen that from a certain astrophysical theory a universal hypothesis is deduced concerning the character of the phenomena which would take place under certain specified extreme conditions. A hypothesis of this kind need not (and, as a rule, does not) imply that such extreme conditions ever were or will be realized; it has no existential import. (25)

Later Hempel examines other possible criticisms against the third premise. One of the possible criticisms would be that the hypothesis is about ravens and not about non-black things. That means hypothesis has a field of application and an indicator of field of application should be included in the representation of hypothesis. "..Thus, we might represent the hypothesis that all ravens are black by the sentence '(x) (Raven(x) ⊃Black(x)) ' plus the indication 'Class of ravens' characterizing the field of application; and we might then require that every confirming instance should belong to the field of application" (25). This procedure would exclude contrapositive instances and thus it dissolves the paradox.

But Hempel dismisses such alternative formulations for the representation of hypothesis. His main objection is the following:

> ...the consistent use of a domain of application in the formulation of general hypotheses would involve considerable logical complications, and yet would have no counterpart in the theoretical procedure of science, where hypotheses are subjected to various kinds of logical transformation and inference without any consideration that might be regarded as referring to changes in the fields of application. This method of meeting the paradoxes would therefore amount to dodging the problem by means of an ad hoc device which cannot be justified by reference to actual scientific procedure. (26)

Though Hempel dismisses the validity of alternative formulations such approaches found resonance in later discussions of the paradox. At the outset of the discussion, paradoxical conclusion can be derived from two premises (E.C&N.C). Hempel's exploration of hidden premises reformulates the derivation of paradoxical conclusion as follows:

1. The hypothesis 'All Ravens are Black" can be expressed as (x) (Rx ⊃ Bx)

2. Equivalence Condition

3. Nicod's Criterion

1-3 entail the conclusion: ∼Ra. ∼Ba confirms (x) (Rx⊃Bx).

II.4. Attempts to reject / modify Nicod's Criterion

Though both premises are well grounded, the premise N.C is found to be more vulnerable to attacks. One of the reasons for intensive criticism against N.C is its qualitative frame work.

The criterion is primarily formed within the qualitative frame work; therefore quantitative theories of confirmation do not find the criterion of any significant use. Though many principles or conditions are primarily qualitative, they play a significant role as conditions of confirmation even in the quantitative frame work (e.g. Equivalence condition). Since theories of confirmation can express N.C confirmation (instance confirmation) in their own terms[19], there is no need to maintain N.C as a criterion of confirmation. In a sense, confirmation theorists do not find any compulsion to maintain N.C in its own form. Though its ability as a useful criterion of confirmation is doubtful, there is no cogent reason to suspect its validity as a criterion. But obviously the strength of N.C criterion is its closeness to our intuition of confirmation. Instance of the hypothesis is the most widely used category of evidence in the cases of confirmation of universal hypotheses which deals with observable predicates. Moreover, often a confirmation relation is not obvious to agents and the confirmation relation between two propositions can be claimed only on the basis of strong arguments. But in the case of instance confirmation, confirmation relation is comparatively more obvious to agents and easily cognisable. An instance of hypothesis would be the one which comes to our mind promptly when we talk of evidence of the hypothesis. Critics may, of course, point out that N.C has a standing as an intuitive idea of confirmation, but it does not have any place as criterion in formal or general accounts of confirmation. And for various other reasons like impreciseness of definition, N.C faces constant challenge in the debates of confirmation.

[19] For example, Bayesian can express the instance confiramtion in terms of prior and posterior probability and H-D model can express instance as deductive consequences.

II.4.1.Quine's attempts to restrain Nicod's Criterion

Among the earlier attempts to reject Nicod's Criterion, the much noted one is suggested by W.V.O. Quine. Quine attempts to restrain Nicod's Criterion with certain conditions. His position is that there are problems in considering predicate like non-raven or non- black on par with predicates like raven and black. These are different kinds of predicates. According to Quine, 'black' 'raven' etc are natural kinds. And Nicod's Criterion is originally formulated to account for instantial confirmation of hypothesis which involves natural kinds. And there is no reason to assume Nicod's Criterion is applicable to non-natural kinds too. "Only instances falling under natural kinds can warrant instantial confirmation of universal laws" (Fitelson and Hawthorne, How Bayesian Confirmation Theory Handles the Paradox of the Ravens, 249). Thus, for Quine, the source of the problem is (N.C). He suggests that the unrestricted version of (N.C) is false, and must be replaced by a restricted version that applies only to natural kinds:

Quine−Nicod Condition (QNC): "For any object 'a' and any natural kinds F and G, the proposition that a has both F and G confirms the proposition that every F has G. More formally, (Fa·Ga) confirms (x)(Fx ⊃ Gx), for any individual term a, provided that the predicates 'F' and 'G' refer to natural kinds" (249).

It is interesting to note that Quine's solution not only restrict ∼Ra. ∼Ba from confirming (x) (Rx⊃Bx), but it also restrict ∼Ra. ∼Ba from confirming (x) (∼Bx⊃∼Rx). The major problem with Quine's approach is the following: it is true that confirmations by contra positive instances are often counter-intuitive. But it is unwise to state that no contrapositive instances confirm the hypothesis. Lipton, Rosenkrantz and many others provide examples where contrapositive instances confirm hypothesis more strongly than direct instances.[20] Quine's solution cannot account for such possibility.

II.4.2. Bayesian Rejection of Nicod's Criterion

Bayesians have various attempts to resolve the paradox of confirmation. Among the various solutions, major one is the rejection of Nicod's Criterion. They clarified that Nicod's Criterion is an imprecisely formulated one. There are forms of Nicod's Criterion which would not be even accepted by its proponents. According to them, to make it precise, N.C needs to be quantified. On the basis of quantification of Nicod's Criterion, there are four possible classifications:

1. $N.C_w$: Instances confirm the hypothesis, relative to some possible background knowledge.
2. $N.C_{\propto}$: Instances confirm the hypothesis relative to our actual background knowledge.
3. $N.C_T$: instances confirm the hypothesis relative to tautological background knowledge.
4. $N.C_S$: instances confirm the hypothesis relative to all possible background knowledge.

On the basis of this classification, Bayesians sought to know which of the N.C is exactly Hempel or the proponents of instantial confirmation had in mind. $N.C_w$ is too weak to be of much use. Unless, one could specify the conditions which restrict the valid application of N.C to some background information, $N.C_w$ is philosophically insignificant.

---

[20] Following is the Rosenkrantz's example: "Finding a rare infected subject who was not vaccinated in an experimental population provides stronger evidence that the vaccine in 100 per cent effective than finding a vaccinated subject who is uninfected." (Rosenkrantz, "Bayesian Confirmation: Paradise Regained", 468) Lipton says that "… some things that are neither rubbed nor heated do support the hypothesis that friction causes heat." (Lipton, *Inference to the Best Explanation*, 94)

In the same way $N.C_{\propto}$ is also philosophically insignificant unless it is specified that what are the conditions which distinguish actual background information from another possible background. So, philosophically significant versions of Nicod's Criterion are $N.C_T$ & $N.C_S$. But it is somewhat obvious to see that $N.C_s$ is too strong to hold. If the background information contains the knowledge that hypothesis is false, then no instances can support the hypothesis. Consider the hypothesis that 'H: All squares in the board are white' and consider the evidence that first square is white. But if we have background information that 'board is a chess board', then the instance that 'First Square is white' does not confirm the hypothesis.

Though, $N.C_s$ is philosophically significant, it is hard to hold it as a valid criterion. I. J. Good offers very good argument and a counter-example to show that $N.C_s$ is false. The following is Good's counter-example.

> Suppose that we know we are in one or other of two worlds, and the hypothesis, H, under consideration is that all the crows in our world are black. We know in advance that in one world there are a hundred black crows, no crows that are not black, and a million other birds; and that in the other world there are a thousand black crows, one white one, and a million other birds. A bird is selected equiprobably at random from all the birds in our world. It turns out to be a black crow. This is strong evidence that we are in the second world, wherein not all crows are black. Thus the observation of a black crow, in the circumstances described, undermines the hypothesis that all the crows in our world are black. Thus the initial premise of the paradox of confirmation is false, and no reference to the contrapositive is required. (Good, "The White Shoe is a Red Herring", 322)

Good's counter example intends to argue that $N.C_s$ is not a valid criterion. Indeed, $N.C_s$ cannot be defended as a valid form of instance confirmation, though it is doubtful that

Good's counter example was successful in arguing that point. $N.C_s$ cannot be defended because there could be cases where background information 'K' itself contains the information that H is false. However, the proponents of N.C. may easily defend such counter example by adding the constraint that 'K' does not contain the information that H is false. The challenge would be to formulate a counter example to $N.C_{(S-(\sim H))}$. '$N.C_{(S-(\sim H))}$' is the criterion which states that instance confirms the hypothesis relative to all possible background information except that the background information contains '∼H'. And it seems Good is attempting to counter $N.C_{(S-(\sim H))}$ since $N.C_s$ is obviously invalid.

In my opinion, Good's example does not pose any challenge to $N.C_{(S-(\sim H))}$. Good's position is that the hypothesis which is under consideration is (x) (Rx⊃Bx). And his argument is that instance (Ra. Ba) confirms ∼(x) (Rx ⊃Bx) and it disconfirms (x) (Rx ⊃Bx). From the narration of the counter example, we hold that H is (x) (Rx ⊃ Bx); then the other information is 'background knowledge'. Either of the two hypotheses ((x) (Rx ⊃ Bx) and ∼(x) (Rx ⊃ Bx)) are true: What is deceiving in Good's example is that, two kinds of background information are employed for two hypotheses. For (x) (Rx ⊃Bx) 'K' is 100 black raven, ∼ (Rx ⊃ ∼ Bx)(no non-black raven) and million other birds. And for the hypothesis ∼ (x) (Rx ⊃Bx), 'K' is 1000 black raven and one white raven and million other birds. Obviously Good is employing two different background information for two hypotheses and then attempting to compare the two hypotheses. What Good's counter-example establishes is hardly obvious. His example would be a counter example only if he chooses one background knowledge and states that, within that particular background information, E confirms H_2 more than H_1 even though E is the instance of H_1 and thus H_1 is disconfirmed since H_2 is contrary hypothesis of H_1. It is not clear how the confirmation which happens in two different background information can be compared. It is the same to say that E_1 confirms H_1 but E_2 confirms H_2. Therefore, E_2 disconfirms H_1. The comparison between two confirmations is possible only if 'K' and 'E' are same in both confirmation relations.

One could argue that my interpretation of Good's example is completely wrong, since the hypothesis which is under consideration is neither (x) (Rx ⊃Bx) nor ~(x) (Rx ⊃Bx). Though, Good says that the hypotheses are H_1: (x) (Rx ⊃Bx) and H_2:~(x) (Rx ⊃Bx), we could neglect that for a better interpretation. In the better interpretation, our background information K is that either of the world is true (actual); and the claim is that, observation of black raven raises the probability of the world which contains a non-black raven. In such an interpretation, of course, the example is a sound one but the problem again arises as what the evidence confirms is neither (x) (Rx ⊃Bx) nor ~ (x) (Rx ⊃Bx), but W_1(world 1) or W_2 (world 2). In this case, the hypotheses are different; they are not merely about ravens. W_1 is the following:

W_1: ($R_1.B_1$… $R_{100}.B_{100}$). (~ (∃x) Rx ⊃Bx) . ($X_1.Y_1$ ….), where 'X' is any bird other than raven and 'Y' is any colour.

W_2: ($R_1.B_1$…$R_{100}.B_{100}$). ($R_a.W_a$) ($X_1.Y_1$…)

Then the claim is E: Ra.Ba confirms W_2 than W_1 because

P(Ra.Ba|W_1) < P(Ra.Ba|W_2). i.e $\frac{100}{1000100} < \frac{1000}{1001000}$

Good is right in establishing the claim, but the doubt is that, how a claim C(W_1, Ra.Ba) < C(W_2, Ra.Ba) would be a counter-example to $N.C_S$ or $N.C_{(S\text{-}\sim H)}$. It is because, the E:(Ra.Ba) is not at all an instance of either W_1 or W_2. And Hempel articulates this point as follows:

> But whatever the merits of this argument may be, it clearly does not refute the assumption (Nicod criterion). For to do so, it would have to show that an evidence sentence of the form E: 'c is a crow and is black' considered by itself and without reference to any other information, may fail to support the hypothesis that all crows are black; whereas Dr. Good's example concerns the confirmatory role, not of E, but of the vastly stronger evidence sentence S stated above. That an evidence sentence which contains E as one conjunctive component may be disconfirmatory for the hypothesis in question is perfectly obvious; the sentence 'c is a crow and is black, and d is a crow and is non-black' will do. But the assumption (N.C) concerns only the case where the given evidence consists exclusively of a sentence of the form E; and that this sentence must be taken to support the hypothesis seems to me undeniable. (“The White Shoe: No Red Herring.” 239-240)

Hempel’s position is that Good’s discussion is not about the N.C which he envisaged. What is discomforting regarding the Bayesian analysis is that its classification of Nicod’s Criterion is wrongly headed. Bayesians are right in pointing out that Nicod’s Criterion is imprecisely formulated. But it seems that Bayesian classification is also inadequate. One major problem with the Bayesian classification is that two of the categories ($N.C_w$& $N.C_\propto$) are philosophically insignificant. The purpose of the classification is to expose the structures of validity. If two categories are philosophically insignificant, then it means it is structured or formed not in the way in which its validity or invalidity can be determined. ‘Philosophically insignificant’, implies that the form of the classification is not formulated as its validity can be determined. It suggests that the ground of classification itself is wrongly headed.

But an unnoticeable consequence of the Bayesian analysis of N.C is that it fails to account for valid instantial confirmation because the remaining categories are $N.C_S$ & $N.C_T$. It is obvious to see that $N.C_S$ is not a valid version and for a subjective Bayesian there is no confirmation relation with tautological background information ($N.C_T$). In short, none of the versions of N.C is valid. Such a position implies that no instance confirms its corresponding hypothesis in relation with any of the background information. That is, there is not even a single background information where an instance can confirm a hypothesis. Such a position is highly counter intuitive. So, the challenge is that if there is at least one single case where an instance confirms the hypothesis, then we should have a Bayesian version which accounts for it. Bayesian classification is inappropriate because it admits that there would be some cases of instantial valid confirmation, yet its classification fails to arrive at a valid form of N.C

Certainly, Bayesian analysis has contributed much to the clarification of Nicod's Criterion. But I think the Bayesian attempt to resolve the paradox by rejecting Nicod's Criterion is bound to fail unless one holds there cannot be even a single case of instantial confirmation (confirmation by N.C). Present general form of Nicod's' criterion can be challenged but such a challenge is not sufficient to refute the instance as category of confirming evidence. Many of the apparent instances of a hypothesis may not be confirming evidence to the hypothesis but that point out only to the need of reformulation of Nicod's criterion with more stringent condition. If there is a single case of valid confirmation by instances, it poses a philosophical challenge to the Bayesians to come up with a valid form of its general version. And if there is a valid general form of Nicod's criterion then the paradox remains.

II.4.3.Counter Examples to N.C

Swinburne formulates the following counter example: "Let h assert that killer bees cannot live north of the 38th parallel north. Let e assert that killer bees have been observed at

37.99◦N. This is a positive instance of the hypothesis that disconfirms H" (Clarke "The Raven's Paradox is a Misnomer", 437). Clarke responds to it as follows:

> The trouble here is that e says too much. Remember that Nicod's Condition says that evidence of the form "a is an FG" confirms a hypothesis of the form "All Fs are Gs." With h as given in the example, we must take the Fs to be killer bees, and the Gs to be things that cannot live north of 38◦N. Then Nicod's Condition says only that the evidence sentence e*, "a is a killer bee who cannot live north of 38◦N," confirms h; but e asserts more, that "a is a killer bee who lives at 37.99◦N." If we have this further information about where exactly a does live, it becomes more likely that there will be another killer b who lives very slightly north of a, across the 38th parallel north. But without that added information, it is perfectly plausible to hold that h is confirmed by e*, and this is all Nicod's Condition requires. Now, if we held that objects or states of affairs confirmed, rather than sentences, it would be understandable why we should not make the distinction between e and e*. For considering the bee itself who lives at 37.99◦N, or the state of affairs of having a bee live there, would be enough to make us lower our probability of h. But this is an untenable view of confirmation.... (437)

Clarke's response resembles Hempel's response to I. J. Good. That is, what Swinburne evaluates is not merely the instance, but the instance along with other evidence or background information. Therefore, if background information or other evidence is attached to the instance, then it could be the case that the instance ceases to be an instance. In Swinburne's example the background information lowers the likelihood of the hypothesis (probability of evidence given hypothesis is true: $P(E|H)$). That is, $P(E|H) > P(E|H.K)$ Moreover, even after information is added, we cannot conclusively say that the instance along with the added information is disconfirmatory. Instead, I can say that still the instance is confirmatory, but the degree of confirmation is reduced.

Rosenkrantz forms a number of counter examples against N.C in his book, *Inference, Method, and Decision*. He states the following counter example. “Three philosophers leave a party, picking up their hats from the closet on their way out. The closet was dark, and so each picked a hat at random (They were the only three with hats at the party). Philosopher 1 forms the hypothesis, ‘Each philosopher has someone else’s hat.’ As evidence, she sees that Philosopher 2 has Philosopher 3’s hat, and Philosopher 3 has Philosopher 2’s hat. These are both instances of the hypothesis, and so she takes them to confirm her hypothesis. But clearly, given these two pieces of evidence, the hypothesis must be false” (Rosenkrantz, 35). But Clarke criticises the above example.

> The trouble, again, is that the evidence sentences are too precise. Nicod’s Condition only tells us about the confirmational import of sentences of the form ‘Philosopher n has someone else’s hat.’ It’s just greedy to ask exactly whose hat Philosopher n has. Again, if we thought that the hats did the confirming, rather than sentences about the hats, we would be tempted to reject Nicod’s Condition here. (“The Raven’s Paradox is a Misnomer”, 438)

Through Swinburne’s example I maintain that it is controversial to conclude that the instance along with the added information becomes disconfirmatory. It can only be said that it is confirming evidence, but with less degree of confirmation. But in Rosenkrantz’s example, instance clearly becomes disconfirmatory when certain other information are added. It is because, unlike in Swinburne’s case, here the added information is inconsistent to the statement which expresses the instance. In Swinburne’s case the added information only lowers the likelihood of hypothesis.

II.5. Attempt to dispel the paradoxical impression

II.5.1. Bayesian Explanation of Paradoxical Impression

Apart from the attempt to reject Nicod's Criterion another Bayesian attempt was to clarify our intuitive notion which prompts us to think that the paradoxical conclusion is wrong. Subjective Bayesians' argument was that the evidence that 'a' is both non-black and non-raven confirms the hypothesis that 'All ravens are black'. So, there is nothing paradoxical in holding that claim. But a solution needs to explain why such claim (confirmation by contra positive instances) seems as paradoxical, even if it is not paradoxical in fact. Subjective Bayesian answers it in two ways:

1. Though contra positive instances confirm the hypothesis, confirmatory strength by contra positive instances is comparatively less than confirmatory instances. That is why generally it is understood that contra positive instances are neutral in confirmatory strength.

Certainly here, the Bayesians provide an intuitively appealing clarification to the paradoxical impression. In the presence of confirmatory instances, contra positive confirmatory instances became insignificant due to its less confirmatory power. This relative insignificance is interpreted as absolute insignificance. Then the important question to be asked why contra positive instances often seem insignificant even if confirmatory instances are absent. Bayesians answer the above criticism by formulating a quantitative claim:

2. Contra positive instances are not only relatively insignificant in the absolute sense also they are close to being insignificant. That is, confirmatory power is not only less than instances but also it confirms H only to a very minute degree.

"Bayesians argue that E does confirm H—but only to a minute degree. Bayesians also claim that PC (Paradoxical Conclusion) looks unacceptable (i.e., we have the impression that E does not confirm H at all) because we implicitly realize that the degree to which E confirms H is for all practical purposes negligible" (Vrnas 547).

By employing both these claims, Bayesians adequately explain the impression of paradoxicality. Later I analyse in detail, that how such claims are established. Apart from the inadequacies in establishing the above claims, I believe that such an approach in bound to fail and is counter intuitive too. Both claims are counter-intuitive because while these try to remove the paradoxically related to the confirmation of contra positive instances, these assume that all cases of confirmation by contra positive instances are paradoxical. That is a counter- intuitive position. Literature already provides examples of cases where contra positive instances confirm a hypothesis to a high degree and some times more than direct instances. The following is the one example: "Finding a rare infected subject who was not vaccinated in an experimental population provides stronger evidence that the vaccine is 100 per cent effective than finding a vaccinated subject who is uninfected." (Rosenkrantz, *Inference, Method and Decision,* 468)

Quantitative analysis of Bayesian solution rests heavily upon the qualitative claim without making it precise; and as a consequence, even the perfect solution of their approach contradicts themselves. Bayesians are committed to a position that equivalence condition (which is formulated in qualitative frame work) is a necessary condition. However, the standard Bayesian solution claims that contrapositive instances confirm hypothesis only to a minute degree, which runs against the spirit of the condition (E.C). E.C is a necessary condition because equivalent hypothesis is only a different formulation of the hypothesis.

And the confirmation should not be dependent on the way hypothesis is formulated. But the Bayesian solution at the end admits that, degree of confirmation is dependent upon the way it is formulated. That is, confirmation in the quantitative sense depends upon the way it is formulated. And which in essence is equal to denying the E.C which they themselves uphold. And the trouble here is that when the Bayesians provide a quantitative solution it is dependent upon a principle which is not made precise in quantitative frame work. So, the E.C condition can be held in quantitative frame work only as follows: (as suggested by Clarke, "The Raven's Paradox is a Misnomer", 428): if E confirms H, E also confirms its equivalent hypothesis to the same degree. That is the reason, I think, Bayesian approach is bound to be a failure.

Other than these fundamental difficulties the bad news for the Bayesians is that they fail to establish their claims (Quantitative claim and comparative claim) because those claims are built upon controversial assumptions. The following are their controversial claims:

1. $P(\sim B_a|K) > P(R_a|K)$

Given the background information we have the probability of an object being non-black is greater than probability of an object being raven.

2. $P(Ra|(x)(Rx \supset Bx).K) = P(Ra|K)$

The probability of an object being raven is independent of the truth or falsity of hypothesis 'All ravens are black'.

3. $P(\sim Ba\ |(x)(Rx \supset Bx. K) = P \sim Ba|K)$

The probability of an object being non-black is independent of truth of the hypothesis that 'All ravens are black'.

The following is the comparative claim, (COMPc):

$c[(x)(Rx \supset Bx), (Ra \cdot Ba) \mid K] > c[(x)(Rx \supset Bx), (\sim Ba \cdot \sim Ra) \mid K]$.

Here c(H, E | K) is some Bayesian measure of the degree to which E confirms H, relative to background corpus K.

Following are the three measures of confirmation which are defended in the Bayesian literature:

- The Difference: $d[H, E \mid K] = P(H \mid E \cdot K) - P(H \mid K)$
- The Log-Ratio: $r[H, E \mid K] = \log(P(H \mid E \cdot K) / P(H \mid K))$
- The Log-Likelihood-Ratio: $l(H, E \mid K) = \log(P(E \mid H \cdot K) / P(E \mid \sim H \cdot K))$

Fitelson proved the comparative claim by using the three measures of confirmation. And his proof is based on a principle which is largely held as desideratum of measure of confirmation.

Desideratum of measures:

If $P(H \mid E1 \cdot K) > P(H \mid E2 \cdot K)$, then $c(H, E1 \mid K) > c(H, E2 \mid K)$. Measures d, r, and l all satisfy this desideratum. So Fitelson states that one can establish the desired comparative claim simply by demonstrating that:

$(COMP_P)\ P((x)(Rx \supset Bx) \mid Ra \cdot Ba \cdot K) > P((x)(Rx \supset Bx) \mid \sim Ba \cdot \sim Ra \cdot K)$

Peter Vranas pointed out that quantitative claim is sufficient for a Bayesian solution and for a quantitative claim assumption 2, i.e. $(P(Ra \mid (x)(Rx \supset Bx).K) = P(Ra \mid K)$ is not necessary.

Following is the quantitative claim

<u>Theorem 1</u>

Quantitative claim: where measure of confirmation is difference measure (c = d)

(QUANTc): c((x)(Rx ⊃ Bx), (~Ba · ~Ra) | K) > 0, but very nearly 0.

<u>Theorem 2</u>

Quantitative claim: where measure of confirmation is ratio measure (c = r)

(QUANTc): r[(x)(Rx ⊃ Bx), (~Ba · ~Ra) | K] > 0, but very nearly 0.

Vranas claims that the assumption (P(~Ba|(x)(Rx ⊃ Bx.K) = P (~Ba|K)) is necessary for the claim but there is no good reason to assume that it is true. According to him, the only argument which explicitly defended the assumption is of Woodward. "In the absence of some special reason for supposing otherwise, it seems reasonable that my estimate of the number of masses in the universe should not go down (or up) when I learn that they all obey the inverse square law" (Vranas 549). Vranas counters the argument by saying that "my estimate of the percentage of black objects should go up or down when I learn that they include every raven" (549). Vranas' attempt is not to state that the assumption is false, but instead to suggest that it is a disputed one.

> There is a general reason why the disputed assumption is hard to defend. Suppose one somehow refutes the claim that my estimate of the percentage of black objects should go up or down when I learn that they include every raven. The disputed assumption does not follow: it does not follow that my estimate should remain the same. What follows instead is that my estimate may remain the same. Indeed: denying that my estimate should go up or down is compatible with affirming that it may go up or down and thus does not entail that it should remain the same. So even if there is no reason why

> P(Ba|H) and P(Ba) should differ, maybe there is no reason why they should be equal either: maybe they may differ and they may also be equal. But why understand the disputed assumption as the claim that P(Ba|H) and P(Ba) should be—rather than may be or are—equal? (550)

II.5.1.1. Fitelson's Modification of Bayesian Solution

Fitelson admits that canonical Bayesian approach is grounded on disputed assumptions; but he disagrees with Vranas on the point that those assumptions are necessary for the Bayesians' claim. According to him, these assumptions are much stronger than what is needed to establish the claim. Fitelson's first step was the withdrawal from holding the quantitative claim. For him, the Bayesians are neither committed to N.C nor committed to the conclusion which is found as paradoxical. That is, the Bayesians are not bound to own the position that contrapositive instances always confirms H. His position is that if at all an instance confirms a hypothesis, then contrapositive instances too confirms, but with lesser degree. So he solely focuses on the comparative claim.

Fitelson's second step is to formulate minimal assumptions which are necessary for the comparative Bayesian claim which is articulated in l measure framework. And he argues that:

> This solution to the ravens paradox is more general than any other we know of, and it draws on much weaker assumptions. It solves the paradox in that it supplies plausible necessary and sufficient conditions for an instance of a black raven to be more favourable to 'All ravens are black' than an instance of a non-black non-raven. Our most general result doesn't depend on whether the Nicod Condition (N.C) is satisfied nor does it assume the more plausible claim that (given background knowledge) a non-black instance is more probable than a raven instance. Indeed, the conditions for this result may be satisfied even if an instance of a black raven lowers the degree of confirmation for 'All ravens are black'. In that case it just

> shows that non-black non-ravens lower the degree of confirmation even more. (“How Bayesian Confirmation Theory Handles the Paradox of the Ravens.” 22)

He calls the weak and plausible assumptions ‘non-triviality conditions’:

Non-triviality Assumptions:

1. P(Ba·Ra | K) > 0 : it is possible that observed object a will turn out to be a black raven
2. P(~Ba·~Ra | K) > 0 : possible that a will turn out to be a non-black non-raven
3. P(~Ba·Ra | K) > 0 : possible that a will turn out to be a non-black raven
4. 0 < P(H | Ba·Ra·K) < 1, : a to be a black raven neither absolutely proves nor absolutely falsifies ‘All ravens are black.
5. 0 < P(H | ~Ba·~Ra·K) < 1 : a to be a non-black non-raven neither absolutely proves nor absolutely falsifies ‘All ravens are black.

And his third step is a claim that new comparative claim can be drawn on three factors which he labels as ‘p’, ‘q’, and ‘r’. And he works out a result which shows that “relative confirmational support for H from to a black raven instance as compared to that from a non-black non-raven instance is merely a function of p, q, and r.” (Fitelson, “How Bayesian Confirmation Theory Handles the Paradox of the Ravens” 24)

Definition of p, q, & r :

- p = P (Ba | Ra· ~H·K) : it represents the likelihood of ‘a’ being black given that ‘a’ is A raven and H (all ravens are black) is false.
- $q = \frac{P(\sim Ba \mid \sim H \cdot K)}{P(Ra \mid \sim H \cdot K)}$: the factor q represents how much more likely it is that ‘a’ will be a non-black thing than be a raven if H is false (if there are non-black ravens)

- $r = \frac{P(\sim Ba \mid H{\cdot}K)}{P(Ra \mid H{\cdot}K)}$: The factor r represents how much more likely it is that a will be a non-black thing than be a raven if H is true (i.e only black ravens).

Theorem 3 (Fitelson, and Hawthorne “How Bayesian Confirmation Theory Handles the Paradox of the Ravens”, 25.) represents how much more a straightforward instance confirms H than contrapositive instances.

<u>Theorem 3</u>

Given Non-triviality, it follows that q $>$ (1−p) $>$ 0 and

$$\frac{P(Ba{\cdot}Ra \mid H{\cdot}K) / P(Ba{\cdot}Ra \mid \sim H{\cdot}K)}{P(\sim Ba{\cdot}\sim Ra \mid H{\cdot}K) / P(\sim Ba{\cdot}\sim Ra \mid \sim H{\cdot}K)} = [q - (1-p)]/(p{\cdot}r) > 0.$$

The L.H.S $= \frac{l\ (H,E)}{l\ (H,E*)}$ where l is the likelihood measure and H is ‘all ravens are black’ and E is an instance $(Ba \cdot Ra)$ and E* is a contrapositive instance (∼Ba.∼Ra). The first part of the theorem says that the comparative confirmation is the function of factors like p, q, r and the second part says that the value of the function is greater than zero. But theorem only states that $\frac{l\ (H,E)}{l\ (H,E*)}$ can be understood as a function of p, q, r. Fitelson’s attempt is to formulate certain necessary and sufficient conditions where $\frac{l\ (H,E)}{l\ (H,E*)} > 1$. In the following corollary of theorem 1 (25 -26) he made explicit the necessary and sufficient conditions for black ravens to support ‘all ravens are black’ more strongly than non-black non-ravens.

<u>Corollary 1:</u> $\frac{P(Ba{\cdot}Ra \mid H{\cdot}K)\ /\ P(Ba{\cdot}Ra \mid \sim H{\cdot}K)}{P(\sim Ba{\cdot}\sim Ra \mid H{\cdot}K)\ /\ P(\sim Ba{\cdot}\sim Ra \mid \sim H{\cdot}K)} > 1$ if and only if q − (1−p) $>$ p·r.

And, more generally, for any real number s,

$\frac{P(Ba \cdot Ra \mid H \cdot K) \,/\, P(Ba \cdot Ra \mid \sim H \cdot K)}{P(\sim Ba \cdot \sim Ra \mid H \cdot K) \,/\, P(\sim Ba \cdot \sim Ra \mid \sim H \cdot K)} = s = (q - (1-p))/(p \cdot r) > 1$ if and only if $(q - (1-p)) = s \cdot p \cdot r > p \cdot r$.

<u>Corollary 2:</u> (Fitelson and Hawthorne. "How Bayesian Confirmation Theory Handles the Paradox of the Ravens", 27)

Given Non-triviality, for real number's 'such that

$$\frac{P(Ba \cdot Ra \mid H \cdot K) \,/\, P(Ba \cdot Ra \mid \sim H \cdot K)}{P\,(\sim Ba \cdot \sim Ra \mid H \cdot K) \,/\, P\,(\sim Ba \cdot \sim Ra \mid \sim H \cdot K)} = s = (q - (1-p))/(p \cdot r),$$

we have the following:

(1) $s > (1/p) > 1$ iff $q - (1-p) > r$

(2) $s = (1/p) > 1$ iff $q - (1-p) = r$

(3) $(1/p) > s > 1$ iff $r > q - (1-p) > p \cdot r$.

> Notice that when q = r, Clause 3 applies (because then r > q−(1−p)); so the value of the ratio of the likelihood-ratios, s, must be strictly between 1/p and 1. Alternatively, when q diminished by (1−p) is greater than r, Clause 1 applies; so the ratio of likelihood-ratios s must be greater than (1/p), possibly much greater. Indeed, looking back at Corollary 1, we see that the value of the ratio of likelihood ratios s can be enormous, provided only that [q−(1−p)] >>> (p·r). (27)

In the last step of his formulation, he comes up with certain additional assumptions than 'non- trivial conditions'. Let's now look at one more theorem, theorem 4, (28-29) that solves the paradox by drawing on additional conditions that restrict the values of q and r in a plausible way. This result is less general than Theorem 1 and its corollaries, but closely related to them.

<u>Theorem 4</u>

Given Non-triviality, both of the following clauses hold:

(4.1) If P(~Ba | H·K) > P(Ra | H·K) (i.e. if r > 1) and

O(H | Ra·K) / O(H | ~Ba·K) > (p + (1−p)/r), then $\quad : O(X|Y) = \frac{P(X|Y)}{P(\sim X|Y)}$

$$\frac{P(Ba \cdot Ra \mid H \cdot K) \;/\; P(Ba \cdot Ra \mid \sim H \cdot K)}{P(\sim Ba \cdot \sim Ra \mid H \cdot K) \;/\; P(\sim Ba \cdot \sim Ra \mid \sim H \cdot K)} > 1.$$

(4.2) If P(~Ba | H·K) + P(Ra | H·K) (i.e. r + 1), but either P(~Ba | K) > P(Ra | K)

or P(~Ba | ~H·K) > P(Ra | ~H·K) (i.e. q > 1), then

$$\frac{P(Ba \cdot Ra \mid H \cdot K) \;/\; P(Ba \cdot Ra \mid \sim H \cdot K)}{P(\sim Ba \cdot \sim Ra \mid H \cdot K) \;/\; P(\sim Ba \cdot \sim Ra \mid \sim H \cdot K)} > 1$$

Following are the additional assumptions (assumptions other than non- triviaility conditions) which is employed to be derive theorem 4.1. and 4.2.

Additional assumptions for 4.1.

1. r > 1(P(~Ba | H·K) > P(Ra | H·K))
2. and O(H | Ra·K) / O(H | ~Ba·K) > (p + (1−p)/r).

Additional assumptions for 4.2.

1. r ≤ 1,(i.e. P(~Ba | H·K) ≤ P(Ra | H·K)
2. and either P(~Ba | K) > P(Ra | K) or q > 1 (i.e.P(~Ba | ~H·K) > P(Ra | ~H·K)

Fitelson argues for plausibility of the assumptions as follows:

> The first antecedent of (4.1) (first assumption of 4.1) draws on the idea that, provided all of the ravens are black, a randomly

> selected object a is more likely (in our world) to be a non-black thing than a raven. This seems really quite plausible. Indeed, not only does it seem that P[∼Ba | H·K] is merely greater than P[Ra | H·K], quite plausibly P[Ra | H·K] is close enough to 0 that P[∼Ba | H·K] is billions of times greater than P[Ra | H·K] (though the theorem itself doesn't suppose that). (Fitelson, "How Bayesian Confirmation Theory Handles the Paradox of the Ravens." 29)

> Now consider the second antecedent to (4.1) (second assumption of 4.1). One wouldn't normally think that the mere fact that an object is black (without also taking account of whether it's a raven) should provide more evidence for 'All ravens are black' than would the mere fact that an object is a raven (without taking account of its colour). Indeed, generally speaking, one would expect O[H|Ra·K] to be very nearly equal to O[H|∼Ba·K]. However, the second condition for Clause (2.1) is even weaker than this. Notice that for r > 1 the term p + (1−p)/r is less than p+(1−p) = 1; and the larger r happens to be (i.e. the greater the ratio r = P[∼Ba | H·K] / P[Ra | H·K] is), the smaller p + (1−p)/r will be, approaching the lower bound p = P[Ba |Ra·∼H·K] for very large r. Thus, the second condition for (4.1) will be satisfied provided that either O[H| Ra·K] is bigger than or equal to O[H | ∼Ba·K] (perhaps much bigger) or O[H| Ra·K] is a bit smaller than O[H|∼Ba·K]. Thus, this second condition can fail to hold only if (without taking account of whether it's a raven) a black object provides more than a bit more evidence for 'All ravens are black' than would a raven (without taking account of its colour). Given realistic background knowledge K, any reasonable probabilistic confirmation function P should surely satisfy the full antecedent of at least one of these two clauses. Thus, a black raven should favour 'All ravens are black' more than a non-black non-raven over a very wide range of circumstances. Furthermore, neither of the usual approximate independence conditions is required for this result. Thus, Theorem 1 and its corollaries together with Theorem (4) dissolve any air of a qualitative paradox in the case of the ravens. (31)

II.5.2. Critical Analysis of Bayesian Solution

Fitelson established the comparative claim on the basis of trivial conditions and on the basis of additional assumptions (P(∼Ba | H·K) > P(Ra | H·K) and O(H | Ra·K) / O(H | ∼Ba·K) > (p +

(1−p)/r)). Certainly, Fitelson's assumptions are absolutely non- controversial and he is successful in establishing the comparative claim also. But there could be some minor objections to Fitelson's project of establishing comparative claim. Only comparative claim can be derived from Fitelson's weaker assumptions. The question is whether the establishing of comparative claim serves the purpose. I think comparative claim alone cannot dispel the paradoxical impression attached to the confirmation by contra positive instances. At best we can defend Fitelson by claiming that If $C(H, E_1) < C(H, E_2)$ then in the background of E_2, E_1 would not be confirming evidence. But the claim of paradox is much bigger than that. It is not merely that in the background of instances, contrapositive instances are considered as neutral or insignificant evidence. But even in the absence of the evidence 'black raven, contra positive instances (red pencil) is considered as insignificant or neutral evidence. But my larger criticism to Bayesian solutions (of both FItelson and traditional) is not about the inadequacies in establishing the comparative and quantitative claim. I believe that Bayesian approach is bound to fail and is counter intuitive too. Both claims are counter-intuitive because while trying to remove the paradoxicality related to the confirmation by contra positive instances, they assume that confirmation by all contra positive instances are paradoxical. That is a counter- intuitive position. Not all contra positive instances (contra positive instances as defined standard definition) may be confirmatory to hypothesis. However, the Bayesian solution states that all contrapositive instances are necessarily less confirmatory than the direct instances. Literature provides examples of cases where contra positive instances confirm a hypothesis to a high degree and some times more than direct instances.

II.5.2.1. Bayesian Solution and Equivalence Condition

I think the Bayesian approach is bound to be a failure because in their solution they violate the E.C (Equivalence Condition), specifically Quantitative Equivalence Condition (Q.E.C.): If E confirms H to the degree α then E also confirms H_1 to the same degree where $H \equiv H_1$. Key point of the Bayesian solution is that if an instance confirms H, then contra-positive instance too confirms H, but with a lesser degree. Apparently it does not even violate the Q.E.C. The Bayesian solution only states that evidence (E_1: instance) confirms H more strongly than another evidence (E_2: contra-positive instance)

i.e, $C(H,E_1) > C(H,E_2)$

While Q.E.C states that

$C(H,E_1) = C(H_1,E_1)$ where $H \equiv H_1$.

Apparently, these are not contradictories. Though Bayesians worked out their solution within the background information of an object being black and an object being raven, its implications are not merely on the realm of the hypotheses: 'All ravens are black' and 'All non-black things are non-raven'. In other words, the raven paradox is not about ravens as stated (Roger Clarke "The Ravens Paradox is a misnomer", 431). The general point of the Bayesian solution is that the contrapositive instances are less confirming evidence than the (direct) instances. To know how the Bayesian solution violates the Q.E.C, we need to work out the relation between instances and contra positive instances. In the context of raven hypothesis, (x) (Rx ⊃ Bx), the Bayesians consider instances and contrapositive instances as two distinct evidence (like Ra . Ba and ∼Ba . ∼Ra) but that need not be the case always. A proposition itself can be both instance and contrapositive instance in relation with two distinct hypotheses which are logically equivalent. My point is that in such cases, the general point of the Bayesian solution (contrapositive instance are less confirmatory than direct instances) violates the equivalence condition. The evidence (∼Ba.∼Ra) is a direct instance of the hypothesis (x) (∼Bx⊃ ∼Rx) while it is a contrapositive instance of the hypothesis, (x) (Rx ⊃Bx). In the same way, Ra.Ba is direct instance of (x) (Rx⊃Bx) while it is contrapositive of the hypothesis (x) (∼Bx⊃ ∼Rx). My argument is that in the context where both instances and contra positive instances are the same proposition, the Bayesian solution violates the Q.E.C. In such a context, the Bayesian solution is the following:

C((x) (∼Bx⊃ ∼Rx), ∼Ba . ∼Ra) > C((x) (Rx ⊃ Bx), ∼Ba . ∼Ra)

While the Q.E.C stipulates that

C((x) (∼Bx⊃ ∼Rx), ∼Ba . ∼Ra) = C((x) (Rx ⊃ Bx), ∼Ba . ∼Ra)

That is, the Bayesian solution contradicts Q.E.C.

The Bayesian's point is that the paradoxicality arises due to less confirming strength of the contrapositive instances. I think that the larger point of raven paradox is that the subjective Bayesians' basic definition of the confirmation relation is a highly problematic one. For the Bayesians, if the posterior probability of a hypothesis in view of a proposition is higher than the prior probability of the hypothesis, the proposition can be considered as evidence. This is the Bayesian definition of qualitative evidence. White shoe satisfies the definition. Is it evidence of the raven hypothesis in a qualitative sense? A mere updation of probability of hypothesis is not sufficient to consider a hypothesis as confirmed one. Otherwise we would have to agree that white shoe is confirming evidence since it updates the probability of the raven hypothesis and whether the degree of confirmation is minute or large is immaterial in the qualitative definition of evidence as long as the qualitative definition of evidence incorporates a condition of confirmation in terms of degree of confirmation. In the discussion of qualitative notion of evidence, the BCT's solution is a circular one: Raven paradox is a challenge because it is the statement of a unintuitive consequence of the BCT's definition. The BCT solves the paradox by stating that it is not unintuitive because it is the consequence of BCT definition.

II.5.3. Paradoxicality is only an Impression: Hempel's dismissal of the paradox

The above discussed solution was the most convincing answer from the Bayesians to the paradox. As we have seen, such a solution is fraught with many difficulties. The general approach of the solution is dissolving the impression of paradoxicality. Hempel was the pioneer champion of such an approach. Hempel's final solution was that there is nothing paradoxical in holding the conclusion as true. And the paradoxicality is only an impression. Thus, he maintains there is no reason to reject any of the premises or the conclusion. And paradoxicality is only a psychological one. He further explores why such an impression of paradoxicality arises and tries to remove the impression of paradoxicality from the derived conclusion.

According to him, we found a non-black-non-raven as neutral evidence, not as confirming evidence because even prior to the testing we have the knowledge that object is not a raven. And such prior knowledge about evidence is the factor which leads to the paradox. He argues further through the following example: consider the testing of hypothesis "All sodium salts burn yellow" and we are testing an object whose chemical constituents are unknown. While burning the object, the colour of the flame found as non yellow.(E.g.-blue). So, here the object is potentially a counter-example, because it could be the case that the object is a sodium salt. However, further testing shows it is not a sodium salt. Such a finding is a relief to the proponents of hypothesis because the possibility of an object/statement being counter example is averted. So, Hempel argues that such situation provides the support to the hypothesis. Though, Hempel's line of thinking implicitly or explicitly still give rise to various solutions, it is severely criticised as mistaken.

Hempel's position is that ∼Ba. ∼Ra is a confirming evidence of (x) (∼Bx⊃ ∼Rx) but Wa.Sa (White Shoe) is not confirming evidence though Wa.Sa along with certain background information (white is a non-black colour, shoe is a non-raven thing) entails ∼Ba. ∼Ra. That is E confirms H, but if an agent has knowledge that E is a consequence of a proposition 'P' . Then E ceases to be confirming evidence. And Hempel's reasoning is that when we know that another propositions 'P' entails E, we acquire a lot more background knowledge about E, which plays a crucial role in the confirmation relation between H and E. Hempel's position is that a non-black non-raven confirms the hypothesis, H. And any other information we acquire about the object 'a' other than it is non-black non-raven, is a background information. And Hempel holds that in view of background information, confirmation relation between H and E cannot be explicated adequately as background information too acts as another relevant evidence of H in conjunction with E. Thus Hempel envisage of confirmation relation as a relation between H and E and holds that any reference to the background information has to be avoided. Primarily Hempel's insistence for the exclusion of background information is problematic on various counts. Bayesians hold that background information is a necessary constituent of confirmation relation. In other words, no proposition can be evidence of a hypothesis in the absence of any background information. But even if we stick to Hempel's framework of analysis of confirmation (which he calls as methodological fiction, where the constituents of the confirmation relation is only hypothesis and evidence) it is difficult to validate Hempel's solution of the paradox. In Hempel's solution, he considers all information regarding the object 'a' except that 'a' is non-raven and non-black, as background information. But such an analysis about background information is problematic.

A non-raven or non-black is not a predicate which can be observed like the predicates red or pencil. A predicate like non-raven or non-black cannot be observed but can be only inferred. That is, to claim that an object 'a' is a non-raven, we need to have prior information about the object. In other words, the information like the object 'a' is white shoe or red pencil is necessary condition to claim that the object 'a' is non-black non-raven. In that sense, they are necessary constituents to define E as ∼Ba. ∼Ra. So the information which Hempel excluded from the scenario of confirmation is necessary constituents to state the evidence as ∼Ba. ∼Ra. In other words, the information like 'a' is white shoe or red pencil is a necessary constituent to state the evidence as ∼Ba. ∼Ra. So Hempel's exclusion of certain information is inadequate to explicate the confirmation relation.

I think, Hempel's argument consists of an intuition which he failed to articulate adequately. His point is that non-black non-raven is not simply a red pencil. A Red pencil could be a non-black, non-raven but with certain qualifications. Qualification which he added is not that it should not be known at the time of testing as a pencil. Its constitution should be unknown. Moreover, it is not clear how the qualification elevates a red pencil to a non-black non-raven.

One way we can pursue Hempel's argument is that interpreting his position as red pencil is not a non-black non-raven. The predicate non-raven can be interpreted in two ways. One is that anything which is not a raven. Here non raven is considered simply as a negation of the predicate ‘raven’ or as the complement of the set of ravens. Another way is treating non-raven, itself as predicate/property like red or blue not as a negation of one property like ‘Raven’. If we pursue the second interpretation, a pencil is not a non-raven because there is no reason to think that pencil has the property of being a ‘non-raven’ like it has the property, e.g. pencilness, or redness. Then the crucial question is what is the predicate/property ‘non-raven’ means. That is a substance/thing about which we know only one thing that it is not a raven is a non-raven. That is, it must have the possibility of being anything else except raven. Then it represents/reflects the predicate of ‘non-raven’.

It is clear that Hempel is emphasising a point that non-raven as a predicate is something more than simply not being a raven. For Hempel, to qualify an object as non-raven, the agent must assume object has the possibility of having any property other than raven. Therefore, a pencil cannot be a non-raven because the agent cannot see the possibility of the object being a shoe or a swan or book since it is known as pencil. So in Hempel’s account, the possibility/potential being anything other than the raven is the condition to consider an object as non-raven.

But it is doubtful such a privilege alone confers the status of non-raven. Here, Hempel insists that one thing which is found as not a raven, not by finding what exactly it is. But often we found thing as not raven by identifying what exactly that is. So to be a non-raven, we should not know anything about the object (except that it is not a raven) is perhaps too stringent and improper and which consists of some elements of fiction. More than that we have good reason to believe that even if an object is known as x (like a swan, or a crow) we can consider it as non-raven. I think Hempel is right in pointing out that only through satisfaction of certain conditions one can qualify an object (like pencil, shoe) as non-raven. But I think that Hempel is mistaken in formulating what are those conditions.

Hempel understood non-raven as one which has the potential of being anything else than raven. The other way to interpret non-raven is that it is a qualified predicate. That is a context demands certain qualification. And the important question is what are those qualifications. In certain context, we would like to emphasise that it is not raven. What demands such an emphasis? It is important to unearth, what is the necessity of such qualification. One ground for such qualification is the similarity with ravens. In such cases it is easy to see that why contra-positive instances are confirming one. If we mistakenly took an object as raven (because of similarity) and found it as non-black, then an emphasis that it is not raven averts the possibility of a stronger counter example. There are other conditions where one can qualify a thing as non-raven. If one ceases to be a raven (e.g. genetically modified) then it is adequate to qualify it as non-raven. The point in nutshell is that, a thing cannot be non-raven simply because it is not-raven but only if there is a stronger reason to emphasise that it is not a raven. The notion of 'strong reason' is not a well articulated one. But it is reasonable to assume that in practicing realm, we have clear judgement about when we should qualify an object as not raven. Hempel attempts to dismiss the paradox, by tracing the impression of paradoxicality.

II.6. A Bayesian Reformulation of Hempel' Solution

Bayesians seriously take Hempel's solution to the paradox. Fitelson is one among them who clarifies Hempel's point more vividly in the Bayesian Framework.

II.6.1. Fitelson on Hempel

But Fitelson rescues Hempel's point in a different way though he rejects Hempel's theory of confirmation as incapable for articulating his own point. He states that Hempel distinguishes between two claims: P.C and P.C* "As Hempel explains one might be misled into thinking that (PC) is false by conflating (PC) with a different claim (PC*) − a claim that is, in fact, false."[21]

(PC*) If one observes that an object a − already known to be a non-raven − is non-black (hence, is a non-black non-raven), then this observation confirms that all ravens are black.

E confirms H, given A – e.g., $\sim$ Ba confirms (x)(Rx $\supset$ Bx), given $\sim$ Ra.

P.C: (E · A) confirms H, unconditionally –

e.g., ($\sim$ Ba · $\sim$ Ra) confirms (x)(Rx $\supset$ Bx) unconditionally.

And Fitelson states that such distinction cannot be made in classical logic and consequently in Hempel's theory of confirmation.

II.6.2. A New Bayesian Definition of Instance

Among the solutions, my lineage is to the solutions which attempt to reject or modify N.C But I have strong reason to state that any attempt to invalidate N.C in toto is doomed to failure. Thus I am inclined to the style of Hempel and Quine who tried to constrain N.C instead of rejecting it. I think constraints on background knowledge help to exclude the invalid instantial confirmation. One of the ways to constrain 'K' is to insist that 'K' itself should not have confirmatory impact on hypothesis. It is formalised as follows: there should

[21] Fitelson, B. "How Bayesian Confirmation Theory Handles the Paradox of the Ravens" P.3

not be any ‘K’ such that K.E confirms (disconfirms) the hypothesis but ∼K.E disconfirms (confirms) the hypothesis. It does not matter if ∼K.E is neutral to hypothesis, but it should not be disconfirming while K.E is confirmatory. I think probably what Hempel had in mind would be such ‘K’ and ‘N.C’ relative to such ‘K’. I think such constraints can prevent many of the counter examples raised against N.C., Obviously such moves do not help to solve the paradox.

At the outset, paradoxical conclusion is shown as derived from two premises (E.C&N.C). In my attempt to resolve the paradox, I would like to construe the paradox as follows (Various discussions hint on such a construal of the paradox):

1. The hypothesis ‘All Ravens are Black” can be expressed as (x) (Rx ⊃ Bx)

2. Equivalence Condition

3. Nicod’s criterion

1-3 entails conclusion 1.

4. Conclusion 1(Con1): ∼Ra. ∼Ba confirms (x) (Rx⊃Bx)

5. ∼Ra. ∼Ba(a non- raven which is non- black) is Wa. Sa (White Shoe)

4-5 entails conclusion 2 (Con2)

6. Conclusion 2 (Con2): Wa. Sa (White Shoe) confirms (x) (Rx⊃Bx).

I hold the position that conclusion_1 (∼Ra. ∼Ba confirms (x) (Rx⊃Bx)) is not paradoxical. And what is paradoxical is conclusion_2 (Wa. Sa (White Shoe) confirms (x) (Rx⊃Bx)) which is derived from conclusion_1 using premise_5 (∼Ra. ∼Ba (a non- raven which is non- black) is Wa.

Sa (White Shoe)). And the premise 5 which is employed to derive the paradoxical conclusion (conclusion 2) is untenable because 'white shoe' or 'red pencil' or 'white board' are not instances of hypothesis that 'All non- black things are non- raven'. In other words they are not 'non- black non- raven'. In order to substantiate my claim, I attempt to make precise the definition of 'Instance'.

In literature, the sole condition of being an instance of H is, that the object 'I' should exemplify the predicate(s) of H. That is an instance of H ((x)(Fx ⊃ Gx)), is an object 'I' who has a pair of predicate 'F' and 'G'. My attempt is to pursue the other conditions of instances which precludes the possibility that 'white shoe' is an instance of the hypothesis 'All ravens are black'. I assume the paradox generated from the imprecise definition of instance. I think the precise analysis of our intuition can unearth much more characteristics of instances.

Another obvious characteristic of instance is that it is deductive evidence. But all deductive evidence are not instances. Then what exactly distinguish instances from other deductive evidence. The other kind of deductive evidence (non- instantial deductive evidence) has larger scopes in the confirmation scenario. E.g. While instances can confirm only general hypothesis, non- instantial deductive evidence can confirm even existential hypothesis. But that is the distinction in terms of scope and significance. It is not a formal or logical distinction between them. The question I pursue is, what would be the exact difference between instances and non- instantial deductive evidence in the confirmation of general hypothesis. I think though $P(E|H) = 1$ for both instances and deductive evidence, often instances are stronger confirming evidence than other kind of deductive evidence .

Consider the following examples: Case: 1.

H: Light consists of particles, E: Reflection of light.

Case 2:

H: All ravens are black. , E_1: genetic structure of raven is associated with dark colour.

In both cases, evidence E and E_1 are deductive evidence. But that deductive evidence could be the deductive evidence of its competing hypothesis too. E_1 could be a deductive evidence to the hypothesis that 'all ravens are red'. But my claim is that category of 'instance' is stronger than the category of 'non- instance deductive evidence' because an object which is an instance of hypothesis can never be an instance of its competing / contrary hypothesis while it is not the case with the deductive evidence in general and non- deductive evidence in particular.

The distinction which I would like to draw is that instances could be less favourable to competing hypotheses than non- instance deductive evidence. I think the distinction is the following: If 'I' is instance of H then, 'I' cannot be an instance of any contrary hypothesis of H, though 'I' could be a confirming evidence of a competing hypothesis, but not instances. But deductive evidence of H could be a deductive evidence of it contrary hypothesis. On the basis of above discussion, I add the following conditions which are simple and quite compatible to our intuition regarding the instances.

Condition 2: 'I' is an instance of H only if: $P(I|H.K) = 1$. It only says that instances are deductive evidence which is obvious but it can prevent Rosenkrantz's counter example.

Condition 3: if 'I' is an instance of H then it would not be an instance to any other competing hypothesis (hypothesis which are contrary to H). In general we can say as follows: 'I' is an instance of H only if $P(I|H.K) > P(I|H_c.K)$: H_c is any hypothesis which is contrary to H.

On the basis of the third condition, I say that white shoe is not an instance of H: (x) $(\sim Bx \supset \sim Rx)$. Consider some contrary hypotheses of 'All non- black things are non – raven'.

H_{C1}: All non- yellow things are non- raven, H_{C2}: All non- blue things are non- raven, H_{C3}: all non- white things are non- raven. Here except for the hypothesis H_{C3}, for all hypotheses, the likelihood of white shoe is 1. i.e. $P(Wa.Sa|H.K) = P(Wa.Sa|\ H_{C1}.K) = P(Wa.Sa|\ H_{C2}.K) = 1$. That is, white shoe is a confirming instance (as per traditional definition) for H and for all contrary hypotheses of H except for one. It is a clear violation of Condition 3. Therefore as per the new definition of instance, 'Wa.Sa' is not a confirming instance of H: (x) $(\sim Bx \supset \sim Rx)$. Thus the paradoxical conclusion 2 cannot be derived.

I think it would not be difficult to accept that 'Wa.Sa' is not an instance of H, but the question remains as to what kind of an object is 'non-raven and non- black'. Now the challenge is that we have vast kind of objects (e.g. white shoe, red pencil...) which satisfy the 'Conditions of Instance' 1 and 2.(all objects exemplify the predicates non- raven and non- black and likelihood of hypothesis regarding these objects is 1). Though many of it violates 'Condition 3', it cannot function as heuristic principle to identify the real instances as it is only test of instances. To pin point the real instance, I formulate a reasonable assumption: 'From a set of objects which satisfies condition 1, the object(s) which provides greater support to H and than other object is the real instance(s). In this case, for objects 'a' (red pencil), 'b' (white shoe) ,c (blue water), $P(\sim Ra. \sim Ba|H.K) = 1$, $P(\sim Rb. \sim Bb|H.K) = 1$, $P(\sim Rc. \sim Bc|H.K) = 1$, then the objects which have the lowest probability (P(E)) for its occurrence would provide maximum supports to H and as per our assumption that would be the instance of H. So to determine what is the real instance we need to determine P(∼Ra. ∼Ba), P(∼Rb. ∼Bb), P(∼Rc. ∼Bc). Obviously the probability of white shoe or red pencil being non- black non –raven is high. But consider an object whose colour is near to black and it is so similar to raven, yet not black and not raven; I think it would not be non- controversial to assume that object as an instance of contrapositive hypothesis ‘all non- black things are non- raven’. Consider another hypothesis: Smoking causes cancer. That is, ‘all smokers are cancerous’. So, the contrapositive instance is ‘all non cancerous things are non- smokers’. Here we can have a vast collection of objects which exemplify the predicates; for example, ‘stone and pencil and animals which are both non cancerous and non- smokers’. But the probability of stone being ‘non- cancerous and non- smokers’ is very high compared to the probability of a human being who is non smoker and non- cancerous. And our intuition tells that the human being who is non- cancerous and non- smokers is the real instance of contra positive hypothesis and then it is easy to see that that contrapositive instance confirms original hypothesis ‘all smokers are cancerous’. Thus there is nothing paradoxical in holding that contrapositive instance confirms H and its logical equivalent hypothesis if instances and contra positive instances are understood as per the new definition of instance.

Conclusion

Major attempts (Hempel’s and subjective Bayesians’) to solve the raven paradox assumes that the notion of paradoxicality is only an impression. And they attempt to clarify the impression of the paradoxical nature of the conclusion (contrapositive instance confirms the original hypothesis) in various ways. Hempel attempts to clarify the paradoxicality by providing a more precise characterization of the confirmation relation which he calls a methodological fiction. And the subjective Bayesians attempt to clarify the paradoxical impression by characterizing and comparing various confirmation relations by employing the notion of degree of confirmation.

Though Hempel's solution fails to survive, the subjective Bayesians' solution remains a strong contender for the solution of the paradox. The Bayesian's point is that paradoxicality arises due to less confirming strength of contrapositive instances. But my basic objection to the Bayesian solution is that such a less confirming evidence (as defined by the Bayesians) is not an evidence at all (in qualitative sense). I consider the paradox as a fundamental challenge against the Bayesian definition of confirmation. A mere updation of probability of hypothesis is not sufficient to consider a hypothesis as confirmed one. In the first chapter, I have argued that updation of probability of hypothesis has to be compared with the updation of probability of other competing hypotheses to determine the confirmation of the hypothesis. And my redefinition of the instance-confirmation brings forth instance's or evidence's relation to other competing hypotheses. And through the redefinition of instance-confirmation, I was attempting to reaffirm the point that confirmation relation cannot be explicated without analyzing the competing hypothesis's relation (especially the best competing hypothesis's relation) to the evidence.

Chapter III

The Problem of Old Evidence

Introduction

The ‘problem of old evidence’ claims that the Bayesian Confirmation Theory (BCT) fails to capture the confirmation of a hypothesis in science by a certain kind of evidence called ‘old evidence’. In the strict sense it is not a paradox because the claim is not that certain implications of BCT are paradoxical; but the point is that BCT fails to capture what one intuitively considers as adequate confirmation relation.

The ‘problem of old evidence’ was formulated by Glymour (*Theory and Evidence*) some evidence are discovered prior to the formulation of hypotheses. Such evidence are called old evidence. Citing historical examples, Glymour points out that often old evidence works as confirmatory evidence though it is not the case that all old evidence are relevant. The problem is that the BCT fails to capture the confirmation by old evidence. Glymour elucidates the point as follows:

> Scientists commonly argue for their theories from evidence known long before the theories were introduced. Copernicus argued for his theory using observations made over the course of millenia, not on the basis of any startling new predictions derived from the theory, and presumably it was on the basis of such arguments that he won the adherence of his early disciples. Newton argued for universal gravitation using Kepler's second and third laws, established before the Principia was published. The argument that Einstein gave in 1915 for his gravitational field equations was that they

> explained the anomalous advance of the perihelion of Mercury, established more than half a century earlier. Other physicists found the argument enormously forceful, and it is a fair conjecture that without it the British would not have mounted the famous eclipse expedition of 1919. Old evidence can in fact confirm new theory, but according to Bayesian kinematics it cannot. (*Theory and Evidence*, 85-86)

Glymour depicts the problem as follows:

<u>Theorem 1</u>

<u>First Part</u>

$P(E|K) = 1$ entails that $P(E|H.K) = P(E|\sim H.K) = 1$

<u>Second Part</u>

$P(H|E.K) = P(H|K)$ when $P(E|H.K) = P(E|K) = 1$.

<u>Proof:</u>

<u>First Part</u>

1. $P(E|K) = P(E|H.K) \times P(H|K) + P(E|\sim H.K) \times P(\sim H|K)$
2. $1 = P(E|H.K) \times P(H|K) + P(E|\sim H.K) \times P(\sim H|K)$: Old Evidence (This identity is expressed as *)
3. If $P(E|H.K) < 1$ and $P(E|\sim H.K) < 1$ then R.H.S of $* < 1$: In the multiplication of two decimal factors, where a each is less than one, the product will be less than each individual factor

4. So both $P(E|H.K)$ and $P(E|\sim H.K)$ are not less than 1
5. If $P(E|H.K) = 1$ and $P(E|\sim H.K) < 1$ then R.H.S of * < 1 :Since $P(E|\sim H.K) \times P(\sim H|K) < P(\sim H|K)$ if $P(E|\sim H.K) < 1$.
6. If $P(E|H.K) <$ and $P(E|\sim H.K)$ 1 then R.H.S of $* = 1$: Since $P(E|H.K) \times P(H|K) < P(H|K)$ if $P(E|H.K) < 1$.
7. So if $P(E|H.K) =$ and $P(E|\sim H.K) = 1$. : * , Premises 4, 5 and 6

<u>Second part</u>

1. $P(H|E.K) = \frac{P(E|H.K) \times P(H|K)}{P(E|K)}$: Bayes' theorem
2. $P(H|E.K) = P(H|K)$: Since $P(E|K) = P(E|H.K) = 1$

All standard measures of Bayesian confirmation (d, r, l, and s)[22] are found as vulnerable to the problem as all measures are based on the relation between prior probability and posterior probability of the hypothesis.

Glymour argues that, in the case of old evidence, $P(E) = 1$ because agent is certain about the truth of E since the old-evidence is a description of a phenomena which is observed in past. Even before the Einstein's formulation of theory of relativity, scientists had observed the phenomenon of 'advancement of perihelion of mercury'. And the old evidence is the observation report, which describes the phenomena which occurred sometime in the past. And what is certain or true is the observation report. The Bayesian framework provides the possibility of a more adequate articulation of Glymour's claim of P(old evidence) $= 1$. Bayesians represent probability of evidence as $P(E|K)$ where 'K' is the

[22] d: difference measure, r: ratio measure, *l:* likelihood measure s: Normalized difference measure.

background information an agent holds at the time of testing. In the case of testing of hypothesis by the old-evidence (E), the background information itself contains the E since the agent holds that E is true. So the background information can be characterized as follows:

K: (E. (K-E))

Probability of old-evidence (E) can be expressed as follows: P(E|(E.(K-E))

Since E entails E, it is adequate to argue that $P(E|(E.(K\text{-}E)) = 1$ i.e. $P(E|K) = 1$.

On the following two counts, the problem of old evidence is more damaging than any other paradox/shortcoming:

1. The problem of old evidence is one which is formulated within a quantitative framework. So it does not pose any challenge to other qualitative theories of confirmation like IBE, H-D Model and Hempel's theory.
2. BCT's main strength or uniqueness is that it is able to capture (describe by using Bayesian tools) almost all features of descriptive realm. Its forte is not its ability to solve the paradoxes as it does not solve any paradoxes in a conclusive manner though it provides improved solutions. But the perceived superiority of BCT lies on the point that Bayesian mechanism can explain/ register what is intuitively / in practice considered as confirmation. Problem of old evidence challenges the BCT exactly on the point which is considered as a stronghold of BCT.

In the first section of this chapter I introduce the classification of the problem of old evidence. Bayesian literatures provide a large cluster of classification of the problem and many of the solutions are dependent upon the classification. That is, Bayesians provide different solutions to different categories of the old evidence problem. In the second section, I introduce different suggested solutions; first among them is the Counter Factual Strategy. Collin Howson and Peter Urbach formulate their solution called the Counter-Factual Strategy (C.F.S) and the second solution which I discuss is the one proposed by John Earman. And although Earman introduced various solutions to the problem but later he himself questioned their validity. The important among Earman's solutions is dissolving the problem by assuming that the probability of a proposition would be one only if it is tautology, in all other cases, even if it is a proposition of fact which is known, the highest value it can acquire is less than 1, like .999. Though it straight way dissolves the original problem of old evidence, it gives rise to the quantitative problem of old evidence which states that the BCT cannot register any significant support by the old evidence, since C(H,E)[23] would be very small, when P(E) is very high. But Fitelson dismisses the quantitative problem of old evidence by introducing l measure. He claims that l measure can be arbitrarily large even if P(E) $\approx$1. But in a dismissive attack to such routes of dissolving the paradox, Maher argues that the qualitative problem of old evidence cannot be dismissed by assuming that probability of old evidence is less than 1 since the conceptual difficulty the problem raises is of the inconsistency in the BCT framework of being P(E) = 1 and its confirming status. And the other major solution is suggested by Ellery Eells which mainly attempts to solve the ahistrocial problem by distinguishing E as a confirming evidence and E as a favourable evidence. And in the second last section of the chapter, I discuss Daniel Garber's solution by giving up the logical omniscience aspect of Bayesian Frame work. And in the last section, I discuss the issue of old evidence in the new Bayesian framework.

[23] Degree of confirmation of H by E. C(H,E) is the expression to represent any measure of degree of confirmation.(difference, ratio, likelihood, normalized difference)

III.1. Classification of the Problem

The primary classification of the problem of old evidence is: ahistorical problem of evidence and historical problem of evidence. Historical problem of old evidence is the problem which is originally posed by Glymour. When scientists knows a confirmation relation between a newly invented theory and the evidence which is already known, the historical problem of old evidence arises. Ahistorical problem of old evidence expands the range of the problem to the new evidence also. New evidence means evidence which is discovered after the formulation of the hypothesis. I think, precisely we can say that, new evidence is the evidence discovered in the process of testing of hypothesis. In the ahistorical problem, even after the event of confirmation evidence remains as evidence for H and the BCT cannot explain the time enduring nature of evidence. It is because, after the event of confirmation, P (E) becomes 1. Then an analysis provided the BCT fails to capture the confirmation of H by E at a later point of time.

The historical problem of old evidence is the one which is originally considered as the problem of old evidence. Classification brought up a new category of problem called the ahistorical problem. Significance of the classification is that the problem is expanded or considered as one which affects all kinds of evidence rather than as a particular category of evidence which is discovered before the formulation of the hypothesis, H. But, rather than clarifying the significance/scope of the problem, the classification is used as a tool to support many of the proposed solutions. Bayesians even provided different solutions to both these categories. Such a move is legitimate because both problems incorporate different structures: The structure of the historical problem is the following:

1. Agent considers E as confirming evidence of H at time t_1.

2. But agent knows E since time t_0.

The structure of the ahistorical problem is the following:

1. Agent considers E as confirming evidence of H at time t_1.

2. Agent knows the confirmation relation between H and E at time t_0

3. But agent also knows E since time t_0.

The difference is that in the historical problem, prior to the confirmation event, agent does not discover the confirmation relation between H and E but only discovers E. Eells elaborates the classification. Eells introduces further classification to present his solutions. As we know an ahistorical problem affects not only old evidence but also new evidence. Eells thinks that two categories of the ahistorical problem should be dealt distinctly. Following is his classification:

Categorisation of Problem of Old Evidence

Sl.No.	Eells' Categories	Corresponding Traditional categories	Description
I.	The problem of old new evidence	Ahistorical problem of new evidence.	New evidence is considered as evidence even after the time of discovery of it. (Even after it becomes old).
II	The problem of old evidence		
II.A	The problem of old old evidence.	Ahistorical problem of old evidence	Even after the event of confirmation, old evidence is considered as evidence.

II.B	The problem of new old evidence	Historical problem of old evidence (The original problem posed by Glymour)	When scientists first discover the confirmation relation between H and E where E is known earlier.

Attempted Solutions

III.2. Counter Factual Strategy

Collin Howson and Peter Urbach formulate their solution called Counter Factual strategy (C.F.S). They claim that the problem of old evidence is nothing but misapplication of Bayesian principles. They articulate C.F.S as follows:

> ...Bayesian assesses the contemporary support E gives H by how much the agent would change his odds on H were he now to come to know E. This means that in the context of K he will compute the difference between what his degree of belief would be in H assuming (counterfactually if he already knows E) that his state of knowledge were K- {E}, and what his degree of belief in H would be were he, knowing K- {E}, to come to know E. It follows that the support he regards E as giving H is equal to P(H|E) - P(H), both probabilities being relativised to K- {E} whether he knows E or not. In other words, the theory is explicitly a theory of dispositional properties of the agent's belief-structure, though the implicit conditions may not be realised in practice.(Howson, "Bayesianism and Support by Novel Facts", 246)

The C.F.S position can be understood in the following way: the confirmation relation between evidence and hypothesis consists of a general pattern and the general pattern can be explicated through the Bayesian mechanism. And BCT's task is not to describe how an agent draws/discovers confirmation relation. The pattern lies in the analysis of prior probability and posterior probability of H in the light of evidence. Posterior probability is evaluated by looking on the relation between H and E in the light of background information. The role of background information (K) is very important and limited too. One of the roles of K is making E and H meaningful statements. That is specifying what E and H refer to. And the other role is in outlining the possible ways the relation between E and K hold and showing that some of these relations as improbable. Apart from making E and H meaningful, its main role is elucidating the way in which H and E can be related.

C.F Strategists insist that it is a necessary requirement that propositions of K should be logically independent of H and E. That is, truth/ falsity of H should not be determined by K. And it is not only the requirement regarding the old evidence but also regarding the new evidence. Howson and Urbach hold that as long as BCT's task is to explicate the logic of confirmation, and not the description of the practice of confirmation, the problem of old-evidence would not be a challenge at all.

In the context of confirmation, the agent may have a lot of information regarding which is relevant and which is irrelevant to confirmation relation. It is unintuitive to say that the agent is making use of all information which he/she had acquired till then, in judging the confirmation relation of H and E. So it is intuitive to assume that the agent is distinguishing the information which is required to determine the confirmation from those that are not required. The question is, what is the basis of this distinction? Howson and Urbach extend it and say that agent excludes the information which is logically dependent upon H and E. It is because, the logically dependent information may prejudice the relation in one way or another. Suppose, in confirmation relation, we know that H is false (K), we may prefer to say that E is not a confirming evidence of H even if H entails E. I think, the proponents of C.F.S. may argue that confirmation should be judged on the basis of the relation between H and E.

In the above example, though there is a strong logical relation between H and E, an agent would discard the relation and conclude that there is no confirmation relation because he/she was prejudiced by the logically dependent information, available from K, that H is false.

C.F.S position can be summarised as follows:

1. It is intuitive and widely acceptable that certain information should be excluded from background knowledge (K)
2. And a rational agent ought to exclude the information which is logically dependent on H and E.

So K (background information) is not all the information which an agent possesses but information which ought to be used to judge a confirmation relation. So in the case of old – evidence, K contains an information which is logically dependent on E, that is, E itself (E is true). So it is plausible to assume that the background information is not K itself but K-E. Howson claims that in relation to the background information K-E, P(E) < 1. Hence, the problem of old evidence is dissolved. They dismiss the problem of old- evidence as trivial as it confuses normative account of confirmation with a descriptive account of confirmation. In a descriptive account of confirmation, exclusion of certain background information which an agent possesses might be an invalid one but not in the normative account.

Though C.F.S has a certain kind of strong intuitive appeal, it consists of much deeper problem which even amounts to rejecting the subjective interpretation of BCT. The tone and the style of C.F. strategists is of a logical Bayesian than of a subjective Bayesian. A key point of C.F.S. is that the logically dependent information ought to be excluded from background information. Here it seems, they are guided by a normative idea about what has to be excluded or included. What is the basis of such a normative idea which demands such exclusion?

One way of understanding this is that such exclusion is an acceptable or prevalent practice and the practice can be defended by appealing to reason. Certainly exclusion of certain irrelevant information from K is acceptable as a prevalent practice so it can be adopted as a norm by providing adequate reasons. But there is no proof of acceptable practice which excludes logically dependent information from background knowledge. A reflection of our own practice does not provide any proof that practitioners intentionally/ unintentionally excludes logically dependent information. Howson may respond that proof for such practices is not necessary as there is a sufficient reason to assume that exclusion of logically dependent information is necessary for capturing the confirmation relation in exact way. I think, a deep problem would reveal when we analyse those sufficient reasons.

One reason could be that the information which is logically dependent on H and E may prejudice our conception/understanding of the confirmation relation between H and E. When one says that certain factors prejudice our understanding of confirmation, we suppose that there is an objective confirmation relation and there is only one single correct understanding of confirmation which exists independent of our understanding of confirmation. If prejudice is the reason for the exclusion of logically dependent information, then one has to subscribe to objective/logical theory of confirmation. But Howson dismisses objective/logical theory of confirmation as merely "an ancient habit of thought that dies hard" (Howson, "The Old Evidence Problem", (550). In the subjective framework, there is nothing like a prejudicing factor; all of them are relevant information.

Another possible reason for the exclusion of logically dependent information is that it is the only way to solve the problem of old evidence. Such response is no matter of interest as it is a mere *ad hoc* solution. Moreover, to show that the *ad hoc* solution is necessary, one needs to show that no other solution is possible. A solution would not be *ad hoc* only because it is formulated in response to a particular problem. But it would be *ad hoc*: 1) if it lacks an intuitive appeal or 2) if it is not an account which is derived or extended from basic principles of the framework or 3) if it lacks sufficient reasons to be valid other than solving a particular problem. Sufficient reason/intuitive appeal is necessary for a solution otherwise there is no reason to think that it would be compatible with the other aspects of the theory.

C.F.S. (the proposal to exclude the logically dependent information) is certainly an *ad hoc* solution because it is counter intuitive and incompatible with the subjective framework. It is counter-intuitive because often it is the case that logically dependent information is considered as crucial in capturing the relation of confirmation. In the previously discussed example, intuition says that E would not be the evidence of H if we know that H is false even if H entails E because confirmation means confirming the truth of H. I think, in an extreme way, we can say that none of the relevant background information is independent of H and E. Here independence is assumed as probabilistic independence, not logical. Howson does not specify whether he means probabilistic independence or logical independence. Charitably we can assume that it is logical independence. If probabilistic dependence is an essential feature of K then there is no reason to assume that logically dependent propositions are to be excluded from K. One can even argue that propositions which are logically dependent on H and E are the subsets of the class of propositions which are probabilistically dependent on H and E.

The most meticulous and scathing attack on C.F.S. is that of Charles Chihara. In the above discussion we are discussing the validity of the method called the C.F.S. Chihara' criticism is that C.F.S. cannot even be considered as a method as it is an imprecisely formulated one. Howson' suggestion is that if background knowledge is considered as (K-E) then the problem of old evidence does not arise. But to avoid the problem we need to delete not only E from K, but all proposition that imply E. The problem which Chihara points out is the lack of any reasonable method to identify what are the propositions in K which imply E. One related trouble is that it is almost impossible to identify what are the propositions which K consists of. I think it is more than the practical difficulties of enumerating the beliefs in K. I do not even think that at least we have a fixed class of proposition(s), in a context of confirmation, which works as K. Therefore the question remains: what else should be deleted from K than E?

> Howson attempts to meet this problem by declaring that we should delete from K 'everything in K dependent on' E. But what does that mean? In particular, what does 'dependent on' mean? Are we to delete all propositions in K that are logically dependent on E so that any truth-functional proposition in which E occurs as a component is to be deleted? But that clearly will not be sufficient to yield a stock of background beliefs of the sort Howson wants. For there may be many propositions in K that are logically independent of E (in this sense) but which make E very probable just the same. (Chihara, "Some Problems for Bayesian Confirmation Theory", 553)

"Perhaps Howson would want to delete from K everything that is causally dependent on the observations and/or experiments that gave rise to our coming to know E. But this would make coming up with a specific stock of background beliefs an extremely messy and difficult task." (554*).* The point is, if this sort of defence of Bayesianism is to work, one needs a reasonably clear procedure or rule to determine the required stock of background information.

III.3.Earman's Tentative Suggestions

III.3.1. Old evidence too is uncertain

Earman's one suggestion is, refuting the claim that P(E) =1 in the case of old- evidence. When the value of probability is 1, it implies that we are absolutely certain about the proposition or event. Earman's argument is that except in the case of tautology, the assignment of value 1 is irrational. According to Earman, P(E) can be .999 but not 1. When P(E) = .999, it can be shown that P(H|E) > P(H) in the case of old- evidence. Thus the qualitative problem of old evidence is resolved.

Objections:

When we hold that P(E) cannot be 1, in the case of old- evidence, it implies that none of the observation statements can be held as true or false. There are many cases where an agent is absolutely certain about the correctness of observation; therefore, it is counter-intuitive to hold that the agent is not rational to hold E as true or false. Earman's position cannot explain the wide spread use of truth/ falsity to observation statement.

III.3.2. Quantitative problem of old- evidence

But the major difficulty with the suggestion is not regarding the assumption of value of P(E). Earman himself pointed out that even if the assumption is granted, it does not solve the problem; rather, it makes the problem adopt a quantitative form. Suppose P(E)= .999 then in the case of deductive evidence, evidential support registered by d, r measures of BCT would be too minute to be significant. But history of science tells that often old- evidence provides high degree of evidential support. So the problem of old- evidence still remains in the quantitative form.

> For it remains true that as pr(E) approaches 1, the degree to which E can confirm anything becomes vanishingly small. The more confident we become in our evidence, the less it can be evidence. This quantitative version of the old-evidence problem is not restricted to the standard relevance measure (d). It also infects the ratio and log of ratio based variants mentioned above. As John Earman points out, the problem also infects Haim Gaifman's confirmation measure (l-pr(H))/(l-pr(H/E)). What these measures all have in common is that they ultimately root confirmation in the contrast between pr (H) and pr (H/E). As pr(E) nears 1, this contrast disappears. (Christensen, "Measuring Confirmation", 439)

III.4. Fitelson's modification of Earman's proposal

Fitelson's point is that Earman prematurely dismissed his own proposal to solve the problem of old-evidence. Earman had suggested that as proposition 'E' is a contingent empirical proposition, it is irrational for a subjective Bayesian to assign the value of 1 to P(E) . Obviously it dissolves the paradox as formulated by Glymour, as it disputes the fundamental assumption of the problem that P (E)=1. And the problem arises only when P (E)=1, so it dissolves the problem rather than solving it.

But Earman himself dismissed the solution as inadequate as the solution only makes explicit the sharpest aspect of the problem: quantitative problem of old-evidence. Even if we assume that $P(E)\neq 1$ but only ≈ 1 the problem remains in quantitative form, because the support by E is still insignificant though in practice E is considered as providing significant support even when $P(E) \approx 1$

Fitelson ("Earman on Old Evidence and Measures of Confirmation", 3) formulates Earman's dismissal of the solution as follows:

Theorem: 2

If H $\models$ E and P (E) $\approx$ 1 then c (H, E) $\approx$ 0

Fitelson qualifies Earman's dismissal of the solution as pre-mature and hails Earman's proposal as a potential one to solve the problem. Fitelson objects to Earman on the basis that Earman's dismissal consider only d, (difference measure) and 'r' (relation measure) as the measure of confirmation. If 'd'and 'r' are the only measures of confirmation then Earman is right in saying that the solution makes the problem remain in the quantitative form. When P(E) $\approx$ 1, d (H,E) $\approx$ 0 and r (H,E) $\approx$ 0.

Fitelson ("Earman on Old Evidence and Measures of Confirmation", 3) formulates the claim as follows:

Theorem: 3

If H $\models$ E and P (E) = I-ε then

d(H, E) $\leq \frac{\varepsilon}{\mathrm{I}-\varepsilon}$

Theorem: 4

If H $\models$ E and P(E) = $1-\varepsilon$ then

r(H| E) = $\frac{1}{\mathrm{I}-\varepsilon}$ (for small ε)

Fitelson's point is that Earman's dismissal of the solution is based on theorem: 1 but there are other measures of confirmation like 'l' and 's' which do not satisfy the theorem 1.

Therefore, if we adopt ‘l’ or‘s’ as the measure of confirmation, then the quantitative problem of old-evidence is solved. And Fitelson further argues that though ‘s’ measure does not yield the quantitative problem of old-evidence it cannot be considered as adequate measure of confirmation as it violate one of desideratum of confirmation. Thus Fitelson argues for the adoption of ‘l’ measure and claims that the following theorem, theorem 5, (“Earman on Old Evidence and Measures of Confirmation”, 3) solves the problem of old-evidence.

Theorem: 5.

Even if H⊢E and P(E) ≈ 1, *l*(H,E) can be arbitrarily large.

Fitelson’s point is that following constraints or conditions or procedures to be satisfied would generate a probability model in which H⊢E and P(E) ≈ 1, but *l*(H,E) can be large. Fiteslon is generating certain models in which certain values are constrained yet satisfies the probability axioms. And his point is that in this constrained spaces, which Fitelson called as probability spaces or probability models, the desired result can be derived. It can be derived that *l*(H,E) is large while certain constraints and the assumptions are held.

On the other hand Christensen one who formulates the ‘S’ measure holds that ‘*l*’ measure yields quantitative problem of old-evidence and ‘S’ measure is the only one which is completely immune from the quantitative problem. It is interesting to note that Fitelson’s defense of ‘*l*’ measure is very limited one. Fitelson only says that *l*(H|E) can be larger even when P (E) ≈1 . While attacking ‘d’ and ‘r’ measure, Fitelson’s claim is that d (H|E) and r (H|E) are necessarily small when P(E) ≈1. But in the case of ‘*l*’ measure, his claim is that only *l*(H|E) can be large. In some cases where P (E) ≈1, *l* (H|E) can be large. Christensen had already given counter-example in his paper "Measuring confirmation", where l (H|E) is small when P (E) ≈1. Christensen argues as follows:

> The two likelihoods compared in Good's measure (l measure) do not converge as pr(E) approaches 1 in quite the same way as pr(H) and pr(H/E) do. Nevertheless, this measure, too, turns out to be infected by quantitative old-evidence difficulties. In cases where E confirms H, the likelihood ratio ranges between 1 and ∞. But when pr(E) is high, and pr(H) is moderate-perhaps the paradigmatic situation in which one discusses evidence, the likelihoods are forced to be so close that their ratio falls almost all the way toward the minimum end of this spectrum. (Christensen, "Measuring confirmation", 440)

Christensen describes the counter example as follows:

> For example, let pr(E) be .99. If pr (H) is .75, the highest value the ratio can have is about 1.04. (To see this, note that in general, pr (E) = pr(H)pr(E/H) + pr(∼H)pr(E/∼H). So in the present case, we get: .99 = .75 pr (E/H) + .25 pr (E/ H). From this, it is clear that as pr (E/H) gets higher, pr (E/ H) gets lower, and that the ratio of pr (E/H) to pr (E/∼H) will be the greatest when pr (E/∼H) is as large as possible that is, when it is 1. Assuming the ratio is thus maximized, we get: .99 = .75 + .25pr (E/ ~H), so pr (E/ ~H) = .96. Thus, the maximum value the likelihood ratio can take when pr (E) = .99 and pr (H) = .75 is 1/.96, which is about 1.04.) The value is even lower if we suppose more uncertainty about H; if pr (H) is .6, the ratio maximum drops to under 1.03. And even if we then allow pr(H) to reach .9, the ratio has a maximum of about 1.1." (Christensen, "Measuring confirmation", 440, foot note 8)

Indeed, there is nothing counter-intuitive in Fitelson's claim. l(H, E) is large only in some cases where P(E) ≈1. Moreover, to solve the quantitative problem one need not show that in all cases**,** c(H,E) is high and also showing that c(H|E) is large whenever P(E) ≈1 is certainly counter intuitive too. But the trouble is about showing that the cases where l(H|E) is large is exactly the cases where the old-evidence confirms the hypothesis. That is apart from showing that l(H,E) can be arbitrarily large one needs to show that, on certain conditions, l(H,E) is large when P(E) ≈1 and old-evidence too confirms H only when hypothesis and evidence satisfy the conditions. That is it has to be shows that sufficient conditions for an old-evidence to be a relevant evidence matches with the sufficient conditions where l(H|E) is large when P(E) ≈1. In the absence of the sufficient conditions, the 'l' measure is more unreliable than the'd' and the 'r' in tackling the quantitative problem of old-evidence. l is more unreliable than the standard relevant measures because 'l' could worsen the situation if it is small when old-evidence significantly confirms H and it is large when it provides only insignificant support. Therefore, unless there is a correspondence between the cases where 'l' being large and the cases where old-evidence is relevant evidence, the Bayesians are not in improved situation.

III.5. Maher's Reinstatement of the Problem

> Nevertheless, one does occasionally meet with attempts to defend the simple Bayesian analysis against this objection (the problem of old evidence), by saying that evidence never is completely certain, or should not be taken to be completely certain. However, the objection as I have presented it does not assume that evidence ever is or should be completely certain; the key point is rather that there is no inconsistency between E being certain and E being regarded as evidence for H, and the simple Bayesian analysis implies otherwise. (Maher "Subjective and objective confirmation", 153)

III.6. Eells' Intervention

Bayesian solutions to the problem largely rest on the classification of old- evidence. Eells provides a larger cluster of classifications. But I think that these classifications do not lead us to the solution. Key to the solution of the problem of old- evidence is that all old- evidence is not a confirming/ relevant. I think a classification which distinguishes relevant old-evidence and irrelevant old- evidence may help us to solve the problem. The question is, on what basis can such a distinction be made? Eells has made the following classification. Eells says that if a hypothesis is formulated with the purpose of explaining an evidence which is already known (old- evidence: only in the case of old –evidence such possibility exists), then old- evidence is not a confirming evidence. Eells does not make clear whether except in this case, all old- evidence are confirming or not. My contention is that even in the case, where old evidence is a part of the design of the hypothesis, it is a confirming one. Eells' argument is that confirmation by E is already incorporated in the prior probability of H. Eells is right in saying that after the formulation of H, while describing E as an evidence, the scientists degree of confidence does not increase. But I do not think it is sufficient to deny that E is a relevant evidence. My point is that, a piece of evidence being a part of the design of the hypothesis is not the distinguishing mark of a confirming old-evidence/ a disconfirming old-evidence. And historical examples also run against this.

- Eells' Solution

Eells first attempt is to solve the problem of new evidence (ahistorical problem of new evidence). He distinguishes the notion of confirmation as follows:

1. E actually confirming T
2. E is being (actual) evidence in favour of T.

Eells' claim is that both notions are different. For the first category of confirmation, he sticks to the standard definition of confirmation of BCT; and the second category of confirmation, he defines as follows:

> E is, at time t, part of the body of evidence in favour of theory T relative to history H of background beliefs if and only if, at some time prior to t in the history H (of a set of background

> beliefs), the confirmation event took place between E, T, and the state of H (the relevant probability assignment) at that earlier time. (Eells, "Bayesian Problems of Old Evidence", 209)

Then he states that the ahistorical problem of new evidence arises because we mistake evidence E (e.g. bending of light) in an ahistorical situation as an evidence actually confirming E. But the new evidence E in an ahistorical situation is only in the state of being favourable evidence. Old new evidence E satisfies the definition of 'being a favourable evidence' as confirmation event took place in the past.

I think Eells' distinction can be understood as E which does the activity and E which enjoys the status which was obtained as a consequence of the activity. One needs to be clear about the kind of activity E is doing at the time of the event (first) of confirmation. In subjective framework the activity is: E changes the degree of belief of agent on hypothesis H. Eells claims that because the activity which E did on $t_{0,}$ E is an evidence at $t_{1.}$

Suppose a scientist discovers a confirmation relation between E and H at t_0, then evidence is actually confirming H (First category of Eells). Then after a time period, at t_1, without examining the confirmation relation, scientists holds that E is a favourable evidence of T, then it is of E being evidence in favour of T (second category of Eells). So, it is acceptable to hold that E being evidence depends upon the event of confirmation (first category). Here too the distinction can be held only if the agent does not examine the confirmation relation (either through memory or directly). If a scientist examines the confirmation relation (i.e. if he recollects how H and E are related) then yet another actual event of confirmation happens. But if agent does not try to remember the details of how H and E are related but only remembers that E was an evidence of H, then Eells is right in saying that E, in the state of being evidence, is dependent on the first confirmation event.

But I think, if an ahistorical problem has to be significant, the agents need to examine the confirmation relation (through memory or directly) again. If they only remember the fact that ‘E is an evidence of H’, the question of confirmation, specifically ahistorical problem, does not arise. Just remembering a fact that ‘E was an evidence of H’ is not what we intuitively mean by confirmation. In my view the major trouble with Eells is that he assumes that E would be an evidence of H forever, if E confirms H sometime in the past. In practice, there are many cases where evidence ceases to be an evidence of H. Eells’ account cannot explain this.

III.7. Theoretical Dependence of Evidence

Another important attempt to dissolve the problem of old-evidence is the attempt to clarify the concept of evidence. In the project of confirmation, evidence is not a phenomenon, but it is an observation statement. That is, evidence is not a phenomenon but instead a description of a phenomenon. Description or construal of a phenomenon necessarily consists of a theoretical presupposition. Therefore, though the phenomenon is old, the evidence is new as its construal is different in each theoretical frame work. Therefore, in a sense, there is nothing like an old evidence (evidence which is discovered prior to a theory) because theoretical construals are the necessary constituents of the evidence and the change in the theoretical construal implies change of the evidence itself. That is, evidence is necessarily posterior to the hypothesis, since theoretical suppositions of the hypothesis is a necessary condition to formulate the evidence. That means there is no such evidence called ‘old-evidence’, since all evidence is logically posterior to the hypothesis. No evidence of a hypothesis can be formulated prior to the formulation of the hypothesis itself. The confirming evidence of Einstein's general theory of relativity is not the observation statement recorded half a century ago (since it is construed in the Newtonian theoretical framework) but it is a new observation statement formulated within the framework of general theory of relativity. So the confirming evidence of Einstein's theory is no more an old one.

Though the above argument has a valuable point, it has certain unintuitive consequences. In the structure of the above argument, one significant notion of comparative confirmation would be inapplicable. We have discussed about three kinds of concept of confirmation: qualitative confirmation, comparative confirmation and quantitative confirmation.

Following are the three kinds of comparative confirmation:

1. $C(H, E_1) > C(H, E_2)$

2. $C(H_1, E_1) > C(H_2, E_2)$

3. $C(H_1, E_1) > C(H_2, E_1)$

Among the three, last two are the significant categories of comparative confirmation. And the last one is even more significant because comparative confirmation of different hypotheses by the same evidence is an important tool to choose a more successful/adequate theory in a domain of investigation. History of science shows that, such a comparison is highly significant in the practice of science. The Wave theory of light gains its pre-dominance over the Particle theory, primarily through its explanation of various phenomena like Poisson spot, interference, which could not have been explained by the particle theory of light.

But those who argue for the theory dependence of evidence nullifies the possibility of this significant way of comparative confirmation since there is no common testing of different theories by the same evidence. But they may counter the point by arguing that the significant comparative confirmation between the hypotheses is not the confirmation by the same evidence, but the confirmation by different evidence about the same phenomenon. But that position concedes that though these (e.g. Perihelion of Mercury in Newtonian and Einsteinian theories) are two evidence, they are minimally related since they are about the same phenomenon. But the question is that how can we determine that the evidence are about the same phenomenon. If two descriptions are about same phenomenon then the two descriptions have a common content which is independent of the theoretical framework. Otherwise it cannot be held that the evidence talks about the same phenomenon. That is, the phenomenon or the world of a theoretical frame work is distinct from the world of other theories.

And if there is a common content (like phenomena) between two evidence, it is legitimate to hold that the common content is the one which relates evidence (a statement) to the external world. This discussion leads to the possibility of dissecting the evidence into two content parts:

1. Empirical content (common content of two evidence): A content part of evidence which is independent of the corresponding theoretical frameworks, which relates the statement to the external world (reality).

2. Theoretical content: A content part which relates the empirical content to the theory. The part which ties up the empirical content to a particular theory.

My point is that in the case of confirmation, what is important is the empirical content of the evidence. The scenario might be more complex than I envisage, and the distinction of evidence into two parts might be a naive venture. And certainly the account is too ambiguous as long as I cannot specify what are the empirical content and theoretical content of a particular evidence. However, my sole point is that different descriptions of the same phenomenon constitute the same evidence in its minimal terms across various theoretical frameworks. And the evidence in its minimal terms is the significant notion of evidence in the context of testing of hypothesis.

But even if my redefinition of evidence as evidence in minimal terms is completely wrong, the position of theoretical dependence of evidence cannot dissolve the problem of old-evidence, it can only limit the scope of the challenge of old-evidence. According to this position, descriptions of the same phenomena are different in each theoretical framework, because theoretical pre-suppositions of each theory are different. One of the significant theoretical pre-supposition is the classification or categorization of objects. And the definitions of category (like planets, stars) are different in each theoretical framework. Moreover, the definitions of basic concepts like space, gravitational force are different in each theoretical framework. But there are theories or laws like Copernicus theory, Galileo's law and Kepler's law which share the same definitions of concepts and the same classifications of object. And it is known that Newton's law of gravitation too shares the same theoretical pre-suppositions of Galileo's and Kepler's theories. Galileo's law and Kepler's law can be deduced from Newton's laws of motion and gravitation. The point is that in the case of theories which same theoretical presuppositions, the problem of old-evidence remains since the same observation report can be adopted as evidence for both theories.

III.8. Garber's solution: Overcoming the Logical Omniscience

Among the various solutions, I consider Garber's attempt highly worth considering. Garber understands that the problem stems from the serious inadequacy of the Bayesian framework. His solution to the problem of old evidence starts from the classification of the problem as historical and ahistorical problem. He considers the historical problem as a very serious challenge and considers that the ahistorical problem can be dealt with some kind of counter factual strategy. In a sense, he considers that the ahistorical problem of old evidence (failure of BCT to capture the time enduring nature of evidence) does not pose any challenge to the foundations of BCT and it demands only a clarification of the BCT framework. Though he hopes that some variant of counter factual strategy can dissolve the ahistorical problem, he is cautious in endorsing the C.F.S as solutions. C.F.S states that when E confirms H "If it had been the case that P (e) $<$ 1, then it would also have been the case that P(h|e) > P (h)" (Garber, "Old Evidence and Logical Omniscience in Bayesian Confirmation Theory", 103). Garber cautions the proponents of C.F.S as follows: "There are, to be sure, some details to be worked out here" (103). With these cautionary remarks, he leaves the ahistorical problem to C.F.S and focuses on what he considers as the heart of the problem: the historical problem.

Interestingly, though C.F.S. claims that their strategy is equally applicable to the historical problem, Garber is not impressed by that and states that C.F.S is quite inapplicable in the scenario of historical problem. But the reason he gives for that seems a bit strange as he is the one who endorses C.F.S for the ahistorical problem. His reason for denying C.F.S in the historical scenario is the following: in the scenario of historical problem there is an actual increase of belief, not a counter factual increase.

But how would it be a challenge or criticism? Nobody, including proponents of C.F.S, say that in the case of ahistorical or historical problem, the increase in confidence is counter factual. All admit that increase in confidence is an actual one. The C.F.S position is that the actual increase in confidence can be explained (some may say 'caused') by certain counter factual positions or counter factual considerations. So possibly what Garber had in mind is that the position of the scientists in historical scenario is not caused by counter factual considerations or approach, but by actual considerations. But such a psychological fact of the scientists cannot be stated in such convincing way as Garber has put it.

But his point could be slightly different from what he stated. Irrespective of whether this is the case or not, there is no reason to assume that scientists adopt a counter factual position in order to explain evidence. Moreover, there is strong reason to assume that in the case of confirmation by old evidence (new old evidence) scientist adopt actual considerations. It is because in the historical situations, explaining the old evidence's existence is the driving reason or challenge of the scientists in their theoretical pursuit. In such a scenario, it might be odd to assume that they worked out a confirmation relation by assuming the non- existence of the evidence. The broader point might be the following: There is no reason or historical facts or information to assume or infer that old evidence's existence and prior knowledge about it seriously trouble the scientists in working out a confirmatory relation. So the argument could be that C.F.S is not an adequate explication of the practice of science.

The last point I would like to make is that Garber's resistance to C.F.S. in the case of historical situation is broadly based on the idea that there is an actual event of confirmation and any solution is supposed to account for some elements in the actual situation and C.F.S. fails to do it. I agree with Garber completely on this point and I add that the same consideration would refute his own endorsement of C.F.S in a ahistorical situation because in the ahistorical situation too there is an actual event of confirmation.

Garber presents his ingenious solution to the historical problem by closely following the idea that a theoretical account must have close proximity with the actual situation, which C.F.S. does not have. The problem is that scientists know the evidence even at time t_0 but consider the evidence as relevant only at time t_1. What is the reason for this change of mind (considering evidence as relevant which had been considered as irrelevant evidence)? Otherwise, the question is, what increases the scientist confidence in H at t_1 by considering an old evidence E. I think the obvious answer would be that the scientist discovers a relevant/logical/probabilistic relation between H and E at t_1, which was unknown to the scientist at time t_0. For simplicity Garber assumes that it was a logical relation. That is, at time t_1 scientist discovers that H entails E which (s)he fails to see at t_0.

So the discovery of the logical relation is the reason behind the actual increase of scientist's confidence in H. So in the problem of old evidence the discovery of logical relation works as the basis of confirmation. That is, the problem of old evidence is peculiar because in all other cases the discovery of logical relation between H and E comes along with the discovery of E itself. Therefore we analyse the impact of the discovery of logical relation and the discovery of E together. And the impact is expressed by the quotient $\frac{P(E|H)}{P(E)}$ or by similar measures.

But the problem of old evidence brings a situation where the impact of the discovery of logical relation and the discovery of the evidence needs to be analysed separately. The trouble with the Bayesians is that they don't have the mechanism to analyse the impact of the discovery of the logical relation and the discovery of the evidence separately. Here it is easy to see why C.F.S. is inadequate in addressing the problem of old evidence. The specific trouble is the inadequacy of BCT to analyse the two impacts separately. But C.F.S does not address this inadequacy; instead it creates a hypothetical situation where both impacts are together.

Garber's main point is that we do not have the mechanism to analyse the confirmational impact of the discovery of logical relation. Discovery of a factual statement can be represented in Bayesian framework. But discovery/learning of a logical relation is something which is unthought-of in the Bayesian framework before Garber. It is because Bayesian assumes that if agent knows H then (s)he already cognizes what are the consequences of H. In other words, learning H is learning all consequences of H. Knowing, as per definition, is believing H as true with adequate justification. So if we consider H alone (H as a conjunction) then one of the consequences of H is the conjuncts of H. So it is legitimate to demand that in knowing/learning H, an agent knows or learns what are its conjuncts and the truth of its conjuncts with a minimal justification.

Of course, one must be aware while predicting 'all ravens are black' that (s)he is affirming the truth about each and every raven, though (s)he may not know what are those particulars. That is, the agent must have envisaged the logical possibility. But the situation become complex, when we add a number of auxiliary hypothesis to the evidence. It is not because the infinite logical possibilities may come up. Of course, even in the case of single hypothesis, infinite possibilities would come up. E.g. consider a mathematical hypothesis: the product of two even number is an even number. Here also logical possibilities are infinite. But the possibilities of combinations are finite or single (by possibilities of combinations, I mean the pattern). So the trouble is that if auxiliary hypotheses are added, different kinds of evidence may come up. If H alone is the case, then all consequences of H are confirmatory; but that would not be situation of the consequences of H.A.; the consequences could be confirmatory and disconfirmatory or neutral.

Logical combination can happen in different ways. Before adding the auxiliary hypotheses, conjunction or simplification may be the only rule of combinations or rule to derive consequences. If auxiliaries are added, the rules of deduction could be numerous and the number of times of applications of rules could be large. So Garber's position is that it is practically impossible or beyond human power to conceive what are the consequences of a hypothesis or a theory. I agree that the problem of logical omniscience arises here.

Garber's point is that discovery of some kind of logical relation between H and E after the discovery of E is the distinguishing mark of old evidence. He attempts to establish it by distinguishing the old evidence which are relevant (confirming/ disconfirming) and which are not relevant. All old evidence are not relevant. Among the old evidence only some are confirmatory or relevant. We are identifying old evidence in relation to each particular hypothesis. That means evidence which is old evidence to H may not be a piece of old evidence to H_1 even if H_1 is the competing hypothesis. Garber points out that the particular evidence E would not be relevant old evidence to H only if H is formulated with the purpose of explaining E. That means explanation of E should not be part of the design of the formulation of H. If H is formulated with the purpose of explaining E then E would not be confirming evidence of H. It is because the confirmational impact of E is already included in the prior probability. That is, prior probability of H already includes E. In other sense, the prior probability of H and the confirmational impact of E cannot be distinguished. So Garber's distinction of relevant and irrelevant old evidence is based on the point whether E, the old evidence, is the part of the design of the formulation of H.

Thus, Garber's point is that $P(H \mid E.K) = P(H \mid K)$ does not pose any challenge to BCT if E is the part of the purpose of the formulation. Thus, Garber zeroes down to the exact challenge of old evidence. Relevant old evidence which poses challenge to the Bayesians is the one which is known to the agent/scientist/subject prior to the formulation of hypothesis and its explanation is not the part of the purpose of the hypothesis's formulation.

That is, while formulating hypothesis, agent is not aware whether the confirmation relation/logical relation exists between H and E. That's why, while formulating hypothesis scientist considers E as neutral evidence which does not have any kind of confirmational impact of H. And I think, we could even say not only that the scientist does not have a kind of well worked out explanation or confirmation relation between E and H, but also that scientist does not even think about possible relation which could be worked out later. So Garber's position is that the challenge comes in the cases where agent knows E but does not know the logical relation between H and E. That is, the agent knows E but does not know it as a relevant evidence of H. And it is a challenging case because though E is a relevant evidence in practice $P(H|E.K) = P(H|K)$ because $P(E|H) = 1$.

In order to identify the exact problem of the Bayesian frame work, he takes a cue from the distinction of relevant and irrelevant old evidence. According to Garber, what updates the probability of H in the case of old evidence is the discovery of the logical relation between H and E. To calculate the posterior probability of H, hypothesis needs to be conditioned upon the logical relation between H and E not merely upon E. Thus, he touches upon the exact problem of the Bayesian framework. The BCT has to revise the posterior probability of H as the probability of H conditional upon the logical relation between H and E.

Garber's point is that logical omniscience is the core point of the Bayesian frame work which generates the problem of old evidence. The sole criterion of a Bayesian to evaluate a rational belief is coherence. And the necessary condition of coherence is the axioms of probability. If an agent violates the probability axioms his/ her sets of beliefs are incoherent. One of the axioms of probability P (T) =1; T is tautology. BCT's position is that if P is to be coherent then the subject 'S' must preclude the possibility that an agent accepting bet against the logical truth. Accepting a bet which is bound to lose is equivalent to the violation of probability axioms. That is, suppose an agent knows the propositions H and E, and assumes $H \vdash E$ then to be coherent, rational agent should assign the $P(E) = 1$.

That is, in other words, the agent has to assign the value 1 or 0 to the logical relation $H \vdash E$. $P(H \vdash E)$ should be either 0 or 1. In practice $P(H \vdash E)$ could be between 0 or 1. And if we assume that $P(H \vdash E) = 0$ or $P(H \vdash E) = 1$ that means we are certain about the logical relation between H and E. That is, either we know it does entail or does not entail. Garber's point is that it could be the case as we are uncertain about a logical relation. That is our rational degree of belief of a logical relation could be between 0 and 1. That is, $0 < P(H \vdash E) < 1$. Notion of coherence in Bayesian framework demands us to assign probability value 1 to any logical truth. If we do not, we violate the axioms of probability and thus the conditions of coherence. "Coherence seems to require that S be certain of (in the sense of having degree of belief one in) all logical truths and logical entailments."(104). "….Bayesian account makes it *irrational* to be anything but logically omniscient. The Bayesian agent who is *not* logically omniscient is incoherent, and seems to violate the only necessary condition for synchronic rationality that Bayesians can agree on" (105).

Through the problem of old evidence Garber points out the larger problem of Bayesian frame work as follows: A practicing agent is often ignorant about logical truths or logical relation. Often, though an agent knows two propositions, (s)he may not be aware of any kind of logical relation existing between the two. And later, one may discover the logical relation between the two propositions and the discovery/learning of the logical relation may update the agents' degree of beliefs on both propositions. So having an account of learning of logical truth is necessary for capturing certain features of the practicing realm. More importantly, learning of logical truth or logical proposition has close and high impact on the learning of empirical truth or proposition. In the case of old evidence, the reason for updating the probability is the learning of logical truth that $H \vdash E$. So Garber's point is that without an adequate account of learning of logical truth, we cannot have an account of empirical learning too.

So Garber explores the possibility of having an account of logical learning which is consistent with the Bayesian principles. Garber also points out that there is a deeper problem with the notion of logical omniscience which does not allow logical learning. Garber also points out that the notion of logical omniscience in Bayesian framework brings back the dread which the analytic synthetic distinction had. The notion of logical omniscience in the Bayesian framework points out that the analytic propositions which include logical truth are substantially different from the synthetic propositions. In the BCT frame work, due to logical omniscience it is irrational to be ignorant about logical truth but quite natural and legitimate to be ignorant or uncertain about empirical truths. "This is an asymmetry that smacks of the dreaded analytic-synthetic distinction. But scruples about the metaphysical or epistemic status of that distinction aside, the asymmetry in the treatment of logical and empirical knowledge is, on the face of it, absurd. It should be no more *irrational* to fail to know the least prime number greater than one million than it is to fail to know the number of volumes in the Library of Congress" (105). "The project, then, is clear: if the Bayesian learning model is to be saved, then we must find a way to deal with the learning of logical truths within the Bayesian framework" (105). So Garber's point is that the solution to old evidence is that the Bayesian should be able to formulate an account of logical learning, an account of learning of logical truth.

Garber introduces two models of logical learning:

1) Conditionalisation model and 2) Evolving probability model.

As per Conditionalisation model posterior probability of H after learning about a particular logical relation obtaining is equal to the probability of H conditional upon that logical relation.

Garber's point is that evolving probability model is an inadequate model to be considered as a model of logical learning. And the main strong objection to the model is that it does not provide any algorithm for the belief change in the process of learning. That is, Garber is critical about evolving probability model on the basis of its failure as an account of learning in general. It is not an adequate one even as an account of empirical learning. That is, even after the learning of an empirical truth, it does not provide any robust procedure for change of belief. Certainly it provides some direction or suggestion for the change of belief in view of learning an empirical truth. But the guidance for changes is too insufficient to be qualified as a procedure for change. And as we know, wide range of applicability of Bayesian account of learning and its success mainly rest on the premises that Bayesian account (through its conditionalisation model) accurately defines or lays down the procedures to pin point change of belief in view of learning. In addition, an adoption of an evolving probability model instead of conditionalisation model is an undoing of the greatest advantage of the Bayesian account of learning. Therefore, Garber sticks to the point that an adequate Bayesian account of logical learning can be made possible only through the conditionalisation model.

Garber explores the possibility of using conditionalisation model as a model of logical learning. But the immediate difficulty which the conditionalisation model faces in the case of logical learning is its rational constraints. Conditionalisation model is not a mere account of our belief changes; instead, it is about how we ought to change our belief. In a sense, Bayesian account of belief and belief change is not a mere descriptive account; it is a normative account of belief and belief change. It stipulates how we ought to have a belief in a proposition and how we to change our belief in that proposition in the process of learning. There could be an irrational belief and irrational belief change. Bayesian mechanism is not supposed to account for such changes. It is not to be in accordance with such changes. The fundamental premise of a Bayesian account is that it provides a criterion to distinguish rational beliefs and belief changes from irrational one.

Garber turns his attention to conditionalisation model as he finds the evolving probability model as unpersuasive. But even in the conditionalisation model, it is hard to make sense of a probability function which assigns the value which is not 0 or 1 to logical truths. Logical truths are the one, which is true in all possible state of the world. But if we assign a value which is less than one then an agent is free to bet on the logical proposition that the logical proposition is false. Certainly we know that (s)he will lose the bet. That means the probability function is incoherent. So to be coherent, we ought to assign probability value 1 to all logical truths. So Garber points out that without revising the notion of coherence, conditionalisation model too is grossly inadequate as a model of learning of logical truth.

A model of logical learning cannot be made sensible with the standard notion of coherence. Therefore, Garber sought to revise/weaken the notion of coherence in order to facilitate an adequacy to the model of logical learning. The standard notion of coherence is relativised to all logically possible worlds. But as we know, in inferential procedure or in testing of hypothesis, an agent does not conceive of all possible worlds. In testing a hypothesis, agent does not conceive of all possible hypotheses regarding a particular subject matter, all possible evidence and all possible confirmation relation between each hypothesis and evidence. Therefore, Garber argues that coherence ought to be relativised to the context of human agent where an agent is concerned with only a limited set of hypotheses and evidence which are known to him or her or which are selected to examine.

On this basis Garber envisages two conception of Bayesianism: Global and local.

> On this conception, what the Bayesian is trying to do is build a global learning machine, a scientific robot that will digest all of the information we feed it and churn out appropriate degrees of belief. On this model, the choice of a language over which to define one's probability function is as important as the constraints that one imposes on that function and its evolution. (110)

But for Garber, local Bayesianism is the conception which is compatible to the practices of science;

> (O)n this model, the Bayesian does not see himself as trying to build a global learning machine, or a scientific robot. Rather, the goal is to build a hand-held calculator, as it were, a tool to help the scientist or decision maker with particular inferential problems. On this view, the Bayesian framework provides a general formal structure in which one can set up a wide variety of different inferential problems. In order to apply it in some particular situation, we enter in only what we need to deal with in the context of the problem at hand, i.e., the particular sentences with which we are concerned, and the beliefs (prior probabilities) we have with respect to those sentences. (111)

Rather than the incompatibility with scientific practice, the major problem of global Bayesianism is the supposition of "…*ideal language of science,* a maximally fine-grained language L, capable of expressing all possible hypotheses, all possible evidence, capable of doing logic, mathematics, etc. In short, L must be capable, in principle, of saying anything we might ever find a need to say in science" (110).

Local Bayesianism is only concerned about problem relative to a language: a system through which we can express a group of hypothesis and evidence which we know. The major step of the problem of relative language is that all hypotheses and evidence are atomic sentences. That is they are considered as 'unanalyzed whole'. Certainly hypothesis and evidence too do have a structure. Those structures of sentences are necessary to determine the meaning and our degree of belief of the proposition. But those structures are not necessary for the Bayesian mechanism. Bayesian mechanism is only interested in the procedure through which we update our prior probability into posterior probability. And the mechanism is not concerned about the way in which we arrive in the prior probability. The structure of the propositions are about determining degree of our belief (probabilities) in proposition (H and E) but the structures of proposition do not play any role in the updating of belief. By characterizing the problem–relative language as consisting of only atomic sentences, Garber open up a way for the model of logical learning.

In this language, not only H and E but also any logical or probabilistic relation between H and E too are considered as unanalyzed whole. That is sentences like ‘H⊢E’ though have a structure are considered as atomic statements:

> But whatever *extrasystematic* content we give sentences like "h_i⊢ $e_{i,}$" in the context of our problem-relative investigation we can throw such sentences into our problem-relative language as *atomic sentences,* unanalyzed and unanalyzable wholes, and submerge whatever content and structure they might have, exactly as we did for the h_i, and e_i. (112)

Garber’s basic idea is that we learn a logical relation H⊢E like we learn H and E. In BCT, the concern is not to characterize the relation, which we have captured, but the concern is how it updates our probability of propositions. Garber’s point is that our understanding of logical truths or statements as atomic statements allows us to assign uncertainty to logical truth. In local Bayesianism, logical statements too are treated like any other contingent empirical proposition. Since the basic presumption of local Bayesianism is the consideration of sentence as if it is devoid of any structure, in every world, considered by a local Bayesian agent, logical truths are not necessarily true as the necessity of the statement derives from the structure of the statement and local Bayesians considers statements as devoid of structure. In that sense, it does not violate the coherence condition which is the sole rational constraint of BCT.

But it needs to be emphasised that how the logical possible worlds are different from the worlds of local Bayesianism. Garber's point is that

> …in the context of investigation, we are interested in no other sentences; our problem relative states of the world are easily specified: they are determined by every possible distribution of truth values to the atomic sentences of the local language L'. This amounts to replacing the logically possible worlds of the global language with more modest epistemically possible worlds, specified in accordance with our immediate interests. (113)

But in this language too we assume that though an agent is incapable of conceiving all logical relations between various propositions, (s)he recognizes the tautologies and contradiction within the language. That is, still there are certain constraints within the language system that P(Tautology) = 1, P(Contradiction) = 0. By opening up a way for model of logical learning Garber attempts to formalize the above intuition of local Bayesianism.

Garber assumes a language L*where there are some atomic sentence like H_i and E_i and also new kind of atomic sentences of the form (A⊢ B) where A and B are the atomic sentence of the language L*

> So, if the possible states of the world are identified with possible assignments of truth values to the atomic sentences of L*, … then imposing coherence will fix no degrees of belief with respect to the atomic sentences of L*. There will be coherent P-functions that will allow us to assign *whatever* values we like in [0, 1] to the atomic sentences of the form "A⊢B," however these may be interpreted extrasystematically. (114)

Treating the logical statements as unanalyzed whole explains how probability of logical truths can be given a value between 0 and 1. But Garber found this result as trivial because logical truths are considered on par with contingent empirical propositions. Thus he attempts to explore certain special properties to logical truth. He attempts to characterize the structure of logical truth in its minimal way. Garber claims that even minimal characterization of structure did not harm the project of assigning uncertainty to logical truths.

The minimal structure which may be added to logical truth is the structure of implication. That is, in the problem-relative language we interpret $H \vdash E$ as H implies E. That is, the special property which Garber assigns to logical truth is that it obeys modus Ponens. The rule: Modes Ponens is applicable to logical truth. Garber argues this added minimal structure provides stronger constraints over probability function in the problem-relative language.

The obvious form of the constraints, the added structure brings forth is the following:

(K): $P(B|A.A \vdash B)=1$

Garber redefines it as follows since conditional probability $P(E|F)$ is undefined when $P(F)= O$

(K*): $P(A.B.A \vdash B) = P(A.A \vdash B)$

Garber holds K* as an additional constraint on all probability functions defined on problem relative language or that is defined in the local Bayesian frame work and Garber shows this yields following results which are desirable properties in a model of logical learning. He works out various theorems on the basis of the assumption that P is a probability function on L* (problem relative language) and P satisfies (K*). Following are the some of the theorems which considers as constraints on assigning probability value to logical truths.

Theorem 6

If $P(A\vdash B) = 1$, then $P(A \supset B) = 1$ and $P(B|A) = 1$, when defined.

First Clause:

If $P(A\vdash B) = 1$, then $P(B|A) = 1$

Proof:

$P(B|A) = \frac{P(B.A)}{P(A)}$

$= \frac{P(B.A.A\vdash B)}{P(A)}$ when $P(A\vdash B) = 1$

$= \frac{P(A.A\vdash B)}{P(A)}$: K*

$= \frac{P(A.)}{P(A)}$: Since $P(A\vdash B) = 1$

$P(B|A)= 1$

Second Clause

If $P(A\vdash B) = 1$, then $P(A\supset B) = 1$

$P(A\supset B)= P(B|A)$: Definition of Implication

$P(B|A)= 1$: Theorem 1, First Clause.

<u>Theorem 7</u>

P(∼A|∼B . A ⊢B) = 1 when defined.

Proof: $\frac{P(\sim A.\sim B\ .A\vdash B)}{P(\sim B\ .A\vdash B)}$: Definition of C.P

= : $\frac{P(\sim A.\sim B\ )}{P(\sim B\ )}$: The assumption is P(A ⊢B) = 1

= P(∼A|∼B) : Definition of C.P

= P(∼B ⊃∼A): Definition of Implication

= P(A ⊃B): Modus Tollens

= P(B|A) : Definition of Implication

= P(B|A)= 1 : Theorem 1, Clause 1.

<u>Theorem 8</u>

If A and B are truth-functionally inconsistent in L, then P (A. A⊢ B) = 0.

Proof:

P(A.A⊢ B)= P(B.A.A⊢ B) :K*

= P(B.A.): Since P(A⊢ B) = 1

= 0 : since A and B are truth functionally inconsistent.

<u>Theorem 9</u>

P(B| (A ⊢B) & (∼A ⊢B)) = 1, when defined

Proof:

$\frac{P(B.(A\vdash B).(\sim A\vdash B))}{P((A\vdash B).(\sim A\vdash B))}$: definition of C.P

If (A⊢B) and (~A ⊢B) then B is tautology. Then P(B)= 1: Axiom 2

$\frac{P(B.(A\vdash B).(\sim A\vdash B))}{P((A\vdash B).(\sim A\vdash B))} = \frac{P((A\vdash B).(\sim A\vdash B))}{P((A\vdash B).(\sim A\vdash B))}$: since P(B)= 1

$= 1$

Garber's point is that if we treat all sentences as atomic sentences in problem relative language, coherence does not impose any constraints to the probability assignments of atomic sentences of L* .But we have seen that coherence condition along with the condition K* impose reasonable constraints on assigning probability value to logical truths or sentences of form A⊢B. So now local Bayesianism is not devoid of any rational constraints or the exercise of reducing a logical truth to atomic sentences is no more a trivial exercise. The important question to be asked is that whether the constraint (K* in conjunction with coherence) forces us to assign a value 0 or 1 to logical truths or it allows freedom to assign uncertainty to sentences of the form A⊢B. Garber shows that it yields the freedom to assign probability value between 0 and 1 to logical truth (like H ⊢E). That is without giving away the rational constraints of Bayesian model, coherence condition in conjunction K* allows the learning of logical truths.

<u>Theorem 10</u>

"There exists at least one probability function P on L* such that P satisfies (K*) and such that every atomic sentence in L* of the form "A ⊢B" where not both A and ~B are tautologies gets a value strictly between 0 and 1."(117)

From the above theorem we have seen that even for a logical truth we can assign a value which is between 0 or 1. That is for 0<P(H⊢E) < 1 even if P(E) = 1. But such a result a is not sufficient to show that, P(H |H ⊢ E) > P(H). Because the challenge of problem is that BCT cannot show that P(H |H ⊢ E) > P(H) when P(E) = 1. “But luckily it is fairly easy to show that under appropriate circumstances, there is always a probability function on L* (in fact, an infinite number of them) that satisfies (K*) in which, for any noncontradictory e, and for any nonextreme values that might be assigned to P(H) and P(H|H⊢ E), P(E) =1 and P(H|H⊢E) > P(H)” (121)

III.8.1. Analysis of Garber’s solution

In a nut shell Garber’s point is that in the case of old evidence what confirms H is not E but the logical relation between H and E. So to explicate the confirmation by old evidence, we need to conditionalise H upon the logical relation between H and E. And he shows that how the conditionalisation upon the logical relation between H and E can be characterised in a formal frame work. Garber’s attempt is a clear departure from the standard Bayesian position on the point that it changes the variable which is to be conditionalised upon. This departure and its formalization make Garber’s attempt a remarkable one among the various solutions to the problem of old evidence.

However, there are certain fundamental difficulties which Garber’s system faces. In Garber’s system, conditionalisation happens on logical relation between H and E. Such a move clearly presupposes the idea that there exists an objective logical relation between two sets of propositions independent of an agent. Notion of an objective logical relation is quite explicit in his essay ‘Logical Omniscience and Problem of Old Evidence’.

The main thrust of his essay is that knowing the logical relation between H and E brings confirmation of H. Though agent knows the old evidence at one point of time in the past, it remains ineffective in confirmational terms because agent fails to capture the logical relation between the old evidence and hypothesis. So Garber's basic assumption is that confirmation happens when an agent comprehends or capture the relation (logical or probabilistic or causal) which already exists independent of whether any agent perceives it or not.

But such an assumption which is quite legitimate in an intuitive framework runs against the basic tenets of subjective Bayesianism. And such an assumption is a clear adoption of logical Bayesianism. Subjective Bayesianism categorically refutes the idea that there exists an objective logical relation between two propositions. The relation which exists between the two propositions is dependent upon an agent. Basically an agent's argument and its persuasion define the relation between the two propositions. Agent's subjective-rational belief defines the relation between hypothesis and evidence. That is, one fundamental flaw of Garber's system is its assumption of objective logical relation which counters the basic tenets of subjective Bayesianism.

The second fundamental criticism against Garber is that, Garber's project is not on the line of theory of confirmation. The project of theory of confirmation is defining the relation between a piece of relevant evidence and its hypothesis. H-D model argues that the relation is a deductive logical relation and IBE argues that it is an explanatory relation. Probabilitists argue that it is a probabilistic relation and subjective Bayesians argue that it is an updation of our rational - degree of belief about propositions. But Garber's account took curious turn when compared to all such accounts.

Garber's account assumes that the relation which exists between the hypothesis and the evidence is some kind of a logical relation. But instead of characterising the logical relation, his confirmation project characterises the impact(s) of the logical relation. He perceives confirmation relation as something else than the relation which exists between the hypothesis and the evidence. Garber's account characterises how an agent updates his/her belief in H in the light of the consideration of a logical relation. Such move is curious because for a subjective Bayesian the confirmation relation (relation between H and E) is an updation of an agent's belief or it is the defining characteristics of the confirmation relation. But Garber's account views the updation of belief as a consequence of some other relation which exists between H and E. That is, updation of beliefs is not the defining characteristics of the confirmation relation. Rather updation of beliefs is a description of what happens to an agent's belief system once the confirmation happens.

Certainly Garber has an ingenious account in showing that how the conditionalisation upon the logical relation/confirmation relation changes one's degree of belief on a proposition. No matter how perfect that account may be, it cannot be a theory of confirmation as the account only characterises certain effects of confirmation relation (in a person's belief system) not the confirmation relation itself. But one could say that it is an attempt to represent the confirmation relation through characterising one of its consequences. Such position too is problematic since the notion of consequence is employed in a vague sense. As per this interpretation, Garber holds that confirmation relation is a logical relation between propositions. So he defines the confirmation relation in the realm of proposition. Then any consequence of the confirmation relation too has to be characterized in the realm of proposition. But Garber characterizes the consequence of the confirmation relation in another realm which is the realm of beliefs. As long as we cannot define how the two realms are related, it is meaningless to claim that one of the characteristic of a particular realm (updation of agent's degree of belief) is the consequence of a relation which is defined in another realm (i.e. confirmation which is defined as relation between propositions).

Though I differ with Garber's account on certain points, I agree with Garber on the intention of his project. One of the ways to characterise the motive of the project is as follows.: logical relation among the sentences need to be brought forth to explicate its role in the confirmation project. So the question in front of us is that how the concern which Garber attempts to raise can be addressed within the subjective Bayesian framework.

One of the concerns which Garber raises is that the logical relation between H and E imposes certain additional constraints than probability axioms on the updation of belief of H and E. So, at least one needs to minimally characterise the logical relation to bring forth the confirmation picture. Though he comes up with certain additional constraints on belief-change process, which is compatible with probability axioms, for that he adopted a strong assumption of 'objective logical relation'. But the challenge before us is to characterize the logical relation between the relevant propositions in the minimal way.

Rather than the details and fundamentals of Garber's solution, my basic objection to Garber's solution is that it left out the ahistorical problem of old-evidence unsolved. At the outset of his essay "Old Evidence and Logical Omniscience in Bayesian Confirmation Theory" Garber makes clear that his solution is aimed only at solving historical problem of old-evidence, not the ahistorical problem of old-evidence. And he left the ahistorical problem to the C.F.S which he himself criticized as imprecise one. My point is that ahistorical problem of old-evidence is more substantial and fundamental. In one sense, the exact challenge of problem of old-evidence is explicated only in the articulation of the ahistorical problem. The way in which the historical problem is interpreted often camouflages the exact challenge which the BCT faces. In other sense, though the distinction between the historical and the ahistorical problem can be maintained on certain basis, there is no basis for such distinction in the framework of the BCT. Therefore, such a distinction is futile if not misleading for any attempt to solve the problem. Following are my arguments to hold that there is no substantial basis for the distinction in the BCT framework:

The distinction mentioned above holds on the ground that in the historical problem though the agent knows the evidence in advance, agent could not see the (logical) relation between H and E prior to the event of confirmation. But in the ahistorical cases, agent knows both the logical relation and the evidence in advance to the event of confirmation. In a certain sense, the argument for distinction has some appeal. But once again the distinction is based upon the notion of 'objective logical relation'. Consider a case, where an agent considers E as a disconfirming evidence to H, and at a later stage, due to the persuasion of a new kind of argument, he/she changes the opinion of disconfirmation to confirmation. Do we want to say that the agent did not have the understanding of logical relation in the past and at only at present, the agent possesses the knowledge of logical relation? I think, it is not a maintainable position at least in the context of a subjective Bayesian framework. My point is that the moment agent knew about evidence, (s) he possesses knowledge of logical relation between the evidence and the hypothesis. I think, it is naive to assume that agent did not have any kind of knowledge of evidence's (logical) relation to H and later possessed the correct knowledge. Instead, I assume that an agent always possesses a certain kind of knowledge of the relation between E and H but which could be changed at some other point of time.

So, instead of saying that an agent acquires the knowledge of logical relation', I think it is apt to say that the agent makes a rational judgment regarding E's (logical) relation to various other propositions. Once we adopt the line of thinking of subjective Bayesianism, it is hard to see the distinction between the ahistorical and the historical problem of old-evidence. One major point is that historical /ahistorical distinction is based on the point of time of the formulation of hypothesis, which is immaterial to an agent's degree of belief in the event of confirmation. The time point which is relevant to the testing of hypothesis is the time point of confirmation event or time of testing hypothesis. Related to the time point of testing of hypothesis, hardly one can hold the distinction of historical and ahistorical problem.

But still distinct kinds of old evidence cases can be worked out on the basis of change or shift in the agent's rational judgment. On that basis the historical and ahistorical distinction can be reformulated as follows: In ahistorical cases even before the time of testing, agent is aware of the relevance of the proposition E to the hypothesis H and in the testing (s)he only maintain/reaffirm his/her judgment about relevance relation. One reason for the agent is already aware of the confirmation relation is that agent had already tested the hypothesis in relation with E (that is there is prior confirmation event). Another reason could be that a one of the purpose for the formulation of hypothesis is to provide an explanation to the evidence. That is, evidence is the part of the design of the hypothesis. As per my formulation, the testing of newly formulated hypothesis (by old evidence) also comes under the ahistorical problem if agent had a judgment about the relevance relation between hypothesis and old evidence. Historical problem of old evidence is regarding the testing of newly formulated hypothesis where the old evidence is not a driving reason for the formulation of hypothesis. Though in such cases agent knows the evidence in advance but (s)he thought that the proposition (old evidence) is not relevant to the hypothesis. In our knowledge system, we may consider various as irrelevant to a particular hypothesis H. We cannot say that in all such cases agent possess a judgment about relevance relation. But I consider the significant notion of irrelevance is the outcome of agent's judgment. That is, agent evaluate the relation between E and H and decides that E is not relevant to H. we often disregard superstitious explanation of some phenomenon as we consider the supposed E and H are irrelevant.

In both historical and ahistorical cases of old evidence, an agent knows E in advance to the event of confirmation and also possesses a rational judgement about E's relation to H. The only difference is that, in the historical problem, agent had possessed a rational judgement that there is no relevance relation between H and E and in the event of confirmation, (s)he changed the position and claims a relevance relation between H and E. But in the ahistorical cases not only agent had possessed the rational judgement of relation between H and E but also maintained the same judgement in the event of confirmation too.

The proponents of the distinction between the historical and the ahistorical problem hold that only in the case of the ahistorical problem agent possesses the knowledge of the logical relation. But once we denounce the language of objectivism and use the term of 'rational judgment' of an agent instead of the knowledge of the logical relation, we can see that the distinction collapses. That is, unlike in the case of new evidence, in both cases of old-evidence, agent possesses knowledge of E and a rational judgment concerning E's relation to H.

And now the challenge of old evidence can be seen in contrast with the new evidence. In the case of testing of hypothesis by new evidence, agent does not have any prior judgement regarding H and E. In that sense, it is agent's first judgement about E's relation to H**.** But in the case of old evidence, agent had a prior judgement regarding E's relation to H. In the case of old-evidence, challenge is that BCT has to register the change or renewal of judgement on the basis of the same evidence upon which the first judgement happened. And the registering/capturing the second judgement which changes or renews the first judgment is the challenge which old-evidence poses to BCT. Of course BCT's central mechanism: 'Bayes' theorem' is a precise procedure which characterizes belief-change/judgment –change. The problem of old-evidence poses the question of belief-change upon the same evidence on which we had made our judgment in the past. Through Bayes' theorem, BCT can capture only the belief-change by new evidence. But once the evidence registers its support/non-support, BCT does not have any means to alter or renew the registration of the support. Unlike the logical/ objectivist Bayesian, the subjective Bayesian allows for rational change of opinion even in the absence of new additional information. That is, on the basis of new arguments rational agent is entitled to renew or alter its judgment on hypothesis on the basis of the same set of evidence on which prior judgment is based.

It is interesting to note that the objective Bayesianism does not allow for such rational change because, for them there can be only one correct judgment on relation between evidence and hypothesis because there is only one objective logical relation exists between the hypothesis and the evidence. Therefore, if a change of opinion/belief happens on the same set of evidence, it means that the previous judgment/belief was irrational and incorrect one. Therefore, the problem of old-evidence does not pose any serious challenge to the objectivist. That is why the logical Bayesians like Rosenkrantz provide a straight forward solution to the problem of old-evidence.

Since in the Subjective Bayesian framework a change of opinion on the same set of evidence is a rationally entitled one, the problem of old-evidence poses a serious challenge. Therefore, to solve the problem of old-evidence one needs to carefully examine the way rational belief-change occurs on the same set of evidence. Presumably belief-change occurs on the same set of evidence because of the adoption of new argument. New argument prompts an agent to redraw the relation between H and E.

In the absence of a deductive logical relation between two propositions, there can be a rational second opinion on the relation between two propositions. One can legitimately argue that we understand a proposition through its relation with other propositions. That is why our understanding of tautologies always remains the same. But that is not the case with contingent propositions. Whether old evidence in the context of new theory is new evidence or not, it is uncontroversial to hold that evidence affords new understanding in the context of new theory. On the basis of changing of relations, our understanding of evidence also changes. Our understanding of 'perihelion of mercury' on the basis of Newtonian mechanism is radically different from our understanding on the basis of Einstein theory of Relativity. The phenomena of reflection or diffraction acquire an entirely different understanding on the basis of particle theory of light and wave theory of light.

At this point of time, it is worthwhile to recall Garber's ingenious attempt. I call it as ingenious because Garber argues that (at least in the case of old-evidence) hypothesis has to be conditionalised upon the logical relation between H and E, not on E. Since the notion of an independent logical relation is indefensible, I interpret Garber's "logical relation' as 'agent's understanding of evidence'. That is, in confirmation, hypothesis has to be conditionalised upon the agent's understanding of evidence not upon the evidence itself.

I think probability of agent's understanding of E in relation with H is P (H.E). Though an agent is a proponent of a particular hypothesis especially in the case of testing of hypothesis, it is legitimate to assume that an agent is open to the understanding of evidence provided by other hypotheses too.

It is quite plausible to assume that an agent contrasts an understanding provided by a particular theory with understanding provided by another hypothesis and weighs the confirmation of hypothesis. But the case of old-evidence explicates the role of a new component called 'catchall hypothesis'. The catchall hypothesis is the hypothesis which states that all known hypothesis are false. That is, the catchall hypothesis states that true theory would be the one of the hypotheses which would be formulated in the future or which would be never formulated at all.

It is easy to see the role of the catchall hypothesis especially in the case of old-evidence (in the historical scenario). In the historical scenario, old-evidence is the one which is found as hard to be explained by any existing theories. So, the prevalent dominant assumption among the practitioners could be that the correct explanation of the evidence lies with a hypothesis which is yet to be formulated. Evidence which is related to the unexplained phenomena brings forth the role of catchall hypothesis in the event of confirmation.

So once an agent considers the understanding of evidence provided by a particular hypothesis, it is legitimate to consider that his/her consideration of understanding happens in two contexts. Once in the context of the best competing theories which are known; second in the context of the catchall hypothesis and other known theories but which are not the best competitors.

Considering an understanding of E by H means considering the possibility of understanding being incorrect and understanding being correct. In the context of the best competing hypothesis, considering an understanding being incorrect means taking into consideration of the understanding provided by the best competitor as correct. In the context of the rest of other hypotheses, hypothesis considering an understanding being incorrect means taking into consideration of the understanding provided by the catchall hypothesis or other known hypothesis as true.

So our claim is that in confirmation, conditionalisation is not based on evidence itself. So we divide evidence E into three exclusive propositions.

$E= (H.E) \vee (\sim H.E)$

$P(E) = P(H.E) + P(\sim H.E)$

But here we divide $(\sim H.E)$ into two categories set of best competitor and all other hypotheses which are known and which are possible.

H_c= Best Competitor to H.

H_r: all rest of hypotheses which are possible or known.

$P(E) = P(H.E) + P(H_c.E) + P(H_r.E)$

III.9. New Proposal for a solution

My interpretation of Garber's conditionalisation upon logical relation is conditionalisation upon agent's understanding of E provided by H in relation with other understanding. Along with this I would like to incorporate the new Bayesian framework which I have introduced in the first chapter: determining the confirmation of H by comparing the ratio of updation of H and its best competitor hypothesis. In the following section I attempt to characterise Garber's intuition in a different way. In this attempt my concern is to characterize the novel structure of confirmation rather than to formulate a concrete solution to the problem.

Some of the tenets of the novel structure of confirmation are the following:

1. In confirmation of hypothesis, conditionalisation of hypothesis is not based upon the evidence; instead it is based upon agent's understanding of evidence.
2. Probability of understanding of an evidence by a hypothesis is formalized as P(H.E)
3. While the standard Bayesian confirmation employs the contrast between the prior probability and the posterior probability, the proposed structure of confirmation adopts the contrast between the posterior probability of the hypothesis and the posterior probability of hypothesis which is conditionalised upon the understanding evidence.

That is, in novel structure of confirmation posterior probability of hypothesis is one of basic unit of the formalization, not the prior probability. As the notion of prior probability is completely excluded from the framework of confirmation, P(H) is considered as the formal representation of posterior probability of hypothesis. In standard form, the posterior probability is represented P(H|E). But it is doubtful that posterior probability of hypothesis can be represented hypothesis given E is true. But the major reason for the shift is ahistorical problem of old-evidence. In historical problem of old-evidence, the difficulty is that BCT fails to register the increase in probability of hypothesis. But in ahistorical problem, the difficulty is not merely that it fails to register the increase of probability. Instead, in the ahistrocial scenario, there is no increase in probability at all, to be registered. That is, even if we show that in case of ahistorical old-evidence, BCT can register the increase in probability, it does not solve the problem because there is no actual increase in probability. That is, in ahistorical scenario, the problem which we have is not merely that BCT's mechanism (theorems of probability calculus) or its interpretation fails to exhibit the basic intuition/ idea of BCT (basic idea is increase in probability of hypothesis). But the problem is that basic intuition of BCT itself (increase in probability of H) is inadequate to capture the confirmation. I do not suppose that confirmation of hypothesis is not dependent upon the contrast between the prior and the posterior probability. Instead I assume that confirmation is determined by the contrast between the posterior probabilities of various hypotheses which competes with each other. I think that could be the proper lesson one could infer from the challenge of problem of old-evidence.

Moreover, my second point of choosing the posterior probability of hypothesis as the basic unit of confirmation is a necessary consequence of the new approach since the new approach incorporates the notion of the best competitor into the definition of confirmation. Best competitor is defined in relation with evidence. The best competitor which we have defined can be determined only in posterior scenario of testing by evidence. Comparison of different hypothesis in the posterior scenario of testing determines the best competitor of H. Such a scenario also gives us reason to hold that confirmation must be determined on the basis of contrast between the posterior probabilities of the competing hypotheses.

The same method (like betting method) which is used to determine numerical value of the prior probability of hypothesis can be adopted to determine the posterior probability of hypothesis too. And the novel structure of confirmation aimed only to distinguish confirmation and disconfirmation. That is, it does not aim to distinguish relevant evidence from irrelevant evidence. It works on the assumption that the given evidence are relevant to the hypothesis.

But other various definitions are needed to clarify the new structure of confirmation like definition of Best competing hypothesis. And I think that these definitions ensure that only relevant evidences are filtered into the testing of hypothesis in the new Bayesian framework. Though new framework formulated by aiming to solve the problem of old-evidence I hold that, the new structure of confirmation is applicable to all cases of confirmation.

III.9.1. Modified Definition of the Best Competitor

In the first chapter, I introduced the best competitor as the one which predicts/ explain the evidence as closely as the original hypothesis. That is if H_c is the best competitor of H, $P(E|H_c.K)$ has the value closest to $P(E|H.K)$ than any contrary hypothesis.

First Clause

H_c is the best competitor to H regarding evidence E iff

$\mid P(E|H.K) - P(E|Hc.K) \mid \; < \; \mid P(E|H.K) - P(E|Hi.K) \mid$; (where 'i' $\neq$ c is any competing hypothesis in the field of inquiry)[24].

But this definition of best competitor is completely ineffective in the context of old-evidence since $P(E|Hi.K)$ = 1 for any contrary hypotheses. So we need to add another clause to define the best competitor. And it is essential because, the above difficulty is not only faced in the context of old-evidence but also in many cases of new evidence. Even in the case of new evidence also there could be many contrary hypotheses which equally explain or predict the evidence to the same degree. The second clause which I added to the definition of best competitor is applicable only in the cases where first clause of definition is ineffective in determining the best competitor.

As per the second clause, the best competitor is the one whose posterior probability is the closest to the posterior probably of the hypothesis than any other contrary hypothesis. But the second clause also could fail to determine the best competitor since there could be cases like two contrary hypotheses are close to probability of H but one is higher than the original hypothesis and the other is lower. But their differences with the original hypothesis are same. In such cases, I think the reasonable option would be to choose the hypothesis whose value is close to the original hypothesis but lower than it.

Second clause:

$\mid P(H|K) - P(H_c|K) \mid \; < \; \mid P(H|K) - P(\,H_i|K) \mid$; (where 'i' is any competing hypotheses in the field of inquiry and i $\neq$ c).

Third Clause:

[24] Absolute value of $P(E|H.K) - P(E|H_1.K)$

($H.H_i$) must be an inconsistent set. Where H_i is any competing hypothesis to H including H_c .

Fourth Clause:

$P(E|H_i.K) \neq P(E|K)$; where E is not an old evidence. Where H_i is any competing hypothesis to H and H itself.

Certainly still the definition is incomplete but I think it serves the purposes for the present discussion.

Conclusion

Unlike other paradoxes, the problem of old-evidence remains as a crucial challenge to the BCT and especially to the subjective Bayesian confirmation theory. Except Garber's account, all Bayesian attempts aim to challenge the claim that in the context of confirmation, P (old-evidence) $= 1$. In a sense, they are trying to dissolve the paradox by challenging the fundamental assumption of the problem. But all such attempts are unsuccessful because they fail to provide a good reason to claim that P(old evidence) < 1. I too think that, the problem of old evidence can be addressed only by countering the claim that P(old-evidence) $= 1$. But I think the claim can be challenged only by precisely characterizing the relation between evidence and background information. C.F.S (Counter Factual Strategy) is an attempt in this direction. They argue that 'K' should not consist of 'E' when E is evidence. I agree with the C.F.S position, but not with their reasons for the position. Their reasons were too ad-hoc and inadequate. It is legitimate to hold that when an agent forms a belief about a proposition, an agent does not use all the information which is available and relevant to H and E. But how a practicing agent chooses the relevant background information from his/her vast stock of knowledge is a question which is not adequately answered by the Bayesians.

I hope that a sufficient characterization of the relation between K and E can dissolve the problem of old-evidence. I consider that our discussion (in the first chapter) about the definition of irrelevant conjunct would be useful for such a precise characterization, since I consider that ‘E’ is a irrelevant background information when ‘E’ is the evidence. But the notion of a proposition being irrelevant to the background information is different from the notion of being irrelevant to evidence or hypothesis.

Conclusion

In the contemporary discussions, the Bayesian Confirmation Theory is considered as the most adequate theory of confirmation of hypothesis in science and its major reason is its ability to capture various characteristics of confirmation which are prevalent in scientific practice. Though many disagree with the standard Bayesian solution to the paradoxes, it is largely agreed that the Bayesians made an improvement in solving the paradoxes of confirmation. One of the distinguishing characteristics of the Bayesian solutions is that the solutions are precisely formulated or solutions are formulated within a system which is precisely defined. Inference to the Best Explanation (IBE) provides solutions to the paradoxes, which are intuitively appealing but the solutions are not as precisely formulated.

It is interesting to note that none of the standard Bayesian solutions to the paradoxes are not really solving the paradoxes; instead they dissolve the paradoxes, by clarifying the paradoxicality which is attendant to the results. In the case of confirmation of irrelevant conjunction (I.C), the standard Bayesian position is that there is nothing paradoxical in holding that the I.C is confirmed. It is paradoxical only when it is assumed it is confirmed as equally as one of its relevant conjunct (relevant to evidence (E) and background information (K)) is confirmed. And the Bayesians show that though I.C is confirmed, it is not equally confirmed as its relevant conjunct. Same is with the Bayesian approach in the case of raven paradox. Contra-positive instance confirms but not to the same degree of direct instances. The Bayesian solutions never take the intuitive notion of paradoxicality as granted or as sacrosanct. They attempt to problematise the intuition about the paradoxical results and formulate arguments to answer that why the results appear paradoxical. Rather than merely explicating the nature of our intuition and practice, the Bayesian framework is rich enough to problematise, scrutinize and clarify our various intuitive or prevalent notions. Bayesians clarify or scrutinize prevalent notions by comparing/relating the various results or various confirmation relations together.

In the solutions to the paradoxes, the Bayesians relate and compare various confirmation relations and show that certain confirmation relations are weaker than or necessarily weaker than other confirmation relations. Bayesians employ the notion of degree of confirmation for the precise formulation of the comparison of confirmation relations. And the Bayesian's major advantage in working out various results and solutions is the notion of degree of confirmation. My larger point is that the Bayesian mechanism provides the possibility/ opportunity for precise comparison between or relating between various confirmation relations and that possibility provides the insight to scrutinise our various intuitive notions.

The Bayesian mechanism not only allows/enables us for the comparison between various confirmation relations but also enables us to compare the confirmation relation with probability relation explanation and deductive logical relation. The Bayesians relate the confirmation relation defined on the basis of probability to the confirmation defined on the basis of logical relation and explanatory relations. Thus it shows how confirmation is related to other relations like probability, logical relation and explanation relation. Relating/comparing makes our intuition sharper and better.

As we have discussed in the introduction, Hempel's formulation of adequacy conditions are clearly relating various confirmation relation and various logical relation together. IBE model is no exception to that. Lipton's major point is that (best) explanatory relation can be determined only on the basis of set of hypotheses not merely on the basis of the hypothesis and evidence. H-D model was actually stating that how a deductive logical relation is related to confirmation relation. It holds that inductive and deductive relations are conversely related. Hempel's attempt was more meticulous than the H-D model. Hempel works out minimal valid forms of deductive logical relation by relating how the entailment relation is related to the consistency relation and how the entailment relation is related to the equivalence relation. Definition of the concept is formulated by analyzing how confirmation is distinct from or related to deductive logical relation and how the principles/forms of deductive reasoning can be extended/ elaborated to a more general framework. Hempel's ingenious idea of the Development of Hypothesis (DoH) is a clear manifestation of such a move. Carnap's distinction of relevant/absolute confirmation is clearly thinking of how the notion of confirmation is distinct from that of the notion of probability and how the probability relation can be extended to the realm of confirmation. Carnap extended the probability relation to confirmation relation by employing inequality relation between probabilities of a hypothesis. My general point is that all models of confirmation achieved the clarification of the concept by comparing or relating it to various relations and various factors.

The question is that whether the BCT mechanism draws a sufficiently strong relational set of concepts or does it draw a sufficiently adequate relation between various factors. We have seen that Bayes theorem draws a sufficient relation between various factors like prior probability, posterior probability, likelihood and probability of evidence. But absence of standard solution to the problem of old-evidence implies that the BCT did not sufficiently characterize the relation between various factors namely evidence and background information. Till the formulation of problem of old-evidence, the lack of precise relation between the evidence(E) and the background information (K) does not pose any serious challenge, since the probability of evidence is considered as agent's degree of belief of evidence which is rationally constrained. And to form a rational degree of belief in a proposition, it is not required to know the precise relation between the background information and the proposition.

Clarification of a concept is gained through characterizing its relation with various factors[25]. However, subjective interpretation of probability allows the BCT to proceed without forming a precisely characterised relation between E and K. My point is that the Bayesians made an improvement in solving paradoxes and capturing characteristics because the Bayesian mechanism draws relations between various factors and characteristics of confirmation. And this drawing of relation is made possible mainly through Bayes' theorem and various other definitions and rules. But the BCT fails to tackle problems and paradoxes, mainly problem of old-evidence, because the pre-dominant philosophical interpretation, the subjective theory of probability, fails to carry over/proceed with the project of drawing of relations or forming of web relations in reference to the concept of confirmation/probability. Subjective interpretation allows to form a probability of E on the basis of K without even contemplating about the relation between E and K. This is the general nature of concept of probability characterized in the subjective interpretation. Ramsey remarks as follows: "All that I want to remark here is that no one estimating a

[25] This point has striking similarity with unification of model of explanation.

degree of probability simply contemplates the two propositions supposed to be related by it; he always considers inter alia his own actual or hypothetical degree of belief…."(Ramsey, Truth and Probability, 65)

Here Ramsey shifts the emphasis from the relational judgment in determining the probability value of a proposition. As an argument against logical probability, it may have a valid point in the sense that we are not contemplating about all possible relevant propositions, while determining the probability of a proposition. But it is not maintainable to argue that in rational judgment, we are not contemplating the relation between a proposition and certain other relevant propositions while determining probability of the proposition.

Often while we form a degree of belief in a proposition, we may not contemplate its relation to other proposition. My point is as follows: Ramsey is right in holding that while forming a degree of belief in 'P', often we are not contemplating its relations to various other propositions. But Ramsey is wrong in arguing that such beliefs are rational. The fundamental mistake of Ramsey is that he determines rationality of belief/partial belief in the context of formation of belief. It is true that our degrees of beliefs which are considered as rational may not have been formed by contemplating its various relations to other propositions. But my point is that ascription of rationality to a degree of belief cannot be done in the context of formation of degree of belief (context of discovery). We can determine the rationality of degree of a belief only when test/check the correctness/validity of degree of belief which we form. And in the context of testing the correctness of our degree of belief, we do contemplate the proposition's relations to other propositions.

Patrick Maher, a staunch critic of subjective interpretation beautifully articulates the point as follows:

> Suppose, for example, that I claim that scientific theory H is probable in view of the available evidence. This is a statement of inductive probability. If my claim is challenged, it would not be a relevant response for me to prove that I have a high degree of belief in H, though this would be relevant if inductive probability were subjective probability. To give a relevant defense of my claim I need to cite features of the available evidence that support H. (Maher, "Confirmation", 5)

My point is that rationality can be determined only in the context of checking/testing the degree of belief which we already hold. While checking the degree of belief, or proving the correctness of degree of belief, we necessarily contemplate its relation with other relevant propositions.

When we judge or determine a degree of belief in 'P' as a rational, we contemplate P's relation to various other propositions. Certainly one kind of such relevant proposition is evidence. More importantly we also look for other propositions which are called competing hypotheses. I agree with both Ramsey and Keynes that probability of P is rational degree of belief. But rational degree of belief is not determined either on the basis of objective logical relation or on the basis of the way we formulate the degree of belief.

Maher points out that while we test/check the probability of P it is not sufficient to say that my degree of belief in P exactly matches with probability of P, or it is not sufficient to say that I arrived at the probability of P through betting method which precisely measures actual degree of belief in P. For it, we need to show that on what basis we hold the degree of belief is correct rather than showing that it is correct that I hold the degree of belief.

> Suppose now that X wishes to argue that (1)(The probability that Tweety flies is greater than 0.75.) is indeed true. If (1) when uttered by X means (2)(X's degree of belief that Tweety flies is greater than 0.75.) then X could prove that (1) is true by proving that (2) is true; for example, X could show that he is willing to bet at odds of more than 3:1 that Tweety can fly. But we would ordinarily think that this evidence is irrelevant to the truth of (1). We ordinarily think that in order to support (1), X needs to cite features of the available evidence that are relevant to whether Tweety can fly, for example, that Tweety is a bird and that most birds can fly. Thus if (1) uttered by X meant (2), the sorts of arguments that could be used to support (1) would be very different to what we usually suppose. (Maher, The Concept of Inductive Probability, 188)

Logical interpretation, though it fails to establish the logical relation among the propositions, continues to enchant philosophers. One of the remaining interests in logical interpretation is that it determines the probability of a proposition in relation with other proposition. Carnapian system figures out how one statement is related to another statement. Carnap defines this relation between propositions on the basis of overlap of ranges of statements. And overlap of ranges itself determines the probability of a statement. Overlap of ranges, which can be called as relational factor of the Carnapian system, not only relates various propositions on the basis of overlap of ranges, but also relates various relations like deductive logical relation and inductive relations. If an overlap of ranges of two propositions is complete, then it is deductive, if it is partial then it is an inductive relation.

But Carnapian system is too ambitious than it is required. Primarily Carnapian system shows that the overlap of ranges or the relational factor is an objective one. Thus relevance relation became independent of human agent. Thus he was not explicating or defining relevance relation, instead his principles were determining relevance relations. My point is that Carnap is right in assuming that probability of hypothesis can be determined only in its relevance relations with other propositions. But it need not be objectively grounded. Rather than seeking an objective basis to determine relevant relations, my inclination would be to explicate the relevant relations which we employ in our practice.

My position is that there are various kinds of relevant relations which I call as relevant categories which we employ in our practice. A precise analysis can bring forth what are the relations of these relevant categories to a proposition or hypothesis. In the preceding chapter, I discuss about the notion of 'Best competing hypothesis' 'Best competing of H' is one category of relevant proposition. Through my definition of 'Best Competitor', I am attempting to explicate our intuitive notion of relevance relation between hypothesis and its best competing hypothesis. Notion of the catch hall hypothesis is another kind of relevant propositions, which are discussed in the Bayesian literature. In a similar way another category of relevant proposition is 'non-worthy hypothesis/ less plausible hypothesis'. I am not claiming that there are certain hypotheses, which are objectively less-plausible; instead my point is that the agent employs the category of less plausible hypothesis while analysing various hypotheses and evidence relation or probable relations. My concern is to analyse how the category of 'less plausible hypothesis' is related to the hypothesis in concern. Can we work out a relation between P(H) and $P(H_{less\ plausible})$. Do we have reason to hold that the difference between P(H) and $P(H_{less\ plausible})$ is greater than a fixed number like 0.5.

I think that the relation which we have with various categories can be expressed in terms of probability value or probability relation. Certainly various categories of relevant relations can be worked out. And precise analysis can bring forth how they are related to a proposition in terms of probability value and probability relation. I think that it may not be correct to hold while judging a correctness of a degree of belief, we consider all relevant proposition of H. We may not consider all relevant propositions but consider all relevant categories of propositions while checking the correctness of a degree of belief. On the basis of a web of relevance category relation, I think we would be able to determine an interval of probability value of hypothesis.

My point is that subjective interpretation of probability is found as inadequate in capturing the role of relevance relations in determining the probability of a hypothesis. And some of the challenges of the Bayesian confirmation theory, especially the problem of old-evidence stems from the infirmities of subjective interpretation of probability. I consider the challenges to the BCT as serious call for new interpretation, which I would like to call as relational interpretation of probability.

I consider the project of definition of confirmation as an unending task, where improvements are the aim, not the conclusive results or solutions. If a substantial philosophical question is conclusively solved/ answered then it means that the concerned human practice met a dead end and advancement in human enquiries are not or cannot be made. As long as human inquiries and reasoning remain an infinite task, philosophical questions too remain as unresolved and interesting. In a sense, philosophers are bound to ask new question and raise a new criticism to open up new vistas of inquiries and reasoning. And philosophical solutions/improvements are means for better criticism. And the improvement which I have made to the discussion of confirmation is the introduction of the factor ‘best competing hypothesis’ into the structure of definition of the BCT. And I think the notion ‘best competing hypothesis’ would contribute to the formulation of a new interpretation of probability, which I consider as one of the complementary move for the improvement of the Bayesian confirmation theory.

Appendix 1: A

Proofs of Theorems: Chapter I

A.1. Theorem 2

In the case of deductive evidence, $(P(E|H.K) = 1)$

$P(H.X|E.K) > P(H.X|K)$ where X is any statement.

Proof:[1]

1. $P(H.X|E.K) = P(E|H.X.K) \times P(H.X|K) / P(E|K)$
2. $P(H.X|E.K) = P(H.X|K) / P(E|K)$ (since $P(E|H.X.K) = 1$)
3. $P(H.X|E.K) > P(H.X|K)$ (where $P(E|K) < 1$)

A.2. Theorem 3

$P(H|E.K) - P(H|K) > P(H.X|E.K) - P(H.X.|K)$ when H entails E

Proof:[2]

1. $P(H.X|E.K) - P(H.X.|K) = P(E|H.X.K) \times P(H.X|K) / P(E|K) - P(H.X.|K)$: Bayes' theorem

[1] Fitelson "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction", 612.
[2] Earman, *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, 63-65.

2. $= 1 \times P(H.X|K)/ P(E|K) - P(H.X.|K)$: H.X entails E since H entails E

3. $= P(H.X|K) (1/ P(E|K) - 1)$

4. $P(H|E.K) = P(H|K) = P(E|H.K)\times P(H|K)/ P(E|K) - P(H|K)$: Bayes' theorem

5. $= 1 \times P(H|K)/ P(E|K) - P(H|K)$: Since H entails E

6. $= P(H|K) (1/ P(E|K) - 1)$

7. $P(H.X|K) (1/ P(E|K) - 1) < P(H|K) (1/ P(E|K) -1)$: irrelevance of 'X'[3] and $0< P(X|K)>1)$

8. $P(H|E.K) - P(H|K) > P (H.X|E.K) - P(H.X.|K)$: 3,6,7.

A.3. Theorem 4

$d(H.X|E.K) = P(X|H.K)\times d(H|E.K)$ when H as well as H.X entails E

Proof:[4]

1. $d(H.X|E.K) = P(H.X|E.K) - P(H.X.|K)$
2. $= P (E|H.X.K). P (H.X|K)/ P (E|K) - P(H.X.|K)$: Bayes' Theorem.
3. $= P (E|H.K). P(H.X|K)/ P(E|K) - P(H.X.|K)$: Since H entails E
4. $= P (E|H.K). P (X|H.K). P (H|K) / P (E|K) - P(X|H.K). P(H|K)$

   : Since $P(H.X|K) =P(X|H.K).P(H|K)$

[3] X is irrelevant if $(P(H.X|K)=P(H|K)\times P(X|K)$.

[4] Rosenkrantz, "Bayesian confirmation: Paradise regained", 470–471

5. $= P(X|H.K) \{P(E|H.K). P(H|K) / P(E|K)\} - P(H|K)$

6. $= P(X|H.K) \{P(H|E.K) - P(H|K)$: Bayes' Theorem

7. $d(H.X|E.K) = P(X|H.K) \times d(H|E.K)$: Definition of 'd'

A.4. Theorem 5

If E confirms H and X is confirmationally irrelevant to H, E and H.E (relative to K) then E also confirms H.X (relative to K).

Proof:[5]

Assume that K is tautologous.

1. $P(H.X|E) = \frac{P(H.X.E)}{P(E)}$: Definition of Conditional probability.

2. $= \frac{P(H.E) \times P(X)}{P(E)}$: Irrelevance of X. That is X is independent of H.E

3. $= P(H|E) . P(X)$: Definition of Conditional probability

4. $> P(H). P(X)$: in the case where E confirms H

5. So $P(H.X|E) > P(H.X)$: Irrelevance of X

A.5. Theorem 6

If E confirms H, and X is confirmationally irrelevant to H, E and H.E (relative to K) then (provided that $P(X|K) < 1)$: $c(H, E|K) > c(H.X, E|K)$ where 'c' (confirmation measure) may be either

[5] Fitelson "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction", 618-619.

difference measure 'd' or the likelihood ratio measure '*l*' of degree of confirmation. (But, c may not be the ratio measure r, since in cases of irrelevant conjunction we will have *r*(*H, E*|*K*) = *r*(*H&X, E*|*K*). *(*617)

Proof[6]*:*

For the c = *d* case of the theorem (indeed, for all three cases), we assume without loss of generality that *K* is tautologous, and we reason as follows:

1. d (H.X, E) = P(H.X| E) − P(H.X) : definition of d measure
2. $= \frac{P((H.X).E)}{P(E)} - P(H.X)$: definition of conditional probability
3. $= \frac{P((H.E).p(X)}{P(E)} - P(H).P(X)$: irrelevance of X
4. $= P(X)\left\{\frac{P((H.E).}{P(E)} - P(H)\right\}$:definition of conditional probability
5. = P(X). d(H,E) : definition of d
6. d(H.X,E) < d(H,E) : P(x) < 1, d(H,E) > 0

For the c= *r* case of the theorem we will prove that r(H.X,E) = r(H,E) where x is an irrelevant conjunct (hence that r violates the theorem)

1. r(H.X,E) $= \frac{P(H.X|E)}{P(H.X)}$: definition of r
2. $= \frac{P((H.E).X)}{P(E).P(H.X)}$: definition of conditional probability
3. $= \frac{P((H.E).X)}{P(E).P(H).P(X)}$: Irrelevance of X

3.1 $= \frac{P(H|E)}{P(H)}$: definition of conditional probability,

[6] Fitelson "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction", 619-621.

algebra

4. $= r\,(H,E)$: definition of r

Proof of the likelihood-ratio (*l*) case of the theorem (*via reductio ad absurdum,* and using the *d* and *r* results as lemmas).

1. $l(H.X,E) \geq l(H,E)$: [*reductio* assumption]
2. $\frac{P(E|H.X)}{P(E|{\sim}(H.X)} \geq \frac{P(E|H)}{P(E{\sim}H.)}$: definition of l
3. $\frac{\frac{P(H.X|E)P(E)}{P(H.X)}}{\frac{P({\sim}(H.X))|E)\times P(E)}{p({\sim}(H.X))}} \geq \frac{\frac{P(H|E)\times p(E)}{P(H)}}{\frac{P({\sim}H|E)P(E)}{P({\sim}H)}}$: Bayes' theorem
4. $\frac{P(H.X|E)}{P(H.X)} \times \frac{P({\sim}(H.X))}{P({\sim}(H.X))|E)} \geq \frac{P(H|E)}{P(H)} \times \frac{P({\sim}H)}{P({\sim}H|E)}$
5. $r(H.X,E) \times \frac{P({\sim}(H.X))}{P({\sim}(H.X))|E)} \geq \frac{P({\sim}H)}{P({\sim}H|E)} \times r(H,E)$: definition of r
6. $\frac{P({\sim}(H.X))}{P({\sim}(H.X))|E)} \geq \frac{P({\sim}H)}{P({\sim}H|E)}$: $r(H.X,E) = r(H,E)$
7. $1 - P(H|E)\times (1 - P(H.X)) \geq (1 - P(H))\times(1 - P(H.X|E)$: $P({\sim}X|Y) = 1 - P(X|Y)$
8. $d(H.X,E) - d(H,E) + P(H|E)\times P(H.X) \geq P(H).\ P(H.X|E)$: $d(H.X,E) < d(H,E)$
9. $r(H,E) > r(H.X,E)$: definition of r
10. $l\,(H.X,E) < l(H,E)$: $r(H.X,E) = r(H,E)$

Appendix 2: A

Proofs of Theorems: Chapter II

A.1. Theorem 1

(QUANTc): c ((x)(Rx ⊃ Bx), (~Ba · ~Ra) | K) > 0, but very nearly 0.

where measure of confirmation is difference measure (c = d)

d ((x)(Rx ⊃ Bx), (~Ba · ~Ra) | K) > 0, but very nearly 0.

Proof: [7]

1. $P(H| \sim Ra. \sim Ba) - P(H)$:Difference measure
2. $P(H| \sim Ra . \sim Ba) = \frac{P(\sim Ra \cdot \sim Ba\ |H) \times P(H)}{P(\sim Ra \cdot \sim Ba)}$: Bayes' theorem
3. $= \frac{P(\sim Ra\ |\cdot \sim Ba.H) \times P(\sim Ba\ |H) \times P(H)}{P(\sim Ra \cdot \sim Ba)}$:Theorem : $(H.A|E) = P(H|E) \times P(A|H.E)$
4. $= \frac{P(\sim Ba\ |H) \times P(H)}{P(\sim Ra \cdot \sim Ba)}$: $P(\sim Ra\ |\sim Ba.\ H) = 1$. As H entails Ra.Ba
5. $= \frac{P(\sim Ba\ |H) \times P(H)}{P(\sim Ra\ |\sim Ba) \times P(\sim Ba)}$: Definition of Conditional probability
6. $P(H| \sim Ra \cdot \sim Ba) - P(H) = P(H)(\frac{P(\sim Ba\ |H)}{P(\sim Ba)} \times \frac{1}{P(\sim Ra\ |\sim Ba)} - 1)$

[7] Vranas, Peter B. M. “Hempel's Raven Paradox: A Lacuna in the Standard Bayesian Solution”, 548.

7. $= P(H)(\frac{1}{p(\sim Ra\,|\sim Ba)} - 1)$:Assumption 3: $\frac{P(\sim Ba\,|H)}{P(\sim Ba)} = 1$. Probability of object being non- black is independent of the truth of H.

8. $= P(H) \times \delta$ where δ is a minute number : Assumption 1: P(~Ra |~Ba) ≈ 1 because P(Ra |~Ba) is minute. That is given that an object 'a' is a non- black thing, the probability of 'a' being a raven is minute because the world contains overwhelmingly more non-black objects than ravens. i.e. $P(\sim B_a|K) > P(R_a|K)$ or $P(\sim B_a|K) >> P(R_a|K)$. Then P(~Ra |~Ba) ≈1 but < 1. Then $\frac{1}{P(\sim Ra\,|\sim Ba)} \approx 1$ but > 1. Then $\frac{1}{P(\sim Ra\,|\sim Ba)} - 1$ is minute but greater than zero.

9. Therefore $P(H|\sim Ra. \sim Ba) - P(H) > 0$ but ≈ 0.

A.2. Theorem 2

Quantitative claim: Where measure of confirmation is ratio measure ($c = r$)

(QUANTc): $r[(x)(Rx \supset Bx), (\sim Ba \cdot \sim Ra) \mid K] > 0$, but very nearly 0.

Proof: [8]

1. $\frac{P(H|\sim Ra \cdot \sim Ba)}{P(H)} > 0$

2. $\frac{P(H|\sim Ra \cdot \sim Ba)}{P(H)} = \frac{P(\sim Ra \cdot \sim Ba \,|H)}{P(\sim Ra \cdot \sim Ba)}$: $\frac{P(H|E)}{P(H)} = \frac{P(E|H)}{P(E)}$: Bayes' Theorem

3. $\frac{P(\sim Ra|\sim Ba\,.H)}{P(\sim Ra|\sim Ba\,)} \times \frac{P(\,\sim Ba\,|H)}{P(\,\sim Ba)} =$ Definition of C.P & theorem $P(H.A|E) = P(H|E) \times P(A|H.E)$

4. $\frac{P(\,\sim Ba\,|H)}{P(\,\sim Ba) \times P(\sim Ra|\sim Ba\,)}$: $P(\sim Ra|\sim Ba\,.H) = 1$ since H entails Ra.Ba

5. $\frac{P(\,\sim Ba)}{P(\,\sim Ba) \times P(\sim Ra|\sim Ba\,)}$: Assumption 3: $P(\,\sim Ba\,|H) = P(\,\sim Ba)$ because probability of ~Ba is independent of the truth of the hypothesis, H.

6. $\frac{1}{P(\sim Ra|\sim Ba\,)} \approx 1$ but > 1. :Assumption 1. $P(\sim Ra|\sim Ba) \approx 1$ because $P(Ra|\sim Ba)$ is minute. That is given that an object 'a' is a non-black thing, the probability of 'a' being raven is minute because the world contains overwhelmingly more non-black objects than ravens. i.e. $P(\sim B_a|K) > P(R_a|K)$

[8] Vranas, Peter B. M. "Hempel's Raven Paradox: A Lacuna in the Standard Bayesian Solution", 548.

7. $P(H|\sim Ra.\sim Ba) \approx P(H) > 1$.

A.3. Theorem 3

Given Non-triviality, it follows that $q > (1-p) > 0$ and

$$\frac{P(Ba\cdot Ra \mid H\cdot K) / P(Ba\cdot Ra \mid \sim H\cdot K)}{P(\sim Bb\cdot\sim Rb \mid H\cdot K) / P(\sim Bb\cdot\sim Rb \mid \sim H\cdot K)} = [q-(1-p)]/(p\cdot r) > 0.$$

$$\frac{P(Ba\cdot Ra \mid H\cdot K) / P(Ba\cdot Ra \mid \sim H\cdot K)}{P(\sim Bb\cdot\sim Rb \mid H\cdot K) / P(\sim Bb\cdot\sim Rb \mid \sim H\cdot K)} = \frac{l(H,E)}{l(H,E^*)}$$ where l is the likelihood measure and H is 'all ravens are black' and E is an instance (Ba.Ra) and E* is a contrapositive instance (~Ba.~Ra).

Proof:[9]

$$\frac{(P(Ba\cdot Ra \mid H\cdot K] / P(Ba\cdot Ra \mid \sim H\cdot K])}{(P(\sim Bb\cdot\sim Rb \mid H\cdot K] / P(\sim Bb\cdot\sim Rb \mid \sim H\cdot K])} : \frac{l\,(H,E)}{l\,(H,E^*)}$$

$$= \frac{P(Ra|H\cdot K) \times P(Ba|Ra.H\cdot K) / P(Ra|\sim H\cdot K) \times P(Ba|Ra\cdot\sim H\cdot K)}{P(\sim Bb|\, H\cdot K) \times P(\sim Rb|\sim Bb.H\cdot K) / P(\sim Bb|\sim H\cdot K) - P(\sim Bb\cdot Rb \mid \sim H\cdot K)}$$

: Theorem: $P(H.A|E) = P(H|\, E) \times P(A|H.E)$ and $P(\sim Bb|\sim H.K) = P(\sim Bb.Rb|\sim H.K) + P(\sim Bb.\sim Rb \mid \sim H.K)$

$$= \frac{P(Ra|H\cdot K)\times 1 / P(Ra \mid \sim H\cdot K)\cdot p)}{P(\sim Bb \mid H\cdot K)\times 1 / P(\sim Bb \mid \sim H\cdot K) - P(\sim Bb\cdot Rb \mid \sim H\cdot K)} :$$

$p = P(Ba \mid Ra. \sim H.K)$, $P(Ba|Ra.H.K) = 1$ and $P(\sim Rb|\sim Ba.H.K) = 1$)

[9] Fitelson, and James Hawthorne. "How Bayesian Confirmation Theory Handles the Paradox of the Ravens", 25.

$$= \frac{(1/p)\cdot(P(Ra \mid H\cdot K)}{P(Ra \mid \sim H\cdot K)} \cdot \frac{P(\sim Bb \mid \sim H\cdot K) - P(\sim Bb\cdot Rb \mid \sim H\cdot K)}{P(\sim Bb \mid H\cdot K)}$$

$$= \frac{(1/p)\cdot(P(Ra \mid H\cdot K)}{P(Ra \mid \sim H\cdot K)} \cdot \frac{P(\sim Bb \mid \sim H\cdot K) - P(\sim Bb \mid Rb.\sim H\cdot K) \times P(Rb \mid \sim H\cdot K)}{P(\sim Bb \mid H\cdot K)}$$

$$= \frac{(1/p)\cdot(P(Ra \mid H\cdot K)}{P(Ra \mid \sim H\cdot K)} \cdot \frac{P(\sim Bb \mid \sim H\cdot K) - ((1-p)\times P(Rb \mid \sim H\cdot K))}{P(\sim Bb \mid H\cdot K)}$$: $p = P(Bb \mid Rb. \sim H.K)$ and

$(1\text{-}p) = P(\sim Bb \mid Rb. \sim H.K)$

$$= \frac{(1/p)\cdot(P(Ra \mid H\cdot K)}{P(Ra \mid \sim H\cdot K)} \cdot (q-(1-p)) \cdot \frac{P(Rb \mid \sim H\cdot K)}{P(\sim Bb \mid H\cdot K)}$$: (where $q = \frac{P(\sim Ba \mid \sim H\cdot K)}{P(Ra \mid \sim H\cdot K)}$

$$= (1/p)\cdot(q-(1-p))\cdot \frac{P(Ra \mid H\cdot K)}{P(\sim Bb \mid H\cdot K)} \cdot \frac{P(Rb \mid \sim H\cdot K)}{P(Ra \mid \sim H\cdot K)}$$

$$= (1/p)\cdot(q-(1-p))/r.$$: (where $r = \frac{P(\sim Ba \mid H\cdot K)}{P(Ra \mid H\cdot K)}$)

A.4. Corollary 1

$$\frac{P(Ba\cdot Ra \mid H\cdot K) \,/\, P(Ba\cdot Ra \mid \sim H\cdot K)}{P(\sim Ba\cdot \sim Ra \mid H\cdot K) \,/\, P(\sim Ba\cdot \sim Ra \mid \sim H\cdot K)} > 1 \text{ if and only if } q-(1-p) > p\cdot r$$

And, more generally, for any real number s,

$$\frac{P(Ba\cdot Ra \mid H\cdot K) \,/\, P(Ba\cdot Ra \mid \sim H\cdot K)}{P(\sim Ba\cdot \sim Ra \mid H\cdot K) \,/\, P(\sim Ba\cdot \sim Ra \mid \sim H\cdot K)} = s = (q-(1-p))/(p\cdot r) > 1 \text{ if and only if } (q-(1-p)) = s\cdot p\cdot r > p\cdot r.$$

Proof[10]

The first biconditional follows from Theorem 1 together with the obvious point that

$(q-(1-p))/(p\cdot r) > 1$ iff $q-(1-p) > p\cdot r$.

To get the second biconditional, just observe that (for any real numbers),

$s = (q-(1-p))/(p\cdot r) > 1$ iff $s\cdot p\cdot r = (q-(1-p)) > p\cdot r$.

A.5. Theorem 4

Given Non-triviality, both of the following clauses hold:

Theorem (4.1)

If $P(\sim Ba \mid H\cdot K) > P(Ra \mid H\cdot K)$ (i.e. if $r > 1$) and

$O(H \mid Ra\cdot K) / O(H \mid \sim Ba\cdot K) > (p + (1-p)/r)$, then $\quad : O(X|Y) = \frac{P(X|Y)}{P(\sim X|Y)}$

$$\frac{P(Ba\cdot Ra \mid H\cdot K) / P(Ba\cdot Ra \mid \sim H\cdot K)}{P(\sim Ba\cdot \sim Ra \mid H\cdot K) / P(\sim Ba\cdot \sim Ra \mid \sim H\cdot K)} > 1.$$

Theorem (4.2)

If $P(\sim Ba \mid H\cdot K) \leq P(Ra \mid H\cdot K)$ (i.e. $r \leq 1$), but either $P(\sim Ba \mid K) > P(Ra \mid K)$

or $P(\sim Ba \mid \sim H\cdot K) > P(Ra \mid \sim H\cdot K)$ (i.e. $q > 1$), then

$$\frac{P(Ba\cdot Ra \mid H\cdot K) / P(Ba\cdot Ra \mid \sim H\cdot K)}{P(\sim Ba\cdot \sim Ra \mid H\cdot K) / P(\sim Ba\cdot \sim Ra \mid \sim H\cdot K)} > 1$$

[10] Fitelson, and James Hawthorne. “How Bayesian Confirmation Theory Handles the Paradox of the Ravens”, 25-26.

Proof: [11]

Assume Non-triviality.

Both parts of the theorem draw on the following observation:

Theorem 1 tells us that $q > (1-p) > 0$ and

$$\frac{P(Ba \cdot Ra \mid H \cdot K) / P(Ba \cdot Ra \mid \sim H \cdot K)}{P(\sim Bb \cdot \sim Rb \mid H \cdot K) / P(\sim Bb \cdot \sim Rb \mid \sim H \cdot K)} = (q-(1-p))/(p \cdot r).$$

So $\frac{P(Ba \cdot Ra \mid H \cdot K) / P(Ba \cdot Ra \mid \sim H \cdot K)}{P(\sim Bb \cdot \sim Rb \mid H \cdot K) / P(\sim Bb \cdot \sim Rb \mid \sim H \cdot K)} > 1$

iff $(q-(1-p))/(p \cdot r) > 1$ iff

$q > p \cdot r + (1-p)$ iff $q/r > (p + (1-p)/r)$.

We will establish each of the two parts of Theorem 4 by showing that their antecedents imply

$q/r > (p + (1-p)/r)$.

Proof of (4.1)

(4.1): $\frac{P(Ba \cdot Ra \mid H \cdot K) / P(Ba \cdot Ra \mid \sim H \cdot K)}{P(\sim Ba \cdot \sim Ra \mid H \cdot K) / P(\sim Ba \cdot \sim Ra \mid \sim H \cdot K)} > 1$, given that $r > 1(P(\sim Ba \mid H \cdot K) > P(Ra \mid H \cdot K))$

and $O(H \mid Ra \cdot K) / O(H \mid \sim Ba \cdot K) > (p + (1-p)/r)$. [12]

Suppose that $r > 1$ and $O(H \mid Ra \cdot K) / O(H \mid \sim Ba \cdot K) > (p + (1-p)/r)$. Then

$q/r = (P(\sim Ba \mid \sim H \cdot K) / P(Ra \mid \sim H \cdot K)) / (P(\sim Ba \mid H \cdot K) / P(Ra \mid H \cdot K)) =$

[11] Fitelson, and James Hawthorne. “How Bayesian Confirmation Theory Handles the Paradox of the Ravens”, 28-29

[12] $O(X|Y) = \frac{P(X|Y)}{P(\sim X|Y)}$

$(P(Ra \mid H{\cdot}K)/P(Ra \mid {\sim}H{\cdot}K)) \,/\, (P({\sim}Ba \mid H{\cdot}K)/P({\sim}Ba \mid {\sim}H{\cdot}K)) =$

$$\frac{(P(Ra \mid H{\cdot}K)/P(Ra \mid {\sim}H{\cdot}K)) \cdot (P(H \mid K)/P({\sim}H \mid K))}{(P({\sim}Ba \mid H{\cdot}K)/P({\sim}Ba \mid {\sim}H{\cdot}K)) \cdot (P(H \mid K)/P({\sim}H \mid K)} = \frac{O(H \mid Ra{\cdot}K)}{O(H \mid {\sim}Ba{\cdot}K)} > (p + (1-p)/r).$$

<u>Proof of (4.2)</u>

4.2: $\frac{P(Ba{\cdot}Ra \mid H{\cdot}K) \,/\, P(Ba{\cdot}Ra \mid {\sim}H{\cdot}K)}{P({\sim}Ba{\cdot} {\sim}Ra \mid H{\cdot}K) \,/\, P({\sim}Ba{\cdot} {\sim}Ra \mid {\sim}H{\cdot}K)} > 1$, given that $r \leq 1$,(i.e. $P({\sim}Ba \mid H{\cdot}K) \leq P(Ra \mid H{\cdot}K)$

and either $P({\sim}Ba \mid K) > P(Ra \mid K)$

or $q > 1$ (i.e.$P({\sim}Ba \mid {\sim}H{\cdot}K) > P(Ra \mid {\sim}H{\cdot}K)$

Suppose $P({\sim}Ba \mid H{\cdot}K) \leq P(Ra \mid H{\cdot}K)$ (i.e. $r \leq 1$), but either $P({\sim}Ba \mid K) > P(Ra \mid K)$ or

$P({\sim}Ba \mid {\sim}H{\cdot}K) > P(Ra \mid {\sim}H{\cdot}K)$ (i.e. $q > 1$).

First we show that we must have $q > 1$ in any case. This is shown by reductio, as follows:

Suppose $q \leq 1$. Then $P({\sim}Ba \mid K) > P(Ra \mid K)$.

So we have $P({\sim}Ba \mid {\sim}H{\cdot}K) \leq P(Ra \mid {\sim}H{\cdot}K)$ (i.e. $q \leq 1$) and

$P({\sim}Ba \mid H{\cdot}K) \leq P(Ra \mid H{\cdot}K)$ (i.e. $r \leq 1$). Then

$P({\sim}Ba \mid K) = P({\sim}Ba \mid H{\cdot}K) \times P(H \mid K) + P({\sim}Ba \mid {\sim}H{\cdot}K) \times P({\sim}H \mid K) \leq$

$P(Ra \mid H{\cdot}K)\, P(H \mid K) + P(Ra \mid {\sim}H{\cdot}K) \times P({\sim}H \mid K) = P(Ra \mid K) < P({\sim}Ba \mid K)$

Contradiction !!! So both q and r cannot be >1.

Thus we have, if $q > 1$ then $r \leq 1$; so $\frac{1}{r} \geq 1$. Hence $(p + (1-p)/r) \leq p/r + (1-p)/r = 1/r < q/r$.

Appendix 3: A

Proof of Theorems: Chapter III

A.1. Theorem: 3

If H $\models$ E and P (E) $=$ I$-\varepsilon$ then

d(H, E) $\leq \frac{\varepsilon}{\mathrm{I}-\varepsilon}$

Proof:[13]

d(H, E) $=$ P(HE) $-$ P(H). (Definition of 'd')

P(H|E) $= \frac{P(E|H) \times P(H)}{P(E)}$: Bayes' theorem

d(H, E) $= \frac{P(E|H) \times P(H)}{P(E)} -$ P(H)

$=$ P (H) $\times(\frac{P(E|H)}{P(E)} - 1)$

$=$ P (H) $\times (\frac{1}{P(E)} - 1)$: when H entails E

$=$ P (H) $\times \frac{1- P(E)}{P(E)}$:

$=$ P (H) $\times \frac{\varepsilon}{\mathrm{I}-\varepsilon}$: P (E) $=$ I$-\varepsilon$

[13] Fitelson. "Earman on Old Evidence and Measures of Confirmation", 3.

d (H,E) $\leq \frac{\varepsilon}{\mathrm{I}-\varepsilon}$: (Since the maximum value which P(H) can assume is 1)

A.2. Theorem: 4

If H $\models$ E and P(E) $=$ $1-\varepsilon$ then

r(H| E) $= \frac{1}{\mathrm{I}-\varepsilon}$ (for small ε)

Proof:

1. r (H|E) $= \frac{\mathrm{P(H|E)}}{P(H)}$: : Definition of r

2. $= \frac{P(E|H) \times P(H)}{P(E) \times P(H)}$

$= \frac{P(E|H))}{P(E)}$

$= \frac{1}{P(E)}$: When H entails E.

$= \frac{1}{\mathrm{I}-\varepsilon}$: Since P(E) = $1-\varepsilon$

A.3. Theorem: 5.

Even if H⊢E and P(E) ≈ 1, *l*(H,E) can be arbitrarily large.

Following are the assumptions which we hold to derive the desired results

a. H⊢E. Therefore $P(H \,.\sim E) = 0$.

b. $P(E) \approx 1$, that is $P(E) = 1 - \varepsilon$ where 'ε' is very small.

c. So assuming that $\varepsilon \ni (0, \frac{1}{2})$. We can make ε as close to zero as we like.

Following are the constraints on values which he talks of, which generate new probability spaces:

1. $P(H \,.\, E) = \frac{2\varepsilon - 1}{\varepsilon - 1}$
2. $P(\sim H \,.\, E) = \frac{\varepsilon^2}{1-\varepsilon}$
3. $P(\sim H \,.\sim E) = \varepsilon$

It can be derived that l(H,E) is large while the above and the constraints and the assumptions are held

Proof:[14]

$l(H,E) = \frac{P(E|H)}{P(E|\sim H)}$: definition of l

$= \frac{1}{P(E|\sim H)}$: Since H⊢E

$= \frac{1}{\frac{P(E.\sim H)}{P(\sim H)}} = \frac{P(\sim H)}{P(E.\sim H)}$

$= \frac{P(\sim H.E.) + P(\sim H.\sim E.)}{P(E.\sim H)}$

$= 1 + \frac{P(\sim H.\sim E.)}{P(\sim H.E.)} =$

[14] Fitelson. "Earman on Old Evidence and Measures of Confirmation", 3-4.

$=1+\frac{\varepsilon}{\frac{\varepsilon^2}{1-\varepsilon}}$ Since P(∼H . ∼E) = ε, P(∼H . E) = ($\varepsilon^2/(1-\varepsilon)$)

$=1+((1-\varepsilon)(\varepsilon)/\varepsilon^2)$

$=1+((\varepsilon-\varepsilon^2)/\varepsilon^2)$

$=1+((\varepsilon/\varepsilon^2)-1))$

l(H,E) $=(\varepsilon/\varepsilon^2)$

Now it is proven that l(H,E) can be arbitrarily large. When the value of 'ε' diminishes, l(H,E) would be very high. E.g. if ε is .05 then l(H,E) is 20 and if ε is .1 then l(H,E) is 10.

A.4.Theorem 10

"There exists at least one probability function P on L* such that P satisfies (K*) and such that every atomic sentence in L* of the form "A ⊢B" where not both A and ∼B are tautologies gets a value strictly between 0 and 1.

Proof:[15]

Consider L and L* as above. Let P be any strictly positive probability on L. That is, for A in L, P(A) = 0 iff A is truth functionally inconsistent in L. Then extend P to L* as follows:

i. Suppose that A in L is not a tautology. Then let C be any sentence in L which is nontautologous, noncontradictory, and inconsistent with A. If A is not truth-functionally inconsistent in L, then ∼A will do; otherwise let C be any atomic sentence in L. Then, for

[15] Garber, Daniel."Old Evidence and Logical Omniscience in Bayesian Confirmation Theory", 117.

any B in L, let $P(A\vdash B) = P(C)$; and for any D in L*, let $P([A\vdash B] \& D) = P(C \& D)$; $P([A\vdash B]vD) = P(CvD)$; etc.

ii. Suppose that A in L is a tautology and B is not. Then let $P(A\vdash B) = P(B)$; $P([A\vdash B] \& D) = P(B \& D)$; $P([A \vdash B]vD) = P(BvD)$; etc.

iii. Suppose that A and B in L are both tautologies. Then let $P(A \vdash B) = P(a_i)$, where a_i is an arbitrary atomic sentence in L; $P([A \vdash B] \& D) = P(a_i \& D)$; $P([A \vdash B]vD) = P(a_i \vee D)$; etc.

P so extended is clearly a probability on L*. Further, it can easily be shown that P so extended satisfies (K*).

And finally, since P on L is strictly positive, $P(A\vdash B)$ will never have a value of either 0

or 1, except when both A and $\sim$ B are tautologies, in which case it will get a value of 0 by clause (ii)"

Bibliography

Achinstein, Peter. *Concept of Evidence*. New York: Oxford U P, 1983.

Barnes, E.C. "The Quantitative Problem of Old Evidence." *The British Journal of Philosophy of Science* 50.2 (Jun. 1999): 249-264.

Bovens, Luc and Stephan Hartmann. *Bayesian Epistemology*. Oxford: Clarendon Press, 2003.

Carnap, Rudolf. *Logical Foundations of Probability*. Chicago: University of Chicago Press, 1950.

Chihara, Charles S. "Some Problems for Bayesian Confirmation Theory." *The British Journal for the Philosophy of Science* 38.4 (Dec. 1987): 551-560.

Christensen, David. "Measuring Confirmation." *The Journal of Philosophy* 96.9 (Sep. 1999): 437-461.

---. *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford: Clarendon Press, 2004.

Clarke, Roger. "The Raven's Paradox is a Misnomer." *Synthese* 175.3 (Aug. 2010): 427–440.

Douven, Igor. "Testing Inference to the Best Explanation." *Synthese* 130.3 (Mar. 2002): 355-377.

Earman, John and Wesley C, Salmon. "The Confirmation of Scientific Hypotheses." *Introduction to the Philosophy of Science.* Eds. Clark Glymour, *et al.* Indianapolis: Hackett Publishing Company, 1999: 42–103.

Edidin, Aron."From Relative Confirmation to Real Confirmation." *Philosophy of Science* 55.2 (Jun. 1988): 265-271.

Eells, Ellery and Branden Fitelson. "Comments and Criticism: Measuring Confirmation and Evidence." *Journal of Philosophy* 97.12 (Dec.2000): 663-672.

---. "Symmetries and Asymmetries in Evidential Support." *Philosophical Studies* 107.2 (Jan.2002): 129-142.

Fitelson, Branden. "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction." *Philosophy of Science* 69.4 (Dec. 2002): 611–622.

---."Earman on Old Evidence and Measures of Confirmation." Web. 17 August 2010. http://fitelson.org/oe.pdf.

---."Studies in Bayesian Confirmation Theory." Web. 6 March 2008. http://fitelson.org/thesis.pdf

---. "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity." *Philosophy of Science* 66 (Sep. 1999): S362-S378.

---. "A Bayesian Account of Independent Evidence with Applications." *Philosophy of Science* 68.3 (Dec.2000): 1-17

Fitelson, Branden and James Hawthorne. "Discussion: Re-solving Irrelevant Conjunction with Probabilistic Independence." *Philosophy of Science* 71.4 (Oct.2004): 505–514.

---. "How Bayesian Confirmation Theory Handles the Paradox of the Ravens." Web. 11 November 2012. http://fitelson.org/ravens.pdf.

Garber, Daniel."Old Evidence and Logical Omniscience in Bayesian Confirmation Theory." *Testing Scientific Theories*. Ed. John Earman. Minneapolis: University of Minnesota Press, 1983: 99-131.

Gemes, Ken. "Logical Content and Empirical Significance." *The Role of Pragmatics in Contemporary Philosophy.* Eds. Paul Weingartner, G. Schurz, and G. Dorn. Vienna: Hölder-Pichler-Tempsky, 1998: 1-20.

---. "Schurz on Hypothetico-Deductivism." *Erkenntnis* 41.2 (Sep. 1994): 171-181.

---."A New Theory of Content I: Basic Content."*Journal of Philosophical Logic* 23.6 (Dec. 1994): 595-620.

---. "Bootstrapping and Content Parts." *Erkenntnis* 64.3 (May.2006): 345-370.

---. "Carnap-Confirmation, Content-Cutting, and Real Confirmation." Web. 25 July 2011. http://www.bbk.ac.uk/philosophy/our-staff/academics/gemes-work/CarnapConfirmation.doc.

---. "Horwich, Hempel, and Hypothetico-Deductivism." *Philosophy of Science* 57.4 (Dec. 1990): 699-702.

---."Hypothetico-Deductivism, Content, and the Natural Axiomatization of Theories". *Philosophy of Science* 60.3 (Sep. 1993): 477-487.

---."Hypothetico-Deductivism: Incomplete but Not Hopeless". *Erkenntnis* 63.1(Jul. 2005): 139-147.

---. “Hypothetico-Deductivism: The Current State of Play; The Criterion of Empirical Significance: Endgame.” *Erkenntnis*. 49.1 (Jul.1998): 1-20.

Gillies, Donald. *Philosophical Theories of Probability*. London: Routledge, 2000.

---.“In Defense of the Popper-Miller Argument.” *Philosophy of Science.* 53.1 (Jan. 1986) 110-113.

Glymour, Clark. "Hypothetico-Deductivism is Hopeless." *Philosophy of Science* 47.2 (Jun.1980): 322- 325.

---. “Relevant Evidence.” *The Journal of Philosophy* 72.14 (Aug.1975): 403-426.

---.*Theory and Evidence*. Princeton: Princeton U P, 1980.

---. "Explanations, Tests, Unity and Necessity." *Nous* 14.1 (Mar.1980): 31-50

---. *Thinking Things Through: An Introduction to Philosophical Issues and Achievements*, MIT Press, 1997.

Good, I. J. “The White Shoe is a Red Herring.” *The British Journal for the Philosophy of Science 17.4* (Feb. 1967): 322.

---.“The White Shoe *qua* Herring is Pink.” *The British Journal for the Philosophy of Science* 19.2 (Aug. 1968): 156- 157.

Grimes, Thomas R. “Truth, Content and the Hypothetico-Deductive Method.” *Philosophy of Science* 57.3 (Sep. 1990): 514-522.

Hájek, Alan and James M, Joyce. “Confirmation.” *Routledge Companion to the Philosophy of Science*. Eds. Stathis Psillos, and Martin Curd. New York: Routledge, 2008: 115–29.

Harman, Gilbert H. “The Inference to the Best Explanation.” *Philosophical Review* 74.1 (Jan.1965): 88–95.

Hempel, Carl G. *Philosophy of Natural Science*, Englewood Cliffs, NJ: Prentice-Hall, 1966.

---.“Studies in the logic of confirmation.” *The Concept of Evidence.* Ed. Peter Achinstein. New York: Oxford U P, 1983:10-43.

---. “The White Shoe: No Red Herring.” *The British Journal for the Philosophy of Science* 18.3 (Nov. 1967): 239-240.

---. “A Purely Syntactical Definition of Confirmation.” *Journal of Symbolic Logic* 8.4 (Dec.1943): 122–143.

---.*Aspects of Scientific Explanation*. New York: Free Press, 1965.

Howson, Colin. “The ‘Old Evidence’ Problem.” *The British Journal for the Philosophy of Science* 42 (Dec.1991): 547–555.

---.“Bayesianism and Support by Novel Facts.” *The British Journal for the Philosophy of Science*, 35.3 (Sep.1984): 245-251.

Howson, Colin and Peter, Urbach. *Scientific Reasoning: The Bayesian Approach*. Chicago: Open Court, 1993.

Jeffrey, Richard.“Bayesianism with a Human Face”. *Testing Scientific Theories*. Ed. John Earman. Minneapolis: University of Minnesota Press, 1983:133-156.

Keynes, John Maynard. *A Treatise on Probability*. London: Macmillan, 1921.

Kuhn, Thomas. “Objectivity, Value Judgment, and Theory Choice.”.*The Essential Tension*. Chicago: University of Chicago Press, 1977: 320–39.

Lipton, Peter.“Inference to the Best Explanation.” *A Companion to the Philosophy of Science*. Ed. W.H Newton-Smith. Oxford: Blackwell, 2000: 184–193.

Lipton, Peter. *Inference to the Best Explanation*. London: Routledge,2004.

Mackie, John Leslie. “The Paradox of Confirmation.” *The British Journal for the Philosophy of Science* 13.52 (Feb. 1963): 265-277.

Maher, Patrick. “Probability Captures the Logic of Scientific Confirmation”. *Contemporary Debates in the Philosophy of Science*. Ed. Christopher R Hitchcock. Oxford: Blackwell, 2004: 69-93.

---.“Bayesianism and Irrelevant Conjunction”. *Philosophy of Science* 71.4 (Oct. 2004): 515-520.

---. “Subjective and Objective Confirmation.” *Philosophy of Science* 63.2 (Jun.1996): 149-174

---. “Conception of Inductive Logic.” *Philosophy of Science* 73 (Dec.2006): 513–523.

---.“Explication Defended.” *Studia Logica* 86 (Jul. 2007): 331–341.

---. “Bayesian probability.” *Synthese* 172 (Jan. 2010): 119–127.

---. “Physical Probability.” *Logic, Methodology and Philosophy of Science: Proceedings of the Thirteenth International Congress*. Eds. Clark Glymour, Wang Wei, and Dag Westerstahl. London: King's College Publications, 2009: 193 - 210.

---. “Confirmation Theory.” *The Encyclopedia of Philosophy*. Ed. Donald M Borchert.USA: Macmillan, 2005.

Merrill, G. H. "Confirmation and Prediction." *Philosophy of Science* 46.1 (Mar. 1979): 98-117.

Moretti, Luca. "The Tacking by Disjunction Paradox: Bayesianism versus Hypothetico-Deductivism." *Erkenntnis* 64.1 (Jan. 2006): 115-138.

Musgrave, Alan. "Logical Versus Historical Theories of Confirmation." *British Journal for the Philosophy of Science* 25.1 (Mar.1974): 1-23.

Niiniluoto, Ilkka."Novel Facts and Bayesianism." *The British Journal for the Philosophy of Science* 34. 4 (Dec.1983): 375-379.

Norton, John D. "Challenges to Bayesian Confirmation Theory." *Handbook of the Philosophy of Science: Philosophy of Statistics*. Eds. Dov Gabbay, *et al*. Amsterdam: Elsevier, 2011:391–440.

Park, Suck-Jung. "Hypothetico-Deductivism is Still Hopeless." *Erkenntnis* 60.2 (Mar. 2004): 229-234.

Pitt, Joseph C. *Theories of Explanation.* Oxford: Oxford U P, 1988.

Psillos, Stathis. "Inference to the Best Explanation and Bayesianism." *Induction and Deduction in the Sciences* Ed. Friedrich Stadler. Dordrecht: Kluwer, 2004: 83–91.

Quine, W. V. O. "Natural Kinds". *Ontological Relativity and Other Essays*. New York: Columbia U P, 1969: 114–138.

Rosenkrantz, Roger D. "Bayesian Confirmation: Paradise Regained." *The British Journal for the Philosophy of Science* 45.2 (Jun. 1994): 467-476.

---. “Why Glymour is a Bayesian.” *Testing Scientific Theories*. Ed. John Earman. Minneapolis: Univeristy of Minnesota Press, 1983: 69- 97.

---. *Inference, Method and Decision: Towards a Bayesian Philosophy of Science*. Boston: D. Reidel, 1977.

Rozeboom, William W. “Let's Dump Hypothetico-Deductivism for the Right Reasons”. *Philosophy of Science* 49.4 (Dec. 1982): 637-647.

Sainsbury, Richard Mark. *Paradoxes.* New York: Cambridge U P, 2009.

Salmon, Wesely C. “Carnap's inductive logic.” *The Journal of Philosophy* 64.21 (Nov.1967): 725-739.

---. “Partial Entailment as a Basis for Inductive Logic.” *Essays in Honor of Carl G. Hempel*. Ed. Rescher Nicholas. Dordrecht: Reidel,1969: 47–81.

---."Confirmation and Relevance." *The Concept of Evidence.* Ed. Peter Achinstein. New York: Oxford UP, 1983: 95-123.

---."Bayes' Theorem and the History of Science.” *Philosophy of Science: Contemporary Readings.* Eds. Yuri Balashov, and Alex Rosenberg. London: Routledge, 2002. 92-105.

---.“Rationality and Objectivity in Science *or* Tom Kuhn meets Tom Bayes”. *Scientific Theories* Ed. C. Wade Savage. Minneapolis: University of Minnesota Press, 1990: 175-204.

Sarkar, Sahotra and Jessica, Pfeifer. *The Philosophy of Science: An Encyclopedia.* New York: Routledge Taylor &Francis Group, 2006.

Savage, C. Wade. “Bayesian Problems of Old Evidence.” *Scientific Theories*. Ed. Eells Ellery**.**Minnesota Studies in the Philosophy of Science. University of Minnesota Press, 1990: 205-223.

Schilpp, Paul Arthur, ed. *The Philosophy of Rudolf Carnap*. LaSalle IL: Open Court, 1963.

Schurz, Gerhard."Relevant Deduction: From Solving Paradoxes Towards a General Theory." *Erkenntnis* 35 (Jul. 1991): 391-437.

Skyrms, Brain. "Three Ways to Give a Probability Assignment a Memory." *Testing Scientific Theories*. Ed. John Earman. Minneapolis: University of Minnesota Press, 1983: 157-161.

Strevens, Michael. *Notes on Bayesian Confirmation Theory*. Web. 02 November 2007. www.nyu.edu/gsas/dept/philo/user/strevens/Classes/Conf06/BCT.pdf, 2006.

---."Bayesian Confirmation Theory: Inductive Logic, or Mere Inductive Framework?" *Synthese* 141 (Sep 2004):365-379.

Suppes, Patrick."A Bayesian Approach to the Paradoxes of Confirmation." *Aspects of Inductive Logic*. Eds. Jaakko Hintikka, and Patrick Suppes. Amsterdam: North Holland Pub. Co, 1966: 198-207.

Swinburne, Richard G. "The Paradoxes of Confirmation: A Survey." *American Philosophical Quarterly* 8.4 (Oct. 1971): 318-330.

---.*Bayes' Theorem*. Oxford: Oxford U P, 2002: 677-691.

Vranas, Peter B. M. "Hempel's Raven Paradox: A Lacuna in the Standard Bayesian Solution." *The British Journal for the Philosophy of Science* 55.3 (Sep. 2004): 545−560.

Wagner, Carl G. "Old Evidence and New Explanation." *Philosophy of Science* 64.4 (Dec.1997): 677-691.

---."Old Evidence and New Explanation II." *Philosophy of Science* 66. 2(Jun.1999): 283-288.

Waters, C Kenneth. "Relevance Logic Brings Hope to Hypothetico-Deductivism." *Philosophy of Science* 54.3 (Sep. 1987): 453-464.

Woodward, James. "Glymour on Theory Confirmation." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 43.1 (Jan. 1983): 147-152.

Synopsis of the Doctoral Research Work

A Critique of the Bayesian Solutions to Some Paradoxes of Confirmation: Irrelevant Conjunction, Hempel's Paradox and Old Evidence

Justification as a necessary condition of knowledge is a topic that has been much debated in epistemology. In philosophy of science, the question of justification is how to justify theories and laws of science. The justificatory relation between theories (hypothesis) and observation statements (evidence) is called confirmation relation. In this dissertation, my attempt is to analyse or define confirmation relation and its underlying principle. In the introduction of the dissertation, I discuss and evaluate four theories of confirmation: Hempel's Theory of confirmation, the H-D model, Inference to the Best Explanation and the Bayesian Confirmation Theory. And on the basis of the evaluation, I argue that the Bayesian Confirmation theory (BCT) is a more adequate theory of confirmation than other theories of confirmation. Yet the BCT faces crucial challenges from various paradoxes of confirmation: Paradox of Irrelevant conjunction, Hempel's paradox and Problem of old-evidence. In the chapters (1-3) of this dissertation, I mainly analyse the BCT's response to the paradoxes of confirmation.

In the first chapter, my attempt is to analyse the paradox of irrelevant conjunction, which stands as a challenge to various theories of confirmation, and the attempts of these theories to resolve it. The original problem of Irrelevant Conjunction (I.C) was formulated by Clark Glymour as a paradox against the Hypothetico-deductive (H-D) model. According to the H-D model, if H entails E and if E is observed then E confirms H. The paradox of irrelevant conjunction states that if E confirms H then E confirms the conjunction of H and any statement, (X), since H.X entails E. In the first section of my chapter, I analyse how proponents of the H-D model revise their model in their attempt to resolve the paradox. Their attempt is to formulate a new set of conditions by which the confirmation of irrelevant conjunction would be prevented. The section discusses Garry Merrill's, Kenneth Waters' and Ken Gems' solution to the paradox.

In the second section, I discuss the traditional Bayesian approach to the paradox and its inadequacies. In the traditional formulation, paradox arises only in the case of deductive evidence. In the modern formulation of I.C, paradox arises not only in the case of deductive evidence, but in all cases of confirmation. Modern Bayesian approach to the paradox starts with the attempt to define irrelevance. The final defended definition is of Hawthorne and Fitelson and it is as follows: X is (an) irrelevant conjunct to H given K with respect to E just in case $P(E|H.X.K) = P(E|H.K)$. I am claiming that the following formula is an adequate definition of irrelevant conjunct. X is confirmationally irrelevant to H given K with respect to E if and only if $P(X|H.E.K) = P(X|H.K)$. I claim that $P(X|H.E.K) = P(X|H.K)$ should be preferred over Branden Fitelson and Hawthorne's definition, as only the latter can be defended as a sufficient condition of irrelevance. It is because, in the case of conclusive disconfirming evidence and deductive evidence, Hawthorne's and Fitelson's criterion fails to capture the notion of irrelevance.

One of the major steps of Bayesians in resolving the paradox is the softening theorem. Fitelson claims that though irrelevant conjunction is represented as confirmed, the BCT has the richness or tools to soften the impact of that irrelevant conjunction. And it is an elaborative version of John Earman's solution to the paradox. Fitelson's softening theorem is as follows: If E confirms H, and X is confirmationally irrelevant to H, E and H.E (relative to K) then (provided that $P(X|K) < 1$) : $c(H,E|K) > c(H.X, E|K)$ where 'c' (confirmation measure) is either difference measure 'd' or the likelihood ratio measure 'l' of degree of confirmation, but not ratio measure r, because $r(H, E|K) = r(H\&X, E|K)$. Fitelson's claim is that the theorem softens the impact of the paradox. That is, adding an irrelevant conjunct makes a conjunction only less confirmatory than its original hypothesis. However, my claim is that it is too little to claim that it softens the impact of paradox. As per the theorem, due to the irrelevant conjunct, the confirmatory strength of H.X becomes less than the confirmatory strength of H. Such a scenario is counterproductive to the claim of softening attitude because a conjunct which is intuitively ineffective in affecting the confirmation relation is represented as one reducing the confirmatory strength. That is, through the softening theorem, Bayesians are only asserting the point of the paradox. Later I attempt to redefine the 'Softening Theorem'.

Though Bayesians had made an improvement by defining irrelevant conjunct and formulating a theorem which softens the impact of paradox, still they are unable to resolve the paradox as irrelevant conjunction is still considered as confirmed, which is counter intuitive. My claim is that in general, confirmation of hypothesis can be determined only in relation with competing hypothesis in general and the best competitor in particular.

I propose the reformulation of the Bayesian framework of confirmation so as to include the role of the best competitor. My Definition of Bayesian confirmation is as follows:

E confirms H if and only if $\left\{ \frac{P(H|E.K)}{P(H|K)} > \frac{P(H_1|E.K)}{P(H_1|K)} \right\}$ (where H_1 is the best competitor).

Definition of the best competitor is as follows: H_1 is the best competitor to H regarding E iff $|\ P(E|H.K) - P(E|H_1.K)\ | < |\ P(E|H.K) - P(E|H_i.K)\ |$: ('H_i' is any competing hypotheses in the field of inquiry). The best competitor is the notion which is defined in relation with particular evidence. And the best competitor for H may vary in the context of varying evidence. So, the best competitor is a notion which is defined in a context which is posterior to evidence. And as per this definition, the best competitor hypothesis of an irrelevant conjunction H.X is H itself. So H.X would be considered as confirmed only if $\left\{ \frac{P(H.X|E.K)}{P(H.X|K)} > \frac{P(H|E.K)}{P(H|K)} \right\}$. It can be shown that it is not the case and L.H.S and R.H.S are equal. Therefore as per new definition, E is neutral evidence to irrelevant conjunction H.X.

In the second chapter my attempt is to analyse Hempel's paradox which is also known as the Raven paradox. Hempel formulates the paradox as the derivation of a paradoxical / false conclusion from two plausible principles of confirmation. That is, Nicod's criterion (N.C) and Equivalence condition (E.C) entail a paradoxical conclusion that 'A non-black non-raven confirms the hypothesis all ravens are black.' In the first section of this chapter, I have analysed the plausibility and acceptance of the premises from which the paradoxical conclusion is derived. From the second section I have discussed various suggested solutions to the paradox. The attempt to solve the paradox can be largely divided into two categories. (A): Attempts to show that the paradoxical nature of the conclusion is only an impression and (B): attempts to refute the premises which are used for the derivation of the paradoxical conclusion. (B1: Refute

Nicod's Criterion, B2: refute Equivalence Condition, B3: Explore if there is any hidden premise and refute it).

Hempel is the pioneering champion of the approach mentioned in category 'A'. The other major solutions come under the category of B1. Among the earlier attempts to reject Nicod's criterion, the much noted one is due to W.V.O Quine. Quine restrains Nicod's criterion to natural kinds. Thus an object which is non–black and non raven is not an instance at all and thus evades the paradoxical conclusion, but the solution is widely criticised for being stringent. The other major attempts to reject N.C. are those proposed by the Bayesians. Certainly the Bayesians have contributed much towards clarification of Nicod's' criterion. But their outright rejection of Nicod's criterion is fraught with enormous difficulties.

Apart from the attempt to reject Nicod's criterion, another Bayesian attempt was to hold the paradoxical conclusion as valid and explain the apparent paradoxicality by employing comparative and quantitative claims. (Solution comes under the category A).

Comparative claim: (COMPc): $c[(x)(Rx \supset Bx), (Ra \cdot Ba) \mid K] > c[(x)(Rx \supset Bx), (\sim Ba \cdot \sim Ra) \mid K]$.

Quantitative claim: (QUANTc): $c[(x)(Rx \supset Bx), (\sim Ba \cdot \sim Ra) \mid K] > 0$, but very nearly 0.

That is, though contra positive instances confirm the hypothesis, it is generally considered as non –confirmatory because its degree of confirmation is not only comparatively less than direct instances but also very minute. By employing both these claims, Bayesians adequately explain the impression of paradoxicality. The bad news for the Bayesian is that they failed to establish their quantitative claim and comparative claim because those claims are built upon controversial assumptions. Vranas claims that the assumption $(P(\sim Ba|(x)(Rx \supset Bx.K) = P \sim Ba|K))$ is necessary for the quantitative claim, but there is no good reason to assume that it is true.

In response to Vranas' challenge, Fitelson has successfully established the comparative claim based on much weaker assumptions which are absolutely non- controversial, but only comparative claim can be derived from Fitelson's weaker assumptions. And my position is that comparative claim alone cannot dispel the paradoxical impression attached to the confirmation by contra positive instances. And my larger criticism is that Bayesian solutions violate Equivalence condition to which they are committed.

Among the solutions, my commitment is to the solutions which attempt to modify the N.C. I am inclined to the style of Hempel and Quine who have tried to constrain the N.C instead of rejecting it. I hold the position that conclusion_1 (∼Ra. ∼Ba confirms (x) (Rx⊃Bx)) is not paradoxical. And what is paradoxical is conclusion_2 (Wa. Sa (White Shoe) confirms (x) (Rx⊃Bx)) which is derived from conclusion_1 using the premise (∼Ra. ∼Ba (a non- raven which is non- black) is Wa. Sa (White Shoe)). And the premise which is employed to derive the paradoxical conclusion (conclusion 2) is untenable because 'white shoe' or 'red pencil' or 'white board' are not instances of hypothesis that 'All non- black things are non- raven'. In other words they are not 'non- black non- raven'. In order to substantiate my claim, I attempt to make precise the definition of 'Instance'. In literature, the sole condition of being an instance of H is, the object 'I' should exemplify the predicate(s) of H. That is an instance of H ((x)(Fx ⊃ Gx)), is an object 'I' who has a pair of predicates 'F' and 'G'. I add the following conditions which are simple and quite compatible to our intuition regarding the instances.

Condition 2: 'I' is an instance of H only if: $P(I|H.K) = 1$. It only says that instances are deductive evidence (evidence which are the consequences of the hypothesis) which is obvious but it can prevent Roger Rosenkrantz's counter example.

Condition 3: if 'I' is an instance of H then it would not be an instance to any other competing hypothesis (hypothesis which are contrary to H). In general we can say as follows: 'I' is an instance of H only if $P(I|H.K) > P(I|H_c.K)$: H_c is any hypothesis which is contrary to H.

On the basis of the third condition, I say that white shoe is not an instance of H: (x) (∼Bx ⊃ ∼Rx). Consider some contrary hypotheses of 'All non-black things are non–raven'. H_{C1}: All non- yellow things are non- raven, H_{C2}: All non- blue things are non- raven, H_{C3}: all non- white things are non- raven. Here except for the hypothesis H_{C3}, for all hypotheses, the likelihood of white shoe is 1. i.e. $P(Wa.Sa|H.K) = P(Wa.Sa|\ H_{C1}.K) = P(Wa.Sa|\ H_{C2}.K) = 1$. That is, white shoe is a confirming instance (as per traditional definition) for H and for all contrary hypotheses of H except for one. It is a clear violation of Condition 3. Therefore as per the new definition of instance, 'Wa.Sa' is not a confirming instance of H: (x) (∼Bx ⊃ ∼Rx). Thus the paradoxical conclusion 2 cannot be derived. I think it would not be difficult accept that 'Wa.Sa' is not an instance of H, but the question remains as to what kind of an object is 'non-raven and non- black'. Now the challenge is that we have vast kind of objects (e.g. white shoe, red pencil...) which satisfy the 'Conditions of Instance' 1 and 2. (All objects exemplify the predicates non- raven and non- black and likelihood is 1). Though many of these violate 'Condition 3', the 'Condition 3' cannot function as heuristic principle to identify the real instances, as it is only test of instances. To pin point the real instance, I formulate a reasonable assumption: 'From a set of objects which satisfy condition 1, the object(s) which provide(s) greater support to H than another object is the real instance(s). In this case, for objects 'a' (red pencil), 'b' (white shoe) ,c (blue water), P(∼Ra. ∼Ba|H.K) = 1, P(∼Rb. ∼Bb|H.K) = 1, P(∼Rc. ∼Bc|H.K) = 1, then the objects which has the lowest probability (P(E)) for its occurrence would provide maximum supports to H and as per our assumption that would be the instance of H. So to determine what is the real instance we need to determine P(∼Ra. ∼Ba), P(∼Rb. ∼Bb), P(∼Rc. ∼Bc). Obviously the probability of White shoe or red pencil being non- black non –raven is high. But consider an object whose colour is near to black and it is so similar to raven, yet not black and not raven; I think it would not be non- controversial to assume that object as an instance of contrapositive hypothesis 'all non- black things are non- raven'. Thus there is nothing paradoxical in holding that contrapositive instance confirms H and its logically equivalent hypothesis if instances and contra positive instances are understood as per the new definition of Instance.

In the third chapter, I analyse the Problem of old evidence, which is formulated by Glymour. Some evidence are discovered prior to the formulation of hypotheses. Such evidence are called old evidence. Citing the historical examples, Glymour points out that often old evidence works as confirmatory evidence though it is not the case that all old evidence are relevant. The problem is that the BCT fails to capture the confirmation by old evidence.

Glymour shows the problem as follows:

P(H|E.K) = P(H|K) when P(E|K) = 1. It is because P(E|K) = 1 entails that P(E|H.K) = 1.

$$P(H|E) = \frac{P(E|H.K) \times P(H|K)}{P(E|K)} = P(H|K) .$$

All standard measures of Bayesian confirmation (d, r, l, and s) are found as vulnerable to the problem, as all measures are based on the relation between prior probability and posterior probability of the hypothesis.

In the first section I have introduced the classification of the problem of old evidence. Bayesian literatures provide large cluster classification of the problem and many of the solutions are dependent upon the classification. The primary classification of the problem of old evidence is the ahistorical problem of evidence and the historical problem of evidence. When scientists discover a confirmation relation between newly invented theory and the evidence which is already known, the historical problem of old evidence arises. An ahistorical problem is that even after the event of confirmation, (discovering the confirmation relation by scientists) evidence remains as evidence of H and the BCT cannot explain the time enduring nature of evidence. Later, Ellery Eells introduces further classifications of the problem.

In the second section, I have introduced the different suggested solutions: first among them is that of Counter factual strategy. Collin Howson and Peter Urbach formulated their solution called counter- factual strategy (C.F.S). The most meticulous and scathing attack on C.F.S is due to Charles Chihara. His point is that if this sort of defence of Bayesianism is to work, one needs a reasonably clear procedure or rule to determine required stock of background information. In the next section I have discussed Eells' solution. Another important suggestion is due to Earman. Earman's one suggestion is of refuting the claim that $P(E) = 1$ in the case of old- evidence. Earman's argument is that except in the case of tautology, the assignment of value 1 is irrational. According to Earman, $P(E)$ can be .999 but not 1. When $P(E) = .999$, it can be shown that $P(H|E) > P(H)$ in the case of old- evidence. Thus the qualitative problem of old evidence is resolved but the quantitative problem of old-evidence remains. There are various criticisms to such suggestions. But the important one is due to Patrick Maher. Maher's point is that the problem is about the inconsistency between E being certain and E being regarded as evidence for H, and not about the certain nature of old evidence.

Among the various solutions, I consider Garber's attempt highly worth considering. Garber understands that the problem stems from the serious inadequacy of the Bayesian framework. Garber's point is that in the case of old evidence what confirms H is not E but the logical relation between H and E. So to explicate the confirmation by old evidence, we need to conditionalise H upon the logical relation between H and E. And he shows that how the conditionalisation upon the logical relation between H and E can be characterised in a formal frame work. Garber's attempt is a clear departure from the standard Bayesian position on the point that it changes the variable which is to be conditionalised upon. This departure and its formalization make Garber's attempt a remarkable one among the various solutions to the problem of old evidence. Though various criticisms are levelled against Garber's account, at the end, I argue that Garber's account and its ingenious intuition open up various paths to address the problem.

In the conclusion, I evaluate the significance of the new Bayesian model of confirmation, which I have discussed in the first and the second chapter and analyse how the new Bayesian model seriously challenges the subjective interpretation of probability and argue that more robust interpretation of probability than the subjective interpretation would help to address various challenges to the BCT in an improved manner.

Date: 18/06/2013

Place: Hyderabad

Submitted By

Philose Koshy, 06HPPH01

Ph.D Philosophy

University of Hyderabad