SEP #3 (A Noble + Harman) 65327 – 7/21/87

Conditional Probability and Nicod's Conditional

David Lewis argued in 1976 that "there is no way to interpret a conditional connective so that, with sufficient generality, the probabilities of conditionals will equal the appropriate conditional probabilities."[1] His argument assumes that any good logic of conditionals must include certain theorems which, when conjoined with the Kolmogorov axioms of probability, lead to absurdities.

In this paper I present and defend a type of conditional which does not have those theorems and is assigned measures of probability equal to those of conditional probabilities. I have previously been interested in the kind of conditional involved, which I call 'C-conditionals' (as opposed to truth-functional or 'TF-conditionals'); but its relation to conditional probability first came to mind in reading Hempel remarks on Jean Nicod's treatment of conditionals in confirmation theory[2].

Hempel described Nicod's criterion as follows: According to Nicod, a universalized conditional, '(x)(if P(x) then Q(x)', is confirmed by an object

> "...if and only if it satisfies both the antecedent (here: 'P(x)') and the consequent (here 'Q(x)') of the conditional; it disconfirms the hypothesis if and only if the satisfies the antecedent, but not the consequent of the conditional; and (we add this to Nicod's statement) it is neutral or irrelevant, with respect to the hypothesis if it does not satisfy the antecedent..." [3]

Thus Nicod's conditional is true or false only in cases where the antecedent is true, and treats the conditional as being neither true nor false (neutral) in cases where the antecedent is not satisfied. Hempel adds that Nicod "states explicitly what is perhaps the most common tacit interpretation of the concept of confirmation." In this paper ~~I try to~~ I shall ~~explicate~~ flesh out the notion of such a conditional syntactically and semantically and ~~to~~ show that such a conditional ~~can~~ provides a solution to David Lewis's problem.

1. Introduction

~~We~~ I must first distinguish 1) the use of ratios from 0 to 1 to

---

[1]"Probabilities of Conditionals and Conditional Probabilities", **The Philosophical Review**, LXXXV, 3 (July 1976).
[2]See Carl G. Hempel, **Aspects of Scientific Explanation**, Free Press, 1965, "Studies in the Logic of Confirmation" pp 3-53, and Jean Nicod, **The Logical Problem of Induction**, 1923, English translations in Jean Nicod, **Foundations of Geometry and Induction** (translated by P.P.Wiener) London, 1930; (cf. p.219) and also in Jean Nicod, **Geometry and Induction**, (translator, Michael Woods) University of California Press, 1970, (Cf. p. 189)
[3]Hempel, Opus Cit., p.11.

represent relative statistical distributions of sub-classes in a reference class of evidential data from 2) projections of these ratios upon other events, or classes of events beyond the evidential data. Most philosophical discussions of probability are concerned with how to justify projections of ratios to events or classes of events not included in the given evidence.

This paper does not address issues about projectibility. It treats Kolmogorov's axioms solely as rules for assigning ratios from 0 to 1 to statements in canonical logical language on the basis of statistical distributions among subclasses of a finite reference class of data. It does not matter whether the initial assignments are based on empirically established frequencies, or on **a priori** numerical assignments. Thus the **a priori** vs frequency theory debate is also not at issue.

Thus occurrences below of the expression, 'Pr(A)', for 'the probability of A', should be interpreted merely as denoting the frequency of true instances of an expression with respect to a class of initial evidentiary data, without implying projectibility or addressing issues about initial assignments.

The objective is to retain the numerical consequences of the Kolmogorov axioms, while altering the presupposed logic from one which uses a truth-functional conditional only to a logic which includes C-conditionals whose probability ratios are those of standard conditional probability.

Kolmogorov's axioms for probabilitity theory may be stated as follows:

PR1 If '(A **iff** B)' is logically true, then Pr(A)=Pr(B).
PR2 $0 \leq Pr(A) \leq 1$
PR3 If |-A, then Pr(A)=1
PR4 Pr(AvB) = (Pr(A) + Pr(B) - Pr(A&B)
PR5 If P(A)>0 then **Pr(B/A)** = $\frac{Pr(A\&B)}{Pr(A)}$

The changes I propose are that 1) the occurrence of '**iff**' in the object-language in PR1 should be changed from a TF-biconditional to a C-biconditional, and that 2) conditional probability, as expressed by '**Pr(A/B)**' in PR5, be changed to 'Pr(If A then B)' with a C-conditional. My claim is that these changes are intuitively and semantically justifiable, and that the numerical results will be the same as in standard Kolmogorov theory.

2. <u>The truth-functional conditional in probability theory.</u>

Let us first review what the problem is.

All probability assignments greater than 0 presuppose the inclusion of one sub-set of the Reference Class within another sub-set of that class. The probability ratio is gotten by dividing the number of members in the included sub-set by the number of members in the sub-set which includes it. The included sub-set is always the intersection of some sub-set with the including sub-set. In some cases the divisor is the number of members in the <u>improper</u> sub-set (the Reference Class itself).

Such is the case when the Probability operator is applied to a purely truth-functional expression. When the divisor is gotten from some proper sub-set of the Reference class, the probability ratio is established by the formula for conditional probability. The assignment of 0 signifies that the two sub-sets are disjoint.

The function called conditional probability, expressed by the symbol, 'Pr(B/A)', is usually read "the probability of B, given A". Intuitively it seems reasonable to expect that conditional probability is simply the probability of a conditional, i.e., that 'Pr(B/A)' means 'Pr(if A then B)'. But if the 'if...then' in 'Pr(If A then B)' is a truth-functional conditional the results are very different from what is wanted. Thus a separate concept of "conditional probability" has been introduced, and the probability of the truth-functional conditional has been rejected for that purpose.

Why is the truth-functional conditional unsatisfactory?

Suppose we have a bowl of 10 pieces of fruit (= the reference class, R) of which 6 are apples and 4 are pears; and suppose there is just one brown apple (a russet apple), though there are 2 brown pears. Let 'A' stand for 'x is an apple' and 'B' stand for 'x is brown' etc. We may represent the contents of the bowl as follows, with sub-sets R={1,2,3,4,5,6,7,8,9,10}, B={5,6,7} and A={1,2,3,4,6,9}:

Figure 1

R:	1	2	3	4	5	6	7	8	9	10
A:	1	2	3	4		6			9	
B:					5	6	7			

By standard rules for establishing probability ratios, we get

Pr(A)=.6; Pr(-A)=(1-Pr(A))=.4; Pr(B)=.3; Pr(A&B)=.1; Pr(-A&B)=.2; and, by PR4, the probability of the TF-conditional '(A->B)' is:

Pr(A->B) =Pr(-AvB)=(Pr(-A) + Pr(B) - Pr(-A&B)) =(.4 +.3 -.2) = .5

All of these represent the probability of a truth-functional expression with reference to the improper sub-set, R, the Reference class.

Thus the probability of 'if x is an apple, then x is brown' with a truth-functional conditional, in this case, is 3/6, or .5. But, the probability that if x is an apple, then x is brown, should be one out of six, i.e., 1/6 or .167. Using the rule, PR5, for conditional probability, i.e.,

$$\textbf{If } Pr(A)>0\textbf{, then } Pr(B/A) = \frac{Pr(A\&B)}{Pr(A)}$$

we get $Pr(B/A) = \frac{Pr(A\&B)}{Pr(A)} = \frac{.1}{.6} = .167$

which is just what we want.

If we let A be 'x is an apricot', the probability ratio for 'If x is an apricot, then x is brown' interpreted as a truth-functional conditional is, paradoxically, 1.0, or certainty, due to the falsity of the antecedent. For there are no apricots in the bowl, so that Pr(A)=0, Pr(A&B)=0 and Pr(-A&B) = Pr(B) =.33, so that by PR4,

$$\begin{aligned} Pr(A\text{->}B) = Pr(-AvB) &= Pr(-A) + Pr(B) - Pr(-A\&B) = \\ &= 1.0 + .33 - .33 = 1.0 \end{aligned}$$

In contrast, conditional probability, gives no probability at all for apricots in the barrel being brown; this case is ruled out by the "If Pr(A)>0,..." proviso. Like Nicod's conditional, if the antecedent has no true cases, the conditional has no probability ratio, not even 0.

The discrepancies between Pr(B/A) and Pr(-AvB) (i.e., between conditional probability and the probability of the truth-functional conditional) are illustrated in Tables I and II below. The upper figures in each row are Pr(-AvB), the lower figures in **bold** type in each row, are the correct probability, Pr(B/A):

TABLE I

Pr(B,if A)	Pr(A)	Pr(B) .0	.2	.5	.8	1.0
	.0	1.0	.8	.5	.2	0
		Undef.	**Undef.**	**Undef.**	**Undef.**	**Undef**
	.2	.8	.84	.9	.96	1.0
		.0	**.2**	**.5**	**.8**	**1.0**
Pr(A)	.5	.5	.6	.75	.9	1.0
		.0	**.2**	**.5**	**.8**	**1.0**
	.8	.2	.36	.6	.84	1.0
		.0	**.2**	**.5**	**.6**	**1.0**
	1.0	.0	.2	.5	.8	1.0
		.0	**.2**	**.5**	**.8**	**1.0**

TABLE II
Discrepancies

Pr(A) \ Pr(B)	.0	.2	.5	.8	1.0
.0	1.00	.80	.50	.20	.0
.2	.80	.64	.40	.16	.0
.5	.50	.40	.25	.10	.0
.8	.20	.16	.10	.04	.0
1.0	.0	.0	.0	.0	.0

In summary, 1) Only when either antecedent or consequent has a probability of 1, do Pr(B/A) and Pr(-AvB) coincide. 2) As the probability of the antecedent approaches 0 (provided that the probability of the consequent is less than 1), discrepancies between the probability of the truth-functional conditional and conditional probability increase. 3) As the probabilities of both antecedent and consequent approach 0, the discrepancy approaches 1, i.e., an outright contradiction between the two results. 4) If the antecedent has no probability, or 0, then the probability of the TF-conditional is 1.0, but there is no probability ratio at all for the conditional probability.

These discrepancies are consequences of the facts that a) each instantiation of the TF-conditional of standard logic, (Øx ->Ψx), must be assigned either T or F, and b) the TF-conditional is true whenever its antecedent is not true, or its consequent is true.

3. Constraints on the logic of conditionals, based on requirements of the logic of probabilities.

Assume we want the C-conditional (i.e., Nicod's conditional) to have probability assignments that coincide with those of the standard conditional probability. Let the symbol, '(A=>B)', represent such a conditional (as distinct from '(A->B)' for the truth-functional conditional). What constraints must be placed on such a conditional?

If Pr(A=>B) is to capture conditional probability, the following, among others, can not be theorems of logic [4], and hence interchangable in probability expressions, under PR1:

I. ((A=>B) <=> (-B=>-A)) \
II. ((A=>-B) <=> (B=>-A)) | "Transposition"
III. ((-A=>B) <=> (-B=>A)) /
IV. ((A=>(B=>C)) <=> ((A&B)=>C)) "Exportation"

Also the Law of Bivalence (that no statement can be neither True nor False) must be given up and constraints must be laid on the rules permitting nested conditionals.

Consider what would happen if these constraints are violated.

3.1 Failure of Transposition

Returning to Figure I, PR1 would be falsified in this model if conditional probability were to be the probability of some conditional (A=>B) for which logical equivalence yielded Transposition Principles:

Figure 1

R:	1	2	3	4	5	6	7	8	9	10
A:	1	2	3	4		6			9	
B:					5	6	7			

Pr(A)=.6, Pr(B)=.3, Pr(-A)=(1-Pr(A))=.4 Pr(-B)=(1-Pr(B))=.7 Pr(A&B)=.1, and Pr(-A.-B)=.2, Hence, by PR5, with 'Pr(A=>B)' replacing 'Pr(B/A)', we get,

$$Pr(A{=>}B) = \frac{Pr(A\&B)}{Pr(A)} = \frac{.1}{.6} = .167, \quad Pr(-B{=>}-A) = \frac{Pr(-B\&-A)}{Pr(-B)} = \frac{.2}{.7} = .286$$

But, if '((A=>B) <=> (-B=>-A)' is logically true, then, by PR1, Pr(A=>B)=Pr(-B=>-A), i.e., .167=.286, which is absurd.

2.2 Failure of Exportation

We again get inequalities asserted as equalities, by PR1, if Exportation is a theorem of Logic:

(1) '((A=>(B=>C)) <=> ((A&B)=>C))' is logically true. [Export]

(2) If 1>Pr(A)>0,
then Pr(A) = Pr(B=>A)xPrB + Pr(-B=>A)xPr(-B) [by LCP]
[By replacing 'Pr(B/A)' by 'Pr(A=>B)' in the standard

[4]These are theorems David Lewis presupposes must belong to any decent logic of conditionals, Opus cit.

probability theorem "The Law of Compound Probabilities"5]

(3) If 0≤ Pr(A=>B) ≤ 1,
then Pr(A=>B)= Pr(B=>(A=>B))xPr(B) + Pr(-B=>(A=>B))xPr(-B)
[By (2), substituting 'A=>B' for 'A']

(4) If 0≤ Pr(A=>B) ≤ 1,
then Pr(A=>B)= Pr((B&A)=>B))xPr(B) + Pr((-B&A)=>B))xPr(-B)
[By (3), and PR1]

(5) $Pr((B\&A)=>B) = \frac{Pr(B\&A\&B)}{Pr(B\&A)} = \frac{Pr(A\&B)}{Pr(A\&B)} = \frac{.1}{.1} = 1.0$

(6) $Pr((-B\&A)=B) = \frac{Pr(-B\&A\&B)}{Pr(B\&A)} = \frac{.0}{.1} = 0$

(7) Pr(A=>B) = 1xPr(B) + 0xPr(-B) = Pr(B) [By 4),5),6),=s]

(8) Pr(A=>B) = Pr(B) [7),Arith]

Substituting in (8) the values assigned to Pr(A=>B) and Pr(B) in the model from Figure 1 we get:

(9) .167 = .333

Thus, assuming that the other steps (especially (3)) are in order, Exportation would lead to absurd results in the Kolmogorov system.

3.3 The Failure of Bivalence

The proviso in PR5 that the probability of the condition must be greater than 0, is required by the arithmetical principle that nothing is divisible by 0, as would be the case if Pr(A)=0 and Pr(A=>B) = (Pr(A&B)÷Pr(A)) as in PR5.

The imposition of this proviso suggests that a conditional, 'If A then B', could have no probability - neither 1.0 nor something less than 1.0 - and that a probability assignment to it would be neither true nor false, if its antecedent had no probability at all. This proviso avoids the "paradox" of the TF-conditional, that every conditional is made true (has Probability 1.0) if the antecedent is never true. It also coincides with a semantic feature of C-conditionals, namely that a C-conditional is neither true nor false when its antecedent is not true, i.e., the rejection of the Law of Bivalence.

Let us return to our original example with Figure 1, in which the conditional probability of 'x is brown, given that x is an apple' differed from 'Pr(x)(xeA -> xeB) with a truth-functional conditional.

First consider the cases in which conditional probability is not involved. We have a reference class R, and sub-classes A and

---

5I.e., If 0≤ Pr(A) ≤ 1,
then Pr(A) = Pr(A/B)xPr(B) + Pr(A/-B)xPr(-B)]

B distributed in various ways in the reference class.

Figure 1

R:	1	2	3	4	5	6	7	8	9	10
A:	1	2	3	4		6			9	
B:					5	6	7			

We are now construing the assignment of probability ratios as based on the number of members in a sub-class of the reference class, to the reference class itself.

<u>Pr(A)</u>= Pr(xeA) = .6; <u>Pr(-A)</u>= Pr(-xeA) = (1-Pr(xeA)=.4;
<u>Pr(B)</u>= Pr(xeB) = .3; <u>Pr(A&B)</u>= Pr(xeA&xeB)= .1;
<u>Pr(-A&B)</u>= Pr(-xeA&xeB) =.2
<u>Pr(A->B)</u>= Pr(xeA->xeB) = <u>Pr(-AvB)</u> = Pr(-xeA v xeB) =.5

On this construal, the variable x ranges over the set of all members of the Reference class, and the probability ratio, Pr(Øx) is the number of members of x̂Øx divided by the number of members in the reference class, R.

To determine Pr(A), we instantiate the sentential function, 'xeA', for all members of R, count up the total instantiations which are true and divide by the number of cases in which 'xeR' is true. Since there are six instantiations which are true, namely '1eA','2eA', '3eA', '4eA', '6eA' and '9eA' and four which are false, we get the ratio 6/10, or .6 as the probability ratio for (A). This gives us the number of instances of 'xeA' which come out true relative to all instances true and false; this is what "the probability of A (relative to reference class R)" means.

The atomic sentential functions, 'xeR', 'xeA', 'xeB', and quantifier-free truth-functional molecular propositions formed from them are in effect class predicates which define sub-classes (proper or improper) of the set R. In the calculus of probabilities, the negation sign defines class complements relative to R, the ampersand defines class intersections and the velle defines unions of classes.

The molecular sentential function 'xeA -> xeB', with the truth-functional conditional, also yields a sub-class predicate, one equivalent to '(-xeA v xeB)' or '-(xeA.-xeB)'. As in the other cases, to determine the number of members in its sub-class, one instantiates 'x' for each member of the reference class R, and counts up the number of instantiations which come out true. That number over the number of members in the class R gives the probability ratio, Pr(-AvB). This calculation is shown in the left-hand part of TABLE III.

TABLE III

xeA	xeB	TF-Conditional (xeA -> xeB)		C-Conditional (xeA => xeB)	
T	F	(1eA -> 1eB)	= F	(1eA => 1eB) =	F
T	F	(2eA -> 2eB)	= F	(2eA => 2eB) =	F
T	F	(3eA -> 3eB)	= F	(3eA => 3eB) =	F
T	F	(4eA -> 4eB)	= F	(4eA => 4eB) =	F

F	T	(5eA -> 5eB) = T	(5eA => 5eB) = -T&-F
T	T	(6eA -> 6eB) = T	(6eA => 6eB) = T
F	T	(7eA -> 7eB) = T	(7eA => 7eB) = -T&-F
F	F	(8eA -> 8eB) = T	(8eA => 8eB) = -T&-F
T	F	(9eA -> 9eB) = F	(9eA => 9eB) = F
F	F	(10eA -> 10eB) = T	(10eA => 10eB)= -T&-F
		Pr(x)(xeA -> xeB) = .5	Pr(x)(xeA => xeB) = .167

Secondly, consider the standard treatment of conditional probability. And suppose that (A=>B) is some conditional such that Pr(B/A) = Pr(A=>B) always holds. In this case we assume again that Pr(A=>B) is short for Pr(xeA=>xeB). The probability ratio, as before is:

$$\underline{Pr(B/A)} = Pr(xeA{=}{>}xeB) = \frac{Pr(xeA\&xeB)}{Pr(xeA)} = \frac{.1}{.6} = .167$$

An ordinary (non-truth-functional) reading of 'if...then', "What is the probability ratio for B's being present, if A is present?", asks what proportion of the members of A are B. To answer this we look at the class A alone, and determine what proportion of its members are members of the intersection of A and B. The answer is the probability ratio, given by its conditional probability. The C-conditional satisfies this intuition. It differs from the truth-functional conditional in that its antecedent, 'xeA' in '(xeA=>xeB)' fixes the effective range of the variable. When the antecedent of a C-conditional has free x, the range of the variable x is, in effect limited to members of R which belong to the sub-class A. The number of members in A is the denominator of the probability ratio, and the number of members in A which are also B, is the numerator. The right-hand side of TABLE III illustrates how this works.

But how could a conditional in the matrix change the range of the variable in the universal quantifier? Presumably it can not. But if a) the Probability Ratio is defined as the number of cases in which the instances of the universal quantification are True, over the total number of instances which are either True or are false and b) the conditional in the matrix is neither True nor False if the antecedent is not true (not satisfied), then with <u>this</u> type of conditional the range of the variable appears to be restricted to the members of class denoted in the antecedent. Only in this way will the probability of the conditional, Pr(A=>B), equal the standard conditional probability, Pr(B/A).

Restating this explanation somewhat differently:

1) With all purely truth-functional wffs, Øx, including those with truth-functional conditionals, the probability of '(x)(Øx)' is determined by instantiating 'x' for each member of the reference class R, and counting up the number of instantiations which come out true. Since the reference class, by hypothesis here, is a class of actual existing entities collected in the sample, every truth-functional statement about a member of that class will be true, or else false of them. Thus the number of true cases, over the number of true or false cases, is simply the number of true cases over the number of members of R; this gives

the probability ratio, Pr(x)Øx, in all cases, including the cases where Øx is '(-xeA v xeB)' i.e., '(xeA -> xeB)'.

2) In the ordinary (non-truth-functional) sense of 'if...then' one proceeds quite differently. To determine the probability of '(x)(xeA=>xeB)' we instantiate the 'x' with each member of R, but we do not count 'aeA => aeB' true in cases where 'aeA' is false, nor in all cases where 'aeB' is true, as the truth-functional conditional does. Rather we look only at members of R of which '_eA' is true, and having counted those for the denominator of the probability measure, we see how many of that same group are also members for which '_eB' is true, to get the numerator for the ratio. For instantiations in which 'xeA' is false, the conditional 'xeA=>xeB' as a whole is treated as neither true nor false. Thus the right-hand portion of TABLE III shows how this procedure gives us the same probability ratio as demanded by the formula for conditional probability, Pr(B/A).

Thus the semantics of any conditional which can be used to express conditional probability, can not be a semantics which is restricted by Bivalence, i.e., that no statement can be neither True nor False. This conclusion is needed to account for the apparent restrictions on the ranges of variables in conditional probabilities. It also explains the proviso, 'If Pr(A)>0,'in the rule, for conditional probability, PR5. For 'Pr(A)=0'is equivalent on our account to 'Pr(x)xeA =0', and this is the case if and only if no member of the reference class is a member of A, i.e., 'xeA' is not true for any instance at all. Thus there can be no ratio of the number of true instances of '(xeA & Øx)' over true instances of 'xeA' (since 0 is never a denominator), and thus no ratio at all. Thus in cases where Pr(A)=0, Pr(A=>B) is not true in some proportion of its instances, and is not false in some proportion of its instances, it is neither true nor false; i.e., it is both not true and not false, contrary to Bivalence.

All probability ratios are conditional probability ratios.

According to the account we are giving, all probabilities ratios are implicitly conditional probabilities. Where A is a schema of standard logic, the ratio Pr(A) is the ratio of true instances of 'xeA&xeR' over the number of true instances of 'xeR', i.e., the number of members in the reference class. In other words, the probability ratio for a truth-functionally defined sub-class x̂Øx is determined by the number of members of the intersection of x̂Øx with x̂Rx over the number of members of x̂Rx.

In the model of Figure 1, the probabilities of A, B, -B, (A&B), (A&-B), -(AvB), (-Bv(-A&B)), etc. are defined as their probability, given the reference class, R. To find the probability ratio of each, we count the number of members in the intersection of that class with the reference class, and take the ratio of that number over the number of members of R, e.g.,

$$\text{Pr(A) Pr(R=>A)} = \frac{\text{Pr(R\&A)}}{\text{Pr(R)}} = \frac{6}{10} = .6$$

and of course the probability of R itself is P(R/R) = Pr(R=>R) =

$$\text{Pr(x)(xeR=>xeR)} = \frac{\text{Pr(xeR\&xeR)}}{\text{Pr(xeR)}} = \frac{\text{Pr(R\&R)}}{\text{Pr(R)}} = \frac{10}{10} = 1.0.$$

In cases of conditional probability, e.g.,the probability of '(x)(xeA=>xeB)', it is the ratio of the number of members in the intersection of two sub-classes A and B of the reference class R, to the number of R that are members in A. I.e., for every 'Pr(A)' there is some B such that, 'PrA' abbreviates 'Pr(x)(xeB=>xeA)', where 'B' may be 'R'.

$$\text{I.e., } \frac{Pr(x)(xeB\&xeA)}{Pr(x)xeB} = \frac{N\hat{x}(xeB\wedge xeA)}{N\hat{x}\ xeB}$$

Always, then, 'Pr(A)' abbreviates, for some B (possibly R),

$$\frac{\text{the number of members in the intersection of A and B}}{\text{the number of members in B}}$$

I.e., Every ratio 'Pr(A)' is an abbreviation of 'Pr(A/B)', or Pr(B=>A), for some B.

3.4 The Problem of Nested Conditionals

The problem of nested conditionals, i.e., the question, how does one determine the probability ratio for nested conditionals like '(A=>(B=>C))' or '((A=>B)=>C)'?, is often viewed as a difficulty. But there is no problem here. The conditionals needed and used in establishing probability ratios are always 1st-degree conditionals, i.e., conditionals with only the standard truth-functional connectives in their antecedent and consequent. X

This is shown by consideration of the methods outlined above for establishing probability ratios by inspecting distributions among sub-classes of a Reference class or Sample Population.

In the cases where A has only truth-functional connectives, 'Pr(A)' abbreviates 'Pr(A/R)' as mentioned above. With the ~~new~~ C-conditional, Pr(A&B) = Pr(R=>(A&B) and Pr(A)=Pr(R=>A). But the denominator for these initial probability ratios (for A, (A&B), (AvB), etc.) is always Pr(R), and Pr(R) always cancels out. Thus, since Pr(A&B)= Pr(R=>(A&B)), and Pr(R)=1.0,

(x)(Pr(A)=x ⇒ (∃y)(y∈R, Pr(A)=Pr(y⇒A))   (y)(Pr(x∈A)=y ⇒ (∃w)(w∈R, Pr(x∈A)=Pr(x∈w⇒x∈A))

$$\frac{Pr(R{=>}(A\&B))}{Pr(R)} = \frac{\frac{Pr(R\&A\&B)}{Pr(R)}}{\frac{Pr(R)}{Pr(R)}} = \frac{\frac{N\hat{x}(xeR\&xeA\&xeB)}{N\hat{x}(xeR)}}{\frac{N\hat{x}(xeR)}{N\hat{x}(xeR}} = \frac{N\hat{x}(xeR\&xeA\&xeB)}{N\hat{x}(xeR)}$$

In cases of conditional probability, Pr(A=>B) is the ratio of of the number of members of the intersection of A and B to the number things that are A, not R. Given that Pr(A) = Pr(R=>A),

$$Pr(A{=>}B) = \frac{Pr(A\&B)}{Pr(A)},\ = \frac{Pr(R{=>}(A\&B))}{Pr(R{=>}A)};$$ OK   or $\frac{Pr((R.A)\Rightarrow B)}{Pr(R\cdot A)}$

but,

$$\frac{Pr(R{=>}(A\&B))}{Pr(R{=>}A)} = \frac{\frac{Pr(R\&A\&B)}{Pr(R)}}{\frac{Pr(R\&A)}{Pr(R)}} = \frac{Pr(R\&A\&B)}{Pr(R\&A)} = \frac{Pr(A\&B)}{Pr(A)} = \frac{N\hat{x}(xeA\&xeB)}{N\hat{x}(xeA)}$$

There is a category mistake in supposing that there would be any probability ratio for a nested conditional of the sort used to express conditional probability. For while nested truth-functional conditionals - expressions like '(xeA->(xeB->xeC))' (or equivalently, '(-xeAv(-xeBvxeC))') - define sub-classes which have a definite number of the individual members of the Reference class as their members, expressions like '(xeA=>xeB)' do not define sub-classes of R whose members are members of R. They represent a ratio of the two ratios that two sub-classes of the reference class separately bear to the reference class. The number which expresses this ratio does not represent the number of members in any sub-class of R.

In the model of Figure 1, for example, it makes no sense to speak of a sub-class of R, denoted by {x:(xeA=>xeB)}. xεA ⇒ xεB as determining a sub-class.

Since all probability ratios represent conditional probability ratios, all represent in decimal form an ordered pair of two numbers. When calculating the ratios for any finite model, each of the two numbers is the number of individuals in some sub-class of the reference class; a proper subclass to another proper sub-class in the case of conditional probabilities, of any sub-class to the reference class in other cases. Probability ratios are ratios between numbers of members of distinct sub--classes (proper or improper) of the reference class. Initially, working from the sample or data base, all such sub-classes including the Reference class are finite, and the individual members of the reference class are indivisible. If any expression is to yield a probability ratio, its terms must be gotten by counting the number of members in some sub-class of the reference class. All truth-functional expressions are used to represent the ratio between the number of members of some sub-class of R and the number of members in R. In conditional probability, Pr(B/A, or Pr(x)(xeA=>xeB), the terms xeA and xeB must determine a subclass of R, but the conditional itself can not do so. Only if the antecedent, A, is such that Pr(A)=1.0, will it cancel out so as to leave the consequent,B, to determine a sub-class. [handwritten margin: decimal notation]

Pr(A=>B) in the model of Figure 1, does not represent some sub-class of R which has one sixth of the members that R has. There is no such sub-class. In that model the number of its members would have to be 1.67 of the 10 members of R. Multiplying by R, i.e., 10, the numbers of members which would be determined by other conditionals are likewise ridiculous:

since Pr(-B=>-A)=.286, the members of {x:(-xeB=>-xeA)} = 2.86;
since Pr(A=>-B)=.833, the members of {x:(xeA=>-xeB)} = 8.33;
since Pr(B=>-A) =.67, the members of {x:(xeB=>-xeA)} = .6.7; and
since Pr(-B=>A)=.71, the members of {x:(-xeB=>xeA)=7.1.

Thus the expressions

'Pr(A=>(B=>C))' which supposedly = a ratio, $\frac{\text{'Pr(A\&(B=>C))'}}{\text{Pr(A)}}$,

Pr(B⇒C) is not a pair of nos, each of which is the number of members in some sub-class of R.

make no sense. For '(B=>C)' does not denote a subclass of R; it is used only to establish a ratio of two numbers, namely the number of members of the sub-class of R which is the intersection of (B and C) over the number of members in the sub-class B. In contrast, non-nested Pr(A=>B) will make sense, provided it is a

1st-degree conditional with truth-functional propositional functions for B and C. For then it is defined as the ratio of number of members in the sub-class denoted by the numerator over the number of members in the sub-class denoted by the denominato-r.

It not only makes no sense to speak of nested conditionals, it makes no sense to assign probabilities to conjunctions or disjunctions of whatever conditionals are implicitly used in conditional probability (unless the same sub-class is denoted in all antecedents). For if probability ratios were determinable in that way, the result should represent the frequency of a sub-class relative to the Reference class; but that, we have just seen, will not do. Correlatively, adding or subtracting conditional probabilities will not, in general, yield ratios of the number of members in any sub-class to the number of members in any other sub-class.

On the other hand, there is no problem in multiplying conditional probabilities or in finding the probability of the negated conditionals. The ratio for Pr-(A=>B) is determined by the axiom that Pr-(A) = Pr(1-Pr(A)). There is no difficulty in subtracting the Pr(B/A) from 1.0, to get Pr-(B/A). What is interesting, however, is that Pr-(B/A) or Pr-(A=>B) is equal to the Pr(A=>-B). This is shown intuitively as follows. Pr(A=>B) is a ratio of the number of members in the intersection of A and B over the number of members in B. In the model above, Pr(A=>B) = 1/6. Obviously, the number of members of A that are not in B, i.e., the intersection of A and -B = 1-(1/6) or 5/6. Thus in our model from Figure 1,

1) (1-Pr(A=>B)) = Pr-(A=>B) = Pr(A&-B) = Pr(A=>-B).6
2) Pr(A=>B) = 1/6 =.167, and
and 3) Pr(A=>-B) = Pr(A&-B) / Pr(A) = .5 / .6 =.833,

Hence, 4) Pr(1-Pr(A=>B)) = (1 -.167) = .833 = Pr-(A=>B).

Thus making all probabilities into conditional probabilities causes no problems of nested conditionals.

Nicod's Conditional

---

[6]A more formal proof from Kolmogorov axioms, is:

1)	If Pr(A)>0, Pr(A=>-B) = Pr(A&-B) / Pr(A)	[Df, Pr(B/A)]
2)	Pr(A&-B) = (Pr(A) - Pr(A&B))	[Standard Theorem]
3)	If Pr(A)>0, Pr(A=>-B) = (Pr(A) - Pr(A&B)) / Pr(A)	[2),1), Sub of =s]
4)	If Pr(A)>0, Pr(A=>-B) = 1 - Pr(A&B) / Pr(A)	[3), Arithmetic]
5)	If Pr(A)>0, Pr(A=>-B) = 1 - Pr(A=>B)	[4),Df, Pr(B/A)]
6)	If Pr(A)>0, Pr(A=>-B) = Pr-(A=>B)	[5), Axiom,Pr(-A)]

In his "Studies in the Logic of Confirmation"7, Hempel is confronted with the "paradox of confirmation". Hempel's question is this: given a universal conditional sentence of the form, '(x)(if P(x) then Q(x))', what would constitute confirmatory evidence? Hempel interprets the 'if...then' in the universal conditional as a truth-functional conditional, symbolized as '(x)(Px ->Qx)'. But this gets him into the "paradox of confirmation", that every observation that some particular object is not a Raven constitutes confirmatory evidence for the generalization, "All ravens are black".

This "paradox" has the same source as the divergence of the probability of truth-functional conditionals from conditional probability, namely, that in standard logic, the falsity of an antecedent, '(a is a Raven)', logically implies '(a is a Raven -> a is black)', which is an instance of "All ravens are black".

Nicod had proposed a criterion of confirmation which would avoid these paradoxes8. His criterion suggests a conditional whose truth conditions satisfy the requirements for conditional probability. But Hempel rejects such a conditional, in large part because it conflicts with the principle of Transposition, and defends the truth-functional conditional, claiming that the "impression of a paradoxical situation...is a psychological illusion"9.

We take our cue from Nicod's criterion of confirmation, as not only a way out of the "paradoxes of confirmation", but as suggesting a conditional whose probability coincides with conditional probability. I will show how and why, for such a conditional Transposition, Exportation and Bivalence, must fail to be theorems of logic. Then I will explain why Transposition is so commonly thought to express a logical truth.

Hempel describes Nicod's criterion as follows: According to Nicod, a universalized conditional, '(x)(if P(x) then Q(x)', is confirmed by an object

> "...if and only if it satisfies both the antecedent (here: 'P(x)') and the consequent (here 'Q(x)') of the conditional; it disconfirms the hypothesis if and only if the satisfies the antecedent, but not the consequent of the conditional; and (we add this to Nicod's statement) it is neutral or irrelevant, with respect to the hypothesis if it does not satisfy the antecedent..." 10

Thus, like the conditional implicit in conditional probability,

---

7
See Carl G. Hempel, **Aspects of Scientific Explanation**, Free Press, 1965, pp 3-53

[8]**Foundations of Geometry and Induction** (translated by P.P.Wiener) London, 1930; p.219; The Logical Problem of Induction is also included in another translation: Jena Nicod, **Geometry and Induction**, University of California Press, 1970, see p. 189.

[9]Hempel, opus cit., p. 18.

[10]Hempel, Opus Cit., p.11.

Nicod's considers only cases in which the antecedent obtains, treats the conditional as being neither true nor false (neutral) in other cases where the antecedent is not satisfied. Hempel adds that Nicod "states explicitly what is perhaps the most common tacit interpretation of the concept of confirmation."

Hempel's main argument against Nicod's conditional is that it requires abandoning the laws of Transposition:

> "Consider the two sentences
>
> S1: '(x)[Raven(x) -> Black(x)]'
>
> S2: '(x)[-Black(x) -> -Raven(x)]'
>
> (i.e., 'All Ravens are black' and 'Whatever is not black is not a raven'), and let **a**, **b**, **c**, **d** be four objects such that **a** is a raven and black, **b** a raven but not black, **c** not a raven but black, and **d** neither a raven nor black. Then according to Nicod's criterion, **a** would confirm S1, but be neutral with respect to S2; **b** would disconfirm both S1 and S2; **c** would be neutral with respect to both S1 and S2, and **d** would disconfirm S2 but be neutral with respect to S1.
>
> "But S1 and S2 are logically equivalent; they have the same content, they are different formulations of the same hypothesis, and yet, by Nicod's criterion, either of the objects **a** and **d** would be confirming for one of the two sentences, but neutral with respect to the other."[11]

But this argument has force only if one agrees that S1 and S2 are logically equivalent, i.e, that Transposition holds. It begs the question if one accepts the conditional which Nicod's account suggests, and which is necessary for the conditional implicit in conditional probability. For then 'if x is a raven then x is black' is not logically equivalent to 'if x is not black then x is not a raven'. Among other things, substitution of contrapositives does not preserve probability ratios, as we have seen. It also does not lead to the implausible consequence that one way to confirm the proposition that all ravens are black is to set about examining all non-black objects.

Hempel's rejection of Nicod's conditional is presented as a proof that that conditonal would violate one of the basic conditions of adequacy for any theory of confirmation, namely the "Equivalence Condition". The Equivalence Condition is surely one we must accept in some very sound, ordinary sense. It says that if S3 confirms S1, and S1 is logically equivalent to S2, then S3 confirms S2. This seems but another version of PR1, the first of axiom of probability. But also, we assume agreement that 'S1 is logically equivalent to S2' entails that 'S1 if and only if S2' is a logical truth. But with the conditional Nicod suggests the formula '(x)(xeA => xeB) if and only if (x)(-xeB => -xeA)' is not a logical equivalence. For in giving a different meaning to 'if...then', we get a different meaning for 'if and only if', as well as for "logically equivalent". Thus the rejection of Transposition due to a different meaning for the conditional need

[11] Hempel, **Aspects of Scientific Explanation**, p.12.

not entail the rejection of the Equivalence Condition for any adequate theory of confirmation, though it may change the the meaning of 'logically equivalent'. This is indeed what happens with Nicod's conditional.

What is implicit in Nicod's conditional may be made more explicit in the formal logic of these conditionals sketched below and in and related semantic concepts of "analytic containment" and "referential synonymity".

Let us use the same model we used for conditional

R:	1	2	3	4	5	6	7	8	9	10
A:	+	+	+	+	+	+				
B:						+	+	+		

III

For purposes of this paper, I shall use (instead of 'All ravens are black') a variety of universal conditionals about what I shall call a 10L-Figures; these are figures like Figure 1 below which have 10 boxes or "locations", labeled 'L1, L2,..etc.

L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
				a2	a3			a5	a6
		b5		b2	b4		b6	b3	b1

Figure 1

The 10L-Figures vary in having a's {a1,a2,a3,...}, and or b's {b1,b2,b3,...}, or in one or more, or none, of the locations. The first universal C-conditional to be considered is:

1) For any location, L, in Figure 1,
if an 'a' occurs in location L,
then a 'b' occurs in location L. [Form='If A then B']

Here we treat 'if...then' as a Nicod-, or C-conditional, and we intend that 'Figure 1' denotes the Figure 1 above, and that 'L1','L2', etc., are used to denote the rectangles in Figure 1 which contain tokens of those signs.

In this situation, though 1) clearly refers to Figure 1, what it talks about in Figure 1 are only those states of affairs in L5, L6, L9 and L10 in which there are a's; it says that in those locations there are also b's. It does not talk about anything present or absent in any other locations in Figure 1.

In contrast, consider the contrapositive of 1), the C-conditional, of the form 'If -B then -A':

1') For any location, L, in Figure 1,
if it is not the case that a 'b' occurs in location L,
then it is not the case that an 'a' occurs in location L.

What 1') talks about is <u>not</u> what occurs in locations picked out by the antecedent of 1) i.e., locations which have a's in them. Nor does it talk about all locations which have b's in them. It says that in locations which <u>do not have b's</u>, namely, L1, L2, L4, and L7, it is also the case that there are no a's.

Now it happens to be true of Figure 1 that both 1) and its contrapositive 1') are true. In other words, in <u>this case</u>, 1) is true of Figure 1, if and only if 1') is true of Figure 1, and that seems to be an instance of a transposition principle.

But the question is, is '1) is true if and only if 1') is true' a logical truth? They have the forms respectively, of 'If A then B' and 'If -B then -A'; are all pairs of statements having these forms <u>logically equivalent</u>?

IV

The term 'logically equivalent' in standard logic is equated with 1) truth-functional equivalence - wherever the truth-table for the truth-functional biconditional 'A≡B' comes out all T's, and with 2) quantificational equivalence, which is a bit more complicated, but which holds only if truth-functional equivalences hold of all of a universal quantifier's instances.

In the semantic theory I wish to advance, in place of standard truth-functional semantics, two expressions are logically equivalent if and only if they are **referentially synonymous**. We need not go into "referentially synonymous" beyond saying that if A is referentially synonymous to B, then 1) A <u>talks about</u> all and only the same things that B <u>talks about</u>, and 2) A says all and only the same things about whatever it talks about that B says about those things and vice versa. Clearly, not all truth-functionally equivalent statements are referentially synonymous; notably the pairs {'(Fa.-Fa)', '(Gb.-Gb)'} and {'(A.-A.B)', '(A.-A.-B)'}.

In the semantics of C-conditionals, truth-functional equivalents are not always "logically equivalent" in this sense. For 'A is logically equivalent to B' is true iff and only the statement 'A if and only if B' is <u>logically</u> true; but when 'if and only if' is taken as a C-biconditional rather than a truth-functional biconditional, "logically equivalent" gets a different meaning12.

Thus for 1) and its contrapositive, 1'), to have the same meaning, in the sense of being referentially synonymous, they must talk about and refer to the same entities. This is what they do <u>not</u> do. The statement 1) <u>talks about</u> only those states of affairs in L5, L6, L9 and L10 which contains a's in Figure 1, while the statement 1') <u>talks about</u> only the states of affairs in

---

12

On the other hand, all logically true truth-functional biconditionals remain logical truths since they are always negations of inconsistent statements; this is not the same as to say the two components are logically equivalent in our new sense.

locations which do not have b's, namely, L1, L2, L4, and L7. Hence 1) and 1') do not talk about the same things, hence do not mean the same thing (in the sense of referentially synonymy), and hence do not "have the same content" and are not "different formulations of the same hypothesis" (to use Hempel's words). Since they are not referentially synonymous, on this theory of logic they are not logically equivalent.

V

The point can also be made syntactically in terms of the calculus of analytic equivalence - i.e., in terms of its rules for theoremhood. Within standard logic it is possible to distinguish syntactically a sub-class of truth-functionally equivalent pairs, which I shall call the class of "analytically equivalent" pairs. Let the calculus AEQ, have '&' and '-' as primitive sentence connectives, the usual definitions and rules of formation, the axiom schemata,

AEQ1.	A aeq (A&A)	[&-Idempotence]
AEQ2.	(A&B) aeq (B&A)	[&-Commutation]
AEQ3.	(A&(B&C)) aeq ((A&B)&C)	[&-Association]
AEQ4.	(A&(BvC)) aeq ((A&B)v(A&C))	[&v-Distribution]
AEQ5.	--A aeq A	[Double Negation]

and the rule of transformation,

R1. If A aeq B, then C aeq C(B//A)[13]

In this system, two expressions can not be analytically equivalent if either contains a variable the other does not, or if there is an atomic sentence which occurs negatively, or positively, in one but not in the other[14].

By this criterion of "analytic equivalence" 1) is not analytically equivalent to 1'), because the atomic components of 1) and 1') occur positively in 1) but not in 1') and negatively in 1') but not in 1).

All analytically equivalent pairs in standard logic are can be proved to be referentially synonymous and thus logically equivalent in our sense. They are also all truth-functionally equivalent; but, as mentioned, not all truth-functionally equivalent pairs are analytically equivalent, and thus not all such pairs are logically equivalent in our sense. Hence, there is a formal proof that 1) and 1') are not logically equivalent in the logic of C-conditionals.

---

[13]For 'A aeq B' read "statement of the form A is analytically equivalent to statement of the form B"; for 'C(B//A)' read "a statement like C except that one or more occurrences of B in C are replaced by A".
[14]A component occurs negatively, if and only if, in primitive notation it occurs within the scope of an odd number of negation signs; otherwise it occurs positively. Due to Herbrand.

VI

It might thought be that although 1) and 1') are counter-examples (using C-conditionals), for the statements,

'If A then B' is referentially synonymous to 'If -B then -A'
'If A then B' is analytically equivalent to 'If -B then -A'

(thus the contrapositives are not always logically equivalent), that perhaps we could never have 'If A then B' true and 'If -B then -A' not true in the same context. Let us express this principle as

'If A then B' is true if and only if 'If -B then -A' is true.

In the case of Figure 1, we can see that though 1) and 1') are not logically equivalent by the definitions involving referential synonymy and analytic equivalence, nevertheless, they were both true together; one was true if and only if the other was true in the case of Figure 1. The question now is, could it ever be the case that one was true and the other was not true? In other words, may it not be that semantically at least, they are truth-functionally equivalent?

'Figure 1', as used in this paper, is the name of an actual state of affairs on a piece of paper in the actual world. Different conditional statements can be made with reference to Figure 1, some of which would be true, and some of which would be false. The meaning of a C-conditional does not depend on its being true or false. The following is a false C-conditional:

2) For any location, L, in Figure 1,
if an 'b' occurs in location L,
then a 'a' occurs in location L.

for (following Nicod's Criterion), this picks out the locations L3, L5, L6, L8, L9, and L10, which have b's in them, and says that they all have a's in them, which is false. The meaning of the statement 2) would have been the same, however, had Figure 1 been drawn differently so as to make 2) true. The fact that Figure 1 is just what it is, is what makes 2) false..

Figure 1

L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
				a2	a3			a5	a6
		b5		b2	b4		b6	b3	b1

But it also seems clear with respect to Figure 1, that if 2) is false, then its contrapositive,

2') For any location, L, in Figure 1,
if it is not the case that an 'a' occurs in location L,
then it is not the case that a 'b' occurs in location L.

must be false also. Thus in Figure 1 it seems that when a statement is false, its contrapositive must be false also. In other words,

'If A then B' is false <u>if and only if</u> 'If -B then -A' is false.

apparently holds for Figure 1; and this also seems to reinforce this new formulation of transposition.

If the conditional is <u>true</u>, then <u>there exists</u>, in some context being referred to, a set of two or more locations, at least one of which contains the antecedent, and all locations which contain the antecedent contain the consequent. The question now is, Whenever this is the case, must the contrapositive conditional also be true?

But a problem arises. In the semantics of C-conditionals,if no location in the context referred to contains the antecedent, then the conditional can not be true, for then it is talking about something that isn't there, in the field of reference[15]. Nor can it be false; for to be false in the semantics of the C-conditional is to satisfy the antecedent, but not the consequent. A C-conditional can be neither true nor false in some cases.

Now consider Figures 2 and 3 and the statements 3) and 3') below. The statement 3) has the form 'If A then B' and is true of Figure 2 below but is not true of Figure 3; and the statement 3') has the form 'If -B then -A' and is true of Figure 3, but is not true of Figure 2. This situation is due in each case to the fact that the antecedent is not satisfied: what is being talked about does not exist in the field of reference.

L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
				a2	a3			a5	a6
b8	b7	b5	b10	b2	b4	b9	b6	b3	b1

Figure 2

3) For any location, L, in Figure 2,
<u>if</u> an 'a' occurs in location L,
<u>then</u> a 'b' occurs in location L. [Form:'If A then B']

L1	L2	L3	L4	L5	L6	L7	L8	L9	L10

[15] This principle, that a C-conditional is neither true nor false when referred to contexts in which the antecedent is not satisfied, is an extension of the property it must have when serving as the conditional of conditional probability; the probability of B given A, is undefined when the probability of A is 0. This semantic property also comes in handy, however, in dealing with other problems like problems relating to non-referring terms, or inconsistent predicates. It presupposes a semantics which distinguish 'is false' from 'is not true' however; a plausible semantics which is demanded by C-conditionals.

				b2	b4			b3	b1

Figure 3

3') For any location, L, in Figure 3,
if it is not the case that a 'b' occurs in location L,
then it is not the case that an 'a' occurs in location L.
[Form:'If -B then -A']

The statement 3) has the same form and the same meaning as 1) (where meaning does not include contingent actual facts about the contents of 'Figure 1' or 'Figure 2' or 'Figure 3'). But 3), and thus 1), can not be said by virtue of their meanings to imply the truth or falsehood of 3') or of 1'), for though both contrapositives are true of Figure 3, they are not true (or false) of Figure 2, since the antecedent does not apply to anything in the field of reference.

Again, consider the contrapositives, 4) and 4'); 4') is true of Figure 3, but 4) is not true, since the antecedent is not satisfied:

4) For any location, L, in Figure 4, [Form:'If A then -B']
if an 'a' occurs in location L,
then it is not the case that a 'b' occurs in location L.

4') For any location, L, in Figure 4, [Form:'If B then -A']
if a 'b' occurs in location L,
then it is not the case that an 'a' occurs in location L.

Or, consider the contrapositives, 5) and 5'); 5) is true in Figure 2, but 5') is not true of Figure 2 since its antecedent is not satisfied:

5) For any location, L, in Figure 4,
if it is not the case that an 'a' occurs in location L,
then a 'b' occurs in location L.

5') For any location, L, in Figure 4,
if it is not the case that a 'b' occurs in location L,
then an 'a' occurs in location L.

Again, because their antecedents can not be satisfied, neither one of 4) and 4') can be true in Figure 4, though they are each contrapositives of the other, having the forms 'if A then -B' and 'if B then -A' respectively:

Figure 4

L1	L2	L3	L4	L5	L6	L7	L8	L9	L10

And likewise, neither one of 5) or 5'), which are contrapositives of each other, with the forms, 'If -A then B' and 'If -B then A', can be true of Figure 5, since neither antecedent is satisfied.

Figure 5

L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
a1	a2	a3	a4	a5	a6	a6	a7	a9	a10
b1	b2	b3	b4	b5	b6	b7	b8	b9	b10

From these considerations it is clear that even where 'If A then B' is true, 'if -B then-A' may not be true, and thus that transposition for C-conditionals fails even if only truth-functional equivalence is being claimed.

Thus we have shown that none of the following forms of transposition principles are laws in the logic of C-conditionals:

A) 'If A then B' is <u>referentially synonymous</u> with
'(if -B then -A)'
B) 'If A then B' is <u>analytically equivalent</u> to
'(if -B then -A)'3
C) 'If A then B' is true <u>if and only if</u>
'(if -B then -A)' is true.
D) 'If A then B' is <u>truth-functionally equivalent</u> to
'If -B then -A'

Of course if 'if...then' is interpreted throughout as the truth-functional conditional, then all of A) through D) are true. For, the conditionals in quotes are then taken in both standard logic and analytic logic to mean '(-AvB)' and '(Av-B). Thus A) and B), for example, become, analytically equivalent to

A') '(-A v B)' is <u>referentially synonymous</u> with
'(--B v -A)'

B') '(-A v B)' is <u>analytically equivalent</u> to
'(--B v -A)'

and these are easily proven as theorems in analytic logic, or in the logic of referential synonymy. If the 'if and only if' in C) is treated as a truth-functional biconditional as well as the 'if...then's in quotes in C), then C) would be derivable in the truth-functional semantics of standard logic. Further, with truth-functional interpretaions of the 'if...then's in D), D) is a metatheorem of standard logic - it says the same thing as what is meant by standard logicians when <u>they</u> say ''If A then B' is <u>logically equivalent</u> to '(if -B then -A)''.

Despite these results, it can be shown, that **Modus Tollens** is is provable in a logic of C-conditionals with a truth-operator. I.e., "If it is true that (If A then B) and it is true that -B, then it is true that -A" can be sustained. Since the logic of the C-conditional also avoids Exportation, the failure of transposition principles will not lead to a rejection of **Modus**

Tollens16. The proofs of this point will be not be give here.

VII

The principles of transposition have a very strong intuitive appeal. Most people, I think, would say that if 'If A then B' were true, then 'If -B then -A' would have to be true also. On the other hand, the account of conditionals given above is also very close to what people would say was meant by a conditional statement - including the notion that a conditional is neither true nor false where its antecedent does not apply to anything. In my opinion, this conditional is closer to a common sense account of the conditionals than the truth-functional account which makes every conditional true if its antecedent is false, or if its consequent is true. How can these two incompatible intuitions be explained?

We may begin by noting that in all possible Figures in which the the set of locations includes both some locations in which the antecedent is true and others in which it is not true, and similarly for the consequent, both C) and D) will be true even if the 'if and only if' in C) is taken as a C-biconditional. Among the possible 10L-Type Figures with a's and b's distributed among the locations, by far the greater number of possibilities will be those in which some locations have a's and some don't and some locations have b's and some don't17.

In general, when people assert universal conditionals both the antecedent and the consequent, taken singly, describe states of affairs which hold in some locations (at some times, in some contexts) and do not hold in others. Even in standard quantified logic, the range of the individual variables is generally taken to refer to the set of all entities whatever, and it is often asserted that no significant predicates18 (or, in any case, extremely few) either apply, or fail to apply, to all entities. Thus it may be suggested that the reason why the layman has difficulty thinking of counter-examples to principles of transposition is because it would rarely be the case that he would think of an example in which either the antecedent or the consequent held, or failed to hold, universally in his field of reference.

---

[16] I.e.,'If (T(If A then B)) & T(-B)) then T-A)'
does not imply, in the logic of C-conditionals,
'(If T(If A then B) then (If T(-B) then T-A))'
The elimination of Exportation is also necessary to get a conditional which expresses conditional probability.

[17] Of the 1,048,576 possible 10L-Type Figures, 4092 would have all, or no a,s or b's, making one or both of the contrapositives not true, or not-false, or not-true when the other was true, or not-false when the other was false. This is less than 4/10s of 1% of the total possibilities.

[18] 'significant predicates' exclude tautologous or inconsistenct predicates.

A second reason perhaps, is that many people in contemporary semantics insist upon equating 'is false' with 'is not true', and reject in the name of 'excluded middle' the possibility of meaningful indicative sentences which are both not true and not false. Movements like "Free logic", and others, are treated at the present time as 'deviant', and have not yet gained universal acceptance due to the immense power and effectiveness of standard mathematical logic and its generally accepted semantic theories.

VIII

The problem which I would like to present to logicians is whether they want 1) a logic of conditionals which entails a variety of "paradoxes" (of which the "paradoxes of confirmation" is but one sub-class) and preserves transposition as a law of logic, or 2) a logic of conditionals which eliminates this and many other "paradoxes, but does not include transposition and exportation among its laws. Or, putting the question more generously: would it not be worthwhile to recognize, in addition to the "truth-functional conditional" another kind of conditional, the C-conditional, with its own logic and a different semantics which allows us to avoid many, if not all, of the "paradoxes" of the the truth-functional conditional, at the small price of giving up some principles currently considered laws of logic?

<u>Miscellaneous left-overs:</u>

We will discuss this by investigating the truth or falsity of the following statements:

1) 'If A then B' <u>truth-functionally equivalent</u> to 'If -B then -A'
2) 'If A then B' is <u>analytically equivalent</u> to '(if -B then -A)'
3) 'If A then B' is <u>referentially synonymous</u> with '(if -B then -A)'
4) 'If A then B' is true if and only if '(if -B then -A)' is true.
5) ''If A then B'is true' is <u>analytically equivalent</u> to ''(if -B then -A)' is true'.
6) ''If A then B'is true' is <u>referentially synonymous</u> with ''(if -B then -A)' is true'.

<u>5) and 6) have not been discussed: they depend on a semantics of the truth-operator, and its relation to SYN and AEQ.</u>

For, 'If A then B' is talking about A's and claiming a correlation between As and Bs. But 'If -B then -A' is talking about very different things non-Bs, and trying to correlate them with non-As. This brings on the Raven Paradox.

To verify or confirm (x)(If x is R then x is B) we confine our attention to Ravens; this is what the claim is about, what the sentence as a whole refer to - ravens and their properties.

To talk about non-black things is to not to talk about ravens. [File:CONDPRB1 - Angell]

*Projection: If we choose one piece of fruit, x, from the barrel what are the probabilities with respect to its being an apple or being brown? Assuming any individual piece could equally possibly be chosen, these four ratios are viewed as the 'probabilities' of getting a brown apple on a given choice.
Pr(A&B)=.01 Pr(A&-B)=.09 Pr(-A&B)=.10 Pr(-A.-B)=.80.

There is a misleading tendency to identify 'Pr(A)=1.0' with ''A'is True' and `Pr(A)=0.0' with ''A' is False', so that the intervening values, 0<Pr(A)<1, representing degrees of truth or something. Rather, each probability ratio represents the ratio of true instances to total instances of a universally quantified propositional function.

To be sure, if any universally quantified proposition such as (x)(Fx=>Gx) is true, then Pr(x)(Fx=>Gx) = 1.0. But if (x)(Fx=>Gx) has any false instances at all, Pr(x)(Fx=>Gx)<1 and it is not true. Existentially quantified statements are false if they have a probability ratio of 0, but a probability ratio between 0 and 1 tells us nothing about whether they are true. No expression can be both true and false; truth and falsehood are, at least, contraries19. To allow that a quantified statement is false but yet "true to a certain degree", and that its probability ratio measures its degree of truth is **prima facie** incompatible with the contrariety of truth and falsehood. To avoid inconsistency we must either introduce two meanings of 'true' and/or 'false', or drop the principle that 'true' and 'false' are contraries, or perhaps abandon 'truth' and 'falsehood' entirely in favor of probability ratios. To retain both the univocality of 'true' and 'false' and their contrariety, one must hold that universal statements are either true or false (false if any instance is false), and that the "probability ratios" assigned to them represent, not degrees of truth, but ratios (actual or expected) of true instances to total instances.

The initial determination of a such ratios are based on statistical distributions of members of some finite class among its sub-classes, either within a sample, or by some **a priori** assignment of initial ratios. Projections of such ratios onto classes which extend beyond the initial data, or assignments of such ratios to events nor included in the data as their 'chances' or 'probabilities' of occurring are, do not change the initial computations of ratios from the data base.

It can also work for <u>all</u> cases, including conditional probabilities expressed as the probability of a conditional '(A=>B)', if the conditional involved is not subject to the Law of Bivalence, i.e., is allowed to be neither True nor False in instances

---

[19]This does not entail that they are sub-contraries; that no expression can be both not true and not false.

Pr(∃x) Fx = 0 does say (∃x)Fx is ~~false~~ never true (. Pr(x)Fx = 0 } = i.e. -(∃x)Fx is true, (x)-Fx is true

Pr(∃x) Fx = .2 does not say '(∃x)Fx' is true, or false

Pr(∃x) Fx = 1.0 does say '(∃x)Fx' is false

~~Pr (∃x) F~~ Pr(Fx) = .2 ∴ Pr((∃x) Fx) = 1.0 Pr(∃x)-Fx = 1.0

Pr(Fx) = 0 ∴ (x)-Fx, -(∃x)Fx = 1 true

Pr(Fx) = 1.00 ∴ (∃x)Fx, (x)Fx = 1 = true

Pr(x)Fx = .2 ∴ ~(x)Fx, (∃x)-Fx } is true

(x)Fx = 1.0 ∴ (x)Fx is true

(∃x)Fx = 1.0 (∃x)Fx = True

in which its antecedent is not satisfied.