30. WHAT IS WRONG WITH VERISIMILITUDE?

JOSEPH WAYNE SMITH
THE UNIVERSITY OF ADELAIDE

ABSTRACT. Karl Popper introduced the idea of verisimilitude to explicate the intuitive idea that a theory $T_2$, even though it is strictly speaking false, may be closer to the truth than a competitor $T_1$. However, as is now well known, the results of Pavel Tichý, John Harris and David Miller establish that on Popper's qualitative theory of verisimilitude, a theory $T_2$ could be closer to the truth than another theory $T_1$ only if $T_2$ contains no false sentences. This result has been taken universally to show the inadequacy of Popper's original account of verisimilitude, since the Miller-Tichý-Harris Theorem conflicts with the very basic intuition which first led Popper to formulate his theory.

In this paper I shall first review the Miller-Tichý-Harris Theorem and examine a number of attempts to salvage the concept of verisimilitude. It will be argued that none of these attempts is successful. Finally an alternative, simple and intuitively satisfactory account of verisimilitude will be offered. This account will be along the lines first suggested by Popper, but it is not subject to any known limitation theorem. Further, the account is capable of giving verisimilitude orderings between not only scientific theories, but philosophical theories as well. This will be achieved without the use of the excessive formalism which dominates the contemporary verisimilitude research programs.

The Passion for exactness is certainly a noble passion, but, like any other, it may make fools of us.1

I. INTRODUCTION

Science, we are often told, is distinguishable from other cognitive enterprises by its "progressiveness"; scientific knowledge by contrast with philosophical or theological knowledge (if these latter subjects are taken to make cognitively meaningful knowledge claims at all), exhibits "continued growth."2 Science is taken to be cumulative regardless of crisis and revolution and hence "capable of unrestricted growth towards
universal coerciveness of argument and evidence.\textsuperscript{71} Few deny that science is in some sense "progressive"; even Paul Feyerabend claims that his epistemological anarchism "helps to achieve progress in any one of the senses one cares to choose."\textsuperscript{74} For Feyerabend it is the metascience of contemporary philosophy of science which, if consistently adhered to, would evaporate or stagnate scientific progress.\textsuperscript{5}

Various philosophical questions can be asked about the concept of 'scientific progress'. First is the semantical and conceptual question of how the expression 'scientific progress' is to be explicated. Second is the epistemological question of how, given an explication of the expression 'scientific progress', are we to identify progressive theories in science? What are the criteria by which progressive scientific theories are distinguished from non-progressive or degenerating scientific theories? My concern in this paper is with a defense of a realistic view of scientific progress, which proposes that the answer to the semantic and conceptual question is that if scientific progress is taken to occur between the theories $T_n$ and $T_{n-1}$, then $T_n$ is closer to 'the truth' than $T_{n-1}$. This is to say that $T_n$ has a higher verisimilitude than $T_{n-1}$.\textsuperscript{6}

The idea of verisimilitude was introduced by Popper to explicate the intuitive idea that a theory $T_2$, even though it was strictly speaking false, may still be closer to the truth than a competitor $T_1$. In Conjectures and Refutations Popper wrote:

> Ultimately, the idea of verisimilitude is most important in cases where we know that we have to work with theories which are at best approximations—that is to say, theories of which we actually know that they cannot be true. . . In these cases we can still speak of better or worse approximation to the truth (and we therefore do not need to interpret these cases in an instrumentalist sense).\textsuperscript{7}

However, as is now well known, the results of Pavel Tichý,\textsuperscript{8} John Harris,\textsuperscript{9} and David Miller\textsuperscript{10} (henceforth denoted by the expression 'the Miller-Tichý-Harris Theorem') establish that a theory $T_2$ could be closer to 'the truth' than another theory $T_1$, on Popper's qualitative theory of verisimilitude, only if $T_2$ contains no false sentences. This result 'has been universally taken to demonstrate the inadequacy of Popper's qualitative theory of verisimilitude.

In this paper I shall first review the the Miller-Tichý-Harris Theorem and examine a number of attempts to salvage the concept of verisimilitude. It will be argued that none of these attempts is successful. Finally an alternative, simple, and intuitively satisfactory account of verisimilitude will be offered. This account will be along the lines first suggested by Popper, but it is not subject to any known limitation theorem. Further, the account is capable of giving verisimilitude orderings between not only scientific theories, but philosophical theories as well, so that the presented account is more general in its applications than a number of the major competing accounts of verisimilitude in the literature.\textsuperscript{11}
II. The Miller–Tichy–Harris Theorem

Since it has been proposed by Chris Mortensen\(^\text{12}\) that the Miller–Tichy–Harris Theorem can be escaped while retaining Popper’s original theory of qualitative verisimilitude, by modifying the classical logical base on which the results depend, a more general formulation of the Miller–Tichy–Harris Theorem is required. Consider a first order formal system \(L\), which may have a denumerable number of constants, predicates and variables, such that the set of wffs of \(L\) are closed under conjunction \(\&\), disjunction \(\lor\), negation \(\neg\), and implication \(\rightarrow\). The usual syntactical formation rules are presupposed, as is a rich metalanguage containing set-theoretical signs of a standard set-theory (e.g., Zermelo–Fraenkel set-theory). \(L\) may also contain modal operators, various types of functionals and any number of special predicates—whether it does or not will be of interest to our argument here. The logic \(L_0\) of \(L\) is a subset of the set of wffs of \(L\) closed under the rule of uniform substitution. If \(\phi \in L_0\), then \(\phi\) is said to be a theorem of \(L_0\), written as \(\vdash \phi\). \(L_0\) is said to be an implication logic if and only if:

1. If \(\vdash \phi\) and \(\vdash \phi \rightarrow \beta\), then \(\vdash \beta\).

\(L_0\) is said to be an \(L_0\)-theory relative to logic \(L_0\) if and only if both (2) and (3) hold:

2. If \(A \subseteq L\) and \(\phi \in A\) and \(\vdash \phi \rightarrow \beta\), then \(\beta \in A\).

3. If \(A \subseteq L\) and \(\phi \in A\) and \(\beta \in A\), then \(\phi \land \beta \in A\).

If \(L_0\) is classical logic, then \(A\) is said to be a classical theory.\(^\text{13}\) \(A\) is inconsistent if and only if for some \(\phi\), \(\phi \in A\) and \(\neg \phi \in A\). \(A\) is trivially inconsistent if and only if \(A = L\). \(A\) is incomplete if and only if for some \(\phi\), both \(\phi \notin A\) and \(\neg \phi \notin A\) and is complete if and only if it is not incomplete. The rule \(\gamma\) holds for \(A\) if and only if \(\phi \in A\) and \(\neg \phi \lor \beta \in A\), then \(\beta \in A\). \(A\) is prime if \(\gamma\) holds and non-prime if it does not.

Let \(A\) and \(B\) be classical \(L_0\) theories and let \(T\) be the set of true sentences of \(L\), and \(F\) the set of false sentences of \(L\) and \(T \cup F = L\). \(A_T\) is the set of true sentences of \(A\), and \(B_T\) the set of true sentences of \(B\). \(A_F\) is the set of false sentences of \(A\) and \(B_F\) the set of false sentences of \(B\). Then Popper’s qualitative definition of verisimilitude is as follows:

\[ (PQDV) A \text{ has a greater verisimilitude than } B, \text{ i.e., } A >_v B =_{df} (B_T \subseteq A_T) \land (A_F \subseteq B_F) \lor (B_T \subseteq A_T) \land (A_F \subset B_F). \]

The Miller–Tichy–Harris Theorem is now stated and proved:

\[ (MTHT) \text{ If } A \text{ is false (i.e., } (\exists \phi) (\phi \in A) \land (\phi \in F)), \text{ then } \neg (A >_v B), \text{ (i.e., } A >_v B \text{ then } A \subseteq T). \]

\textbf{Lemma 1:} If \(A\) and \(B\) are classical \(L_0\)-theories and \(\vdash \neg \alpha \land a \land b \rightarrow a\) and if \(a \in F\) then \(a \land b \in F\), then if \((B_T \subseteq A_T) \land (A_F \subseteq B_F)\) then \(A \subseteq T\).
Proof of Lemma 1: Suppose that \( A \not\subseteq T \) for reductio ad absurdum. Let \( f \in A \) and \( f \in F \) and let \( a \in A \cap B \), so that \( a \in A, a \in T, \) and \( a \notin B \). Since \( a \in A \) and \( f \in A \), then \( a \& f \in A \) as \( A \) is an \( L_0 \)-theory. Since \( f \in F \), \( a \& f \in F \) and \( a \& f \in AF \). But since \( a \notin B \), then \( a \& f \notin B \) and hence \( a \& f \notin BF \). But as \( a \& f \notin BF \) and \( a \& f \in AF \), we obtain a contradiction from the assumption that \( A \not\subseteq T \), for \( (AF \subseteq BF) \rightarrow ((a \& f \in AF) \rightarrow (a \& f \in BF)) \). Hence there is no \((f \in A) \& (f \in F)\). Hence \( A \not\subseteq T \).

Lemma 2: If \( A \) and \( B \) are classical \( L_0 \)-theories and
\( 1 \) \( a \rightarrow (a \lor b) \); \( 2 \) \( \gamma \) holds for \( A \);
\( 3 \) if \( a \in T \) then \( a \lor b \in T \), then
if \( (B \subseteq A) \) \& \( (AF \subseteq BF) \) then \( A \not\subseteq T \).

Proof of Lemma 2: Suppose that \( A \subseteq T \) for reductio ad absurdum. Let \( f \in A \) and \( f \in F \). Let \( b \in B \cap AF \) so that \( b \in B, b \in F \), and \( b \notin A \). Since \( b \in B \), \( \neg f \lor b \in B \). Since \( f \in F \), then \( \neg f \lor b \in T \), hence \( \neg f \lor b \in Br \).

But \( b \in A \) and \( f \in A \) so \( \neg f \lor b \notin A \). Hence \( \neg f \lor b \notin A \).

But \( \neg f \lor b \notin A \) and \( \neg f \lor b \in Br \) contradicts \( B \subseteq A \).

Hence the assumption that \( A \subseteq T \) leads to contradiction and there is no \( f \in A \) and \( f \in F \). Hence \( A \not\subseteq T \).

Proof of (MTHT): The proof is immediate from Lemmas 1 and 2.

This result is, as I have said, devastating for Popper’s original account of verisimilitude. I shall discuss in this paper the attempts of Tichy, Tuomela and Niiniluoto, Perry, Mott, Bunge, Wójcicki, Krajewski and Rosenkrantz, and Mortensen to avoid the difficulties facing Popper’s original notion of verisimilitude. Each of these accounts, it will be argued, is seriously defective in a number of ways (many being open to straightforward counter-examples).

III. Tichy’s Account of Verisimilitude

Pavel Tichy, in criticizing Popper’s probabilistic theory of verisimilitude, considered an elementary weather language \( Lw \) containing no predicates and only three primitive sentences: ‘it is raining’, ‘it is windy’ and ‘it is warm’—abbreviated as ‘p’, ‘q’ and ‘r’ respectively. Tichy proposed for a simple language like \( Lw \) based only on propositional logic, that the distance between two constituents be defined as the number of primitive sentences negated in one of the constituents, but
WHAT IS WRONG WITH VERISIMILITUDE?

not in the other. The verisimilitude of an arbitrary sentence $a$ can be defined as the arithmetic mean of the distances between the true constituent $t$ and the constituents appearing in the disjunctive normal form of $a$.

David Miller,\textsuperscript{15} criticized Tichý's proposal for being "language-dependent". Stated more precisely, Tichý's proposed orderings by truth-likeness can be reversed by simple linguistic reformulations.\textsuperscript{16} In Tichý's example, $p$, $q$, and $r$ are three independent sentences of $L_w$, all of which are true. There are eight maximally consistent sentences of the sentence algebra of $L_w$, only one of which, $p \& q \& r$, is true. According to Tichý, the maximally consistent sentence $p \& q \& r$ is closer to the truth than is $\neg p \& \neg q \& \neg r$. But let $d = p \iff q$ and $e = p \iff r$. Consider the sentence algebra of the maximally consistent sentences generated from $(p, d, e)$. Then the true maximally consistent sentence is $p \& d \& e$, with the first false maximally consistent sentence now as $\neg p \& \neg d \& \neg e$, and the second as $\neg p \& d \& e$, which reverses Tichý's verisimilitude ordering.

Stated more generally, Popper's definition of verisimilitude is not invariant in its orderings with respect to logically equivalent ways of representing the two theories $A$ and $B$. A simple argument for this, formulated by Chris Mortensen, is as follows. Let $A = \{a_1, a_2, \ldots\}$, $B = \{b_1, b_2, \ldots\}$, $A_1 = \{a_1, a_2, \ldots\}$, $B_1 = \{b_1, b_2, \ldots\}$ and let $A >_B$ just in case $A \subseteq B$. Consider the claim that if $A >_B \rightarrow (B_1 \subseteq A_1) \& (A_1 \subseteq B_1)$. If $A_1 \not\subseteq T$, then there exists an $f \in A_1$. Also $a_1 \in A - B$, so that $A_1 = \{f, a_1, a_2, \ldots, a_n\}$. Hence $A_1$ and $A_1^f$ are logically equivalent. But it is provable that $\neg((B_1 \subseteq A_1^f) \& (A_1^f \subseteq B_1))$, since by the Miller-Tichý-Harris Theorem, $f \& a_1 \in A - B$ but $f \& a_1 \not\in B$. Miller is correct in my assessment in criticizing Tichý's initial proposals for their failure of invariance of verisimilitude orderings with respect to logically equivalent representations of theories. Ideally if the concept of truth is invariant with respect to logically equivalent representations of theories, so ought the concept of verisimilitude, and an account of verisimilitude which does not preserve this intuition is to be regarded as defective.

A second line of criticism of Tichý's proposal was given by Karl Popper,\textsuperscript{17} who presented what he took to be counter-examples to the position. Consider once more Tichý's elementary weather language. In order to determine the distance of a sentence $a$ from the truth, i.e. $d_T(a)$, we first put a into disjunctive normal form, count the negation signs and divide by the number of conjunctive constituents of the disjunction, i.e., the number of the disjunction signs plus 1. If we consider $a = (p \& q \& r) \lor (\neg p \& \neg q \& \neg r)$, $b = a \lor (p \& q \& \neg r)$, $c = b \lor (p \& \neg q \& \neg r)$ and $d = c \lor (\neg p \& q \& r)$, then we obtain $d_T(a) = 1.5$, $d_T(b) = 1.3$, $d_T(c) = 1.25$ and $d_T(d) = 1.20$. The distances from the truth of these statements declines with declining logical strength, although according to Popper's intuition, they ought to increase.

Tichý's account runs into further difficulties. The sentence $p$ is true. Hence its distance from the truth is 0. But let us suppose that $p \iff h$. Hence $h$ is true, and $d_T(h) = 0$. But $\neg h$ is a degenerate disjunction of a degenerate conjunction. Hence $d_T(\neg h) = 1 / 1 = 1$. Also $p \iff \neg p$, where $n$ is an even natural number and we may repeat this argument. It is no good arguing here that since $p \iff \neg p$, this claim
collapses, for we may simply replace \( \neg p \) by \( \emptyset \) and repeat the first argument. This, of course, is an example of the variance of Tichý's verisimilitude ordering with respect to the relation of logical equivalence.

Consider a statement \( e = (p \land \neg p) \lor (q \land \neg q) \) which has \( d(e) = 2/2 = 1 \). We would expect, given classical logical intuitions, that \( e \) would be of maximal distance from the truth. However, for a statement \( f = (\neg p \land \neg q \land \neg r) \lor (\neg p \land \neg q \land r) \), \( d(f) = 7/2 \). Hence a logical truth may be further from the truth than a contradiction, which is strongly counter-intuitive. Tichý no doubt will regard logical truths as uninformative having zero information. Nevertheless, they are true statements, and their distance from the truth must be zero. For \( g = p \lor \neg p \), \( d(g) = 1/2 \), so even the distance from the truth of various tautologies varies.

In a later paper Tichý responds to Miller's criticism of the failure of invariance of verisimilitude orderings by translation into logically equivalent languages. In the case of Tichý's weather language with sentences \( p, q, \) and \( r \) where \( m = p \leftrightarrow q \) and \( a = p \leftrightarrow r \), Tichý denies that \( p \land m \land a \) and \( p \land q \land r \) are equivalent theories. His counter-argument is as follows. Two statements are equivalent if they have the same affirmative force. The affirmative force of a statement is the range of a statement, i.e., the class of possible states of affairs in which the statement holds true. Two statements are equivalent just in case they have the same range. A 'possible state of affairs' is a function which maps atomic propositions to truth-values and the totality of such functions is known as the logical space of the language. Tichý claims that \( \langle p, q, r \rangle \) and \( \langle p, m, a \rangle \) are two distinct sets of propositions, so no function defined on the former set can be identical with one defined on the latter. Tichý's claim here is a *petitio principii* against Miller, since there is no demonstration of the claim that \( p, m, a \) is in "another language". Thus whilst Miller was quite right to point out in his rejoinder to Tichý, that Tichý's argument does not show that sentences in different languages may not have the same assertive power, a more direct rejoinder is in order. We need only present a formalized metalanguage for Tichý's weather language and reformulate Miller's argument in our metalanguage. Alternatively we may grant Tichý his point and simply construct another language in which \( \langle p, m, a \rangle \) does feature in our object-language. Either of these strategies will refute Tichý's later rejoinder to Miller, that Miller's argument involves a surreptitious shift from the object-language to the metalanguage of \( L_w \).

We turn now to a consideration of Tichý's more extensive account of verisimilitude. To do so, we shall first need to define a number of concepts. Consider a first-order language \( L^1 \), with a finite vocabulary \( V \). Then a wff of \( L^1 \), \( w_i \), is said to be of depth \( d \) if its longest string of nested quantifiers (this in turn being a sequence whose every member stands within the scope of each of its predecessors) is of depth \( d \). Formulas of level 0 are those atomic formulas constructible from nothing but the members of \( V \). Atomic formulas constructible from members of \( V \) and variables \( x_1, x_2, \ldots x_{n+1} \) none of which are of levels 0, 1, \ldots \( n \) are called formulas of level \( n+1 \). Let \( q(l) \) be the number of formulas of level \( l \). A conjunction of level \( l \) is any \( q(l) \) way conjunction whose ith conjunct is either the ith formula of level \( l \) in the lexicographic ordering of formulas of level \( l \) or its negation. The \( m = 2^{q(l)} \) conjunctions of level \( l \) are referred to as \( \mathcal{F}_1, \mathcal{F}_2, \ldots \mathcal{F}_m \) in lexicographic order (with the negation sign last in the alphabet).
Any formula of level $d$ is called a $d$-subtree of level $d$. A $d$-subtree of level 0 is called a $d$-tree. Such formulas are capable of being given representation in a tree-diagram. Any formula represented by a $d$-tree is called a $d$-constituent. A node of level 1 of a $d$-tree is any occurrence of $\phi_1$ in a $d$-tree.

Let $c$ be a $d$-constituent, then the worlds in which $c$ is true are known as $c$-worlds, and for any world $w^*$, there is a unique $d$-constituent $c$ such that $w^*$ is a $c$-world. If $F$ is an arbitrary consistent formula of depth $\leq d$, then there are consistent $d$-constituents $c_1 \lor c_2 \lor c_3 \lor \ldots \lor c_k$, which is the distributive normal form of depth $d$ of $F$. $F$ asserts in effect that the actual world is a $c_1$-world or a $c_2$-world ... or a $c_k$-world.

The measure $E((h, w))$ is the number of formulas of level 1 which appear unnegated in one of the nodes and negated in the other. Let $c$ and $c'$ be $d$-trees and $C$ and $C'$ be the respective classes of their nodes. A linkage relation $L$ between $c$ and $c'$ exists if: (1) $C$ is the domain and $C'$ the range of $L$; (2) if $\phi_1$ and $\phi_2$ are arbitrary members of $C$ (resp. $C'$) such that $\phi_1$ is directly subordinated to $\phi_2$ and $\phi_2$ is a member of $\phi_1$ (resp. $\phi_1$ is a member of $\phi_2$) and $\phi_2$ is directly subordinated to $\phi_1$; (3) there is no subrelation $S$ of $L$ which satisfies (1) and (2). The breadth of $L$ is the ratio of the number of actual divergences between linked nodes and the number of all possible such divergences. The distance $\delta(c, c')$ between two $d$-trees $c$ and $c'$ is the breadth of the narrowest linkage between $c$ and $c'$. If $T$ is a consistent theory of depth less than or equal to $d$, and $c_1 \lor c_2 \lor \ldots \lor c_k$ is the distributive normal form of depth $d$ of $T$, and $c^*$ is a true $d$-constituent, then the d-distance $\Delta_d(T)$ of $T$ from the truth is:

$$\delta(c^*, c_i).$$

Consider a propositional language $LA$ with three primitive symbols $h$, $r$, and $w$, such that $h$ is true in the actual world if it is hot, $r$ if it is raining, and $w$ if it is windy. Thus for $A = h \& r \& w$, $\Delta_d(A) = 0/3 = 0$. For $D = -h \& -r \& -w$, $\Delta_d(D) = 3/3 = 1$. The truth-likeness of a sentence $\phi$, $M(\phi)$ is $1 - \Delta_d(\phi)$ so that $M(A) = 1$ and $M(D) = 0$. Now $A \lor D$ however is $(-h \& r \& w) \lor (-h \& -r \& -w)$ and $\Delta_d(A \lor D) = 0.5$. On this account however $\Delta_d(A \lor -A) = 0.5$, whereas for a contingent statement $B = h \& r \& -w$, $\Delta_d(B) = 0.33$. This is strongly counter-intuitive. Further the measure $\Delta_d'$ violates what one would take to be an intuitive criterion of adequacy for a verisimilitude ordering: if $\vdash p \rightarrow q$, then $\Delta_d(p) \leq \Delta_d(q)$. This is to say that $q$ may be equal to but not closer to the truth than $p$. In the case of $\vdash A \rightarrow A \lor D$, $\Delta_d(A) = 0$ and $\Delta_d(A \lor D) = 0.5$ which is satisfactory. But for $\vdash A \lor D \rightarrow (A \lor D) \lor B$, $\Delta_d(A \lor D) = 0.5$, whilst $\Delta_d(A \lor D \lor B) = 0.44$ and $0.5 > 0.44$. This is also strongly counter-intuitive.

Tichy's response to Popper's alleged counter-examples to his position indicates that he would not find that the failure of verisimilitude orderings to be invariant with regard to the relations of either logical
equivalence or implication to be at all problematic. However, the example which he offers in reply to Popper does not at all challenge our intuitions. Tichý asks us to consider three statements, \( p \), 'Snow is white', \( q \), 'Grass is green', and \( r \), 'The Moon is made of green cheese'. Then from \( p \lor (q \lor r) \) we can infer \( p \lor q \). On my position \( \Delta_d(p \lor (q \lor r)) \leq \Delta_d(p \lor q) \). Now it may well be taken to be counter-intuitive to claim that \( \Delta_d(p \lor (q \lor r)) < \Delta_d(p \lor q) \), but the claim that \( \Delta_d(p \lor (q \lor r)) = \Delta_d(p \lor q) \) is not. In the propositional logic, if \( p \) and \( q \) are true, then the falsity of \( r \) makes no difference to the truth value of the whole disjunct, and if this is so, it is quite plausible to claim that under these conditions \( \Delta_d(p \lor (q \lor r)) = \Delta_d(p \lor q) \).

It is concluded that Tichý's position stands open to a number of counter-examples, which indicate the inadequacy of his position.

IV. Tuomela and Niiniluoto on Verisimilitude

The approach to the problem of explicating the notion of verisimilitude adopted by both Ikka Niiniluoto and Raimo Tuomela makes use of Hintikka's notion of constituents to define a quantitative distance between constituents, which is used in turn to define the notion of verisimilitude. These approaches will now be outlined.

Consider a first order language \( L^w \) with a finite vocabulary, but no individual constants. Each generalization \( g \) in \( L^w \) can be expressed as a finite disjunction of mutually exclusive constituents at a depth \( d \). \( L^w \) is a monadic language with logically independent primitive predicates \( O_1, O_2, \ldots, O_k \). Constituents \( C_i \) of \( L^w \) are expressions of the following form:

\[
(4) \quad C_i = (\pm)(\exists x) C_{t1}(x) \land \ldots \land (\exists x) C_{tk}(x),
\]

where '\((4)\)' is to be replaced by the negation sign '\(-\)' or else by no sign at all, and where \( K = 2^k \). The constituents are sentences which claim that certain \( Q \)-predicates are empty, whilst others are non-empty. The \( Ct \)-predicates of \( (4) \) are conjunctions of the form:

\[
(5) \quad C_{tj}(x) = (\pm)O_{tj}(x) \land \ldots \land (\pm)O_{tk}(x).
\]

Assume that the language \( L^w \) is interpreted in a domain \( D^w \), which represents the actual world. Then only one of the constituents \( C_i \) is true in \( D^w \); let this constituent be \( C_a \).

Niiniluoto defines the Clifford-measure or the distance between monadic constituents \( C_i \) and \( C_j \) as follows:

\[
(6) \quad d_C(C_i, C_j) = \frac{|CT_i \Delta CT_j|}{K}
\]

where '\( \Delta \)' denotes the relation of symmetric difference, '\( K \)' is the total number of \( Q \)-predicates in \( L^w \), '\( CT_i \)' denotes the set of \( C_i \)'s \( Q \)-predicates, and '\( CT_j \)' denotes the set of \( C_j \)'s \( Q \)-predicates. Given the distance \( d_C(C_i, C_j) \) between two constituents \( C_i \) and \( C_j \), the distance of a
generalization \( g \) from \( C_* \) is \( d_c(g, C_*) \) and the truthlikeness or verisimil­itude \( M(g, C_*) \) of \( g \) is \( 1 - d_c(g, C_*) \).

If we take account of the errors made by \( C_i \) with respect to \( C_j \) we have:

\[
M(g, C_*) = 1 - \left( \gamma m^*(g, C_*) + (1 - \gamma) m^*(g, C_*) \right)
\]

where \( '1 - m^*(g, C_*)' \) denotes the measure relative to \( C_* \) of the degree of truth in \( g \), \( '1 - m^*(g, C_*)' \) denotes the measure of the degree of information about the truth in \( g \), and \( \gamma \) is a weight parameter for these two factors, such that \( 0 < \gamma < 1 \).

To outline Tuomela's position, we shall still operate with the language \( LTN \). Consider two sentences \( T_1 = C_1 \lor C_2 \lor ... \lor C_m \) and \( T_2 = C_1 \lor C_2 \lor ... \lor D_n \). Then \( d_c(T_1, T_2) \) is defined as follows:

\[
d_c(T_1, T_2) = \frac{1}{K} \sum_{i,j=1}^{m,n} \alpha_{ij} wcard \left( (C_1 \setminus C_2) \right) + \frac{1}{r_s} \sum_{i,j=1}^{r,s} d_c(C_i, C_j) + \frac{1}{t,v} \sum_{i,j=1}^{t,v} d_c(C_i, C_j) + \frac{1}{\beta} wcard \left( (C_1 \cup C_2) \right) + (1 - \beta) wcard \left( (C_1' \setminus C_2') \right.
\]

In (8) \( 'wcard' \) means 'weighted cardinality', these being weights associated with each \( C_t \) predicate, representing the importance of that \( C_t \) predicate for theory distance, and summing up to \( K \); \( 'Y' \) is the name of a parameter \( 0 < Y < 1 \) reflecting the relative importance of the whole error-factor in the characterization of theory-distance. The parameters sum up to one, i.e., \( Y + \alpha_1 + \alpha_2 + \alpha_3 = 1 \), and \( 0 < \alpha < 1 \). In (8) \( m \) is the number of \( C_t \)-predicates in \( C_1 \setminus C_2 \), \( n \) the number in \( C_2 \), \( r \) and \( s \) those in \( C_1 \setminus C_2 \) and \( C_1 \) respectively; \( k, l, t, \) and \( v \) give the cardinalities of \( C_1 \setminus C_2, C_2, C_2 \setminus C_1, \) and \( C_1 \) respectively. The distance \( d_c(C_1, C_2) = 1/k \) consists of a weighted number of predicates \( 0, i = ...k \) which have a different sign in \( C_1 \) and \( C_2 \), assigned on the basis of the importance of \( 0_i \) for the comparison of distance. The concept of verisimilitude is defined as follows:

\[
V(T^{(d)}) = 1 - d_c(T^{(d)}, C_t^{(d)}),
\]

where \( C_t^{(d)} \) is the constituent of \( LTN \) at depth \( d \) representing the truth.

The proposals of Niiniluoto and Tuomela are based upon the intuition that the truthlikeness of some theory can be determined by distance comparisons within the theory. However, the language-relativity (which is fully conceded by Niiniluoto) renders these accounts less than adequate. Let \( L_1 \) and \( L_2 \) be two essentially different first-order languages, and let \( C_1^2 \) and \( C_2^2 \) be the true constituents of \( L_1 \) and \( L_2 \).
respectively. If \( g \) and \( g' \) are generalizations in \( L_1 \) and \( L_2 \) respectively, then it is possible that \( M(g, C_1) < M(g', C_1) \), even if \( M(g, C_2) > M(g', C_2) \). If \( g \) and \( g' \) are not expressed in \( L_2 \) and \( L_1 \) respectively then it is impossible for distance comparisons to be made at all in any non-problematic fashion. The suggestion that we consider a common extension of \( L_1 \) and \( L_2 \), namely \( L_3 \), with vocabulary \( \lambda_3 = \lambda_1 \cup \lambda_2 \) also meets difficulties, if \( T_1 \) and \( T_2 \) are conflicting and mutually inconsistent, for either we would outrightly fail to obtain a true constituent \( C_0 \), or else the overall verisimilitude of \( T_3 \) would be quite low due to the large numbers of inconsistent sentences in it. Any satisfactory account of verisimilitude should avoid this problem; Tuomela's and Niiniluoto's accounts do not.

To conclude this section, I will advance some counter-examples to the accounts of Tuomela and Niiniluoto. Let us consider the situation where \( \delta = 1 \) and \( \varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 0 \) in Tuomela's model. Then his measure of theory distance is equivalent to Niiniluoto's. Consider a simple language \( L_S \), which has only two predicates \( \{O_1, O_2\} \) in its vocabulary and only one variable, \( x \). Then the Q-predicates or Ct-predicates will be:

\[
\begin{align*}
Ct_1(x) &= (\not\exists) O_1(x) \\
Ct_2(x) &= (\exists) O_1(x) & (\not\exists) O_2(x).
\end{align*}
\]

The constituents of \( L_S \) are expressions of the form:

\[
(11) \quad C_i = (\exists) (\not\exists) O_1(x) & (\exists) (\not\exists) O_2(x).
\]

Consider \( C_4 = (\exists) O_1(x) & (\not\exists) O_2(x) \), which is taken to be true. Then \( d(C_4, C_0) = |C_4 \triangle C_0| / K \). Now \( C_T \) is the set of \( \text{Ct}_x \)-predicates of \( C_4 \) which is \( \{O_1(x), O_1(x) & O_2(x)\} \). Now \( C_T \cap \not\exists C_T = \{O_1(x), O_1(x) & O_2(x)\} \) and \( \text{Ct}_x \cap \not\exists \text{Ct}_x = \{\emptyset\} \), i.e., \( C_T \triangle \text{Ct}_x = \{\emptyset\} \) and \( |C_T \triangle \text{Ct}_x| = 0 \). Hence \( d(C_4, C_0) = 0 \). This is as we would expect. But for a counter-example we need only consider a \( C_i \) such that \( C_T \) has the same elements as \( C_T \), e.g., \( C_i = \neg(\exists) O_1(x) \bigoplus (\not\exists) O_1(x) & O_2(x) \). \( C_i \) and \( C_4 \) differ only by virtue of the external negation signs, although their Q-predicates are identical. Hence \( |C_T \triangle \text{Ct}_x| = 0 \), and \( d(C_4, C_0) = 0 \) and \( M(C_4, C_0) = 1 \). This is strongly counter-intuitive and a defect in the proposal.

V. Verisimilitude and Shared Tests

Clifton Perry\(^{28}\) has argued that the Miller-Tichý-Harris Theorem may be questioned if the claim that the comparability of truth and falsity contents with respect to empirical content is in turn questioned. If the comparison between two theories with respect to a given variable is to be meaningful, then both theories must be measured in terms of that variable. He continues:

It may therefore be suggested that although competing theories may be commensurate, it does not follow that, taken as a whole, the truth and falsity contents of a theory with greater empirical content are comparable to the truth and falsity contents of a theory with less empirical content. Insofar as the theory with greater empirical content has po-
tentially more true or false logical consequences which are not also consequences of the theory with less empirical content, reference to differences in empirical content between theories in the ascription of verisimilitude would fail to compare adequately only those consequences which were relevant to both theories. Reference to the differences in empirical content in the ascription of verisimilitude to two theories shall be referred to as ascriptions of 'absolute' verisimilitude. Appraisals of verisimilitude which are restricted to those test situations which are applicable to both theories in question and consequently obviate reference to the different degrees of empirical content in the comparison of truth and falsity contents shall be referred to as ascriptions of 'relative' verisimilitude. It may be said, therefore, that theory B possesses more relative verisimilitude than theory A if and only if B's falsity content is a sub-set of A's falsity content, A's truth content consequently being a sub-set of B's truth content. The above formulation differs from Popper's in that reference to the difference in empirical content in the comparison of truth and falsity contents is omitted.31

Following the conventions adopted earlier in the paper, we shall represent Perry's definition as follows:

\[(PDRV) \ A >_R \ B \iff (A \subseteq B_T) \land (B_T \subseteq A_T).\]

This weakened view of verisimilitude is also subject to a variant of the Miller-Tichý-Harris Theorem:

\[(MTHT)^* \ A >_R \ B \rightarrow A \subseteq T.\]

**Lemma 1:** \(A \subseteq B_F \rightarrow A \subseteq T.

**Proof of Lemma 1:** Consider \(A \subseteq B_F\) and suppose for reductio ad absurdum that \(A \notin T\). Let \(f \in A\) and \(f \in F\). Let \(a \in A_T-B_T\) so that \(a \in A, a \in T\) and \(a \notin B\). Since \(a \in A\) and \(f \in A\), then \(a \land f \in A\). Also \(f \in F\), so \(a \land f \in A_F\). Hence \(a \land f \in A_F\). Now since \(a \notin B\), then \(a \land f \notin B\). Hence \(a \land f \notin B_F\). But \(a \land f \in A_F\) and \(a \land f \notin B_F\) contradict \(A \subseteq B_F\). Hence \(~(A \notin T)\), i.e., \(A \subseteq T\).

**Lemma 2:** \((B_T \subseteq A_T) \rightarrow A \subseteq T.

**Proof of Lemma 2:** Consider \(B_T \subseteq A_T\) and suppose for reductio ad absurdum that \(A \notin T\). Let \(f \in A\) and \(f \in F\). Consider \(b \in B_T-A_F\), \(b \in B, b \in F\) and \(b \notin A\). Since \(b \in B\), \(~f \lor b \in B\). Since \(f \in F\), \(~f \lor b \in T\). But \(~f \lor b \notin A_T\), as \(b \notin A\), but \(f \in A\), so \(~f \lor b \in A\). Now \(~f \lor b \in B_T\) and \(~f \lor b \notin A_T\) contradict \(B_T \subseteq A_T\). Hence \(~(A \notin T)\). Therefore \(A \subseteq T\).

\((MTHT)^*\) follows immediately from Lemmas 1 and 2. Perry's definition does not escape a variant of the Miller-Tichý-Harris Theorem. It too is inadequate.
VI. Verisimilitude and Short Theorems

Peter Mott\textsuperscript{32} has proposed that since the Miller-Tichy-Harris Theorem relies upon sentences that are not 'genuine' theorems in the sense of being characteristic of the theory, then this limitation theorem may be avoided by use of the notion of 'short theorems'. The notion is based upon the idea of organicity: an axiom is organic with respect to a first-order system $X$ if it contains no segment which is in turn a theorem of $X$, or becomes a theorem of $X$ as soon as open variables are bound by any type of quantifier so that a wff of $X$ is produced.\textsuperscript{33} Mott, however, offers the following definition of shortness: "a disjunction of prime formulas, $A$, is short in $X$ provided that there is no disjunction $A$ in $X$ obtained from $A$ by replacing throughout, $P$ by $\neg P$ or $-P$ by $P.$\textsuperscript{34}

There are in the first-order system $X$ denumerably many sentence letters $P_1, \ldots P_n, \ldots$ and the usual logical connectives. For any sentence letter $P_i$, $P_i$ and $\neg P_i$ are prime formulas in $P_i$. The verisimilitude of $Y$ is greater than $X$ just when it preserves the short truths of $X$ and adds new short truths, i.e, $X_{ST} \subseteq Y$ iff $(X_{ST} \subseteq Y) \land \neg (Y_{ST} \subseteq X)$.

Mott recognizes that his account of verisimilitude only preserves transitivity orderings for a set of theories $S$ which is a chain with respect to the relation of $\subseteq$. On this failure for sets of theories $S^\sharp$ which are not chains, but may well be given an intuitively correct verisimilitude ordering he states:

Intuition would have it that verisimilitude is transitive—but then intuition might be conditioned by nothing more substantial than the phrase 'nearer the truth'. On reflection there seems to be no \textit{prima facie} reason why verisimilitude should not be more like '... is indistinguishable from ...' than '... is identical to ...'. Perhaps as theories evolve they gradually drift apart, so that though each improved upon its predecessor, the last is not comparable with the first. Perhaps an early cosmology might contain all mixed up together religious and cosmological truths. Later theories may preserve the secular while forgetting the divine insights. The very subject matter of the theories may gradually drift finally rendering the first and the last about almost entirely different things, though each handles the problems, or most of the problems of its predecessor. In sum, it may be that there are decisive arguments to show that verisimilitude is transitive, but if so they are not known to the writer.\textsuperscript{35}

Such an argument for the plausibility of taking the relation of verisimilitude to be transitive may be either based on the transitivity of truth assessments, or else upon considering the transitivity of Popper's original definition of verisimilitude. Insofar as this account captures many of our intuitive beliefs about verisimilitude (notwithstanding of course the Miller-Tichy-Harris Theorem), the transitivity of verisimilitude orderings is quite marked. If $(A \succ B) \land (B \succ C)$, then it is provable that $A \succ D$. If we expand out $(A \succ B) \land (B \succ C)$, we obtain:

$$\begin{align*}
(12) & \quad \{(B_T \subseteq A_T) \land (A_F \subseteq B_F) \lor (B_T \subseteq A_T) \land (A_F \subseteq B_F)\} \\
& \land \{(C_T \subseteq B_T) \land (B_F \subseteq C_F) \lor (C_T \subseteq B_T) \land (B_F \subseteq C_F)\}
\end{align*}$$
WHAT IS WRONG WITH VERISIMILITUDE?

which by use of distribution rules gives:

\[(13) \quad \{(B_T \subseteq A_T) \& (A_F \subseteq B_F) \& (C_T \subseteq B_T) \& (B_F \subseteq C_F)\}
\v{(B_T \subseteq A_T) \& (A_F \subseteq B_F) \& (C_T \subseteq B_T) \& (B_F \subseteq C_F)\}
\v{(B_T \subseteq A_T) \& (A_F \subseteq B_F) \& (C_T \subseteq B_T) \& (B_F \subseteq C_F)\}
\v{(B_T \subseteq A_T) \& (A_F \subseteq B_F) \& (C_T \subseteq B_T) \& (B_F \subseteq C_F)\}
\v{(B_T \subseteq A_T) \& (A_F \subseteq B_F) \& (C_T \subseteq B_T) \& (B_F \subseteq C_F)\}\]

Now (13) can be readily shown to imply:

\[(14) \quad (C_T \subseteq A_T) \& (A_F \subseteq C_F) \v (C_T \subseteq A_T) \& (A_F \subseteq C_F)\]

i.e., \(A \succ v C\). Hence Popper's original verisimilitude relationship is transitive. We should note that such a relationship concerns the relationship between the sets of true and false consequences of theories. Mott's remarks about cosmological theories "drifting apart" will either mean that many of the religious claims of the early cosmology will be taken to be false (e.g., the world was created less than 10,000 years ago) or else the two cosmologies cannot be compared. Mott requires that subset relationships still hold between the respective sets of short theorems, but if the theories are incomparable, then this relationship will not hold, i.e., \(Y\) will not preserve some of the 'short religious truths of \(X\)', since such truths will not be theorems of \(Y\). In conclusion, Mott's rejection of the transitivity of verisimilitude is not supported by a satisfactory argument, and this is all the worse for his account of verisimilitude.

Mott's account of verisimilitude also fails to avoid one horn of the Miller-Tychy-Harris Theorem, this being the claim that if \((B_T \subseteq A_T) \& (A_T \subseteq B_F)\) then \(A \subseteq T\). The proof of this can be readily given by taking \(a \in A_T - B_T\) and \(f \in A\) and \(f \in F\). The problem is that \(a \& f \in A_T\) but \(a \& f \notin B_F\) which contradicts \(A_T \subseteq B_F\). We can readily let \(a\) be a short theorem. As long as \(f \in A\) and \(f \in F\) this result will follow. Mott's original motivation for his account was based upon the view that if \(b \in X\), then \(a \lor b\) is logical baggage. This is of course a product of the rule of inference \(a \rightarrow a \lor b\). But Mott's claim ignores one part of the Miller-Tychy-Harris Theorem as it has been stated and proved here, where reliance is placed upon \(a \& b \rightarrow a\). Since \(f\) is not a short truth of \(A\) as \(f \in F\), Mott's program is beside the point. It fails to resolve the problem posed by the Miller-Tychy-Harris Theorem.

VII. Bunge's Theory of Partial Truth

Mario Bunge has argued that the notion of 'degree of truth' is extensively employed in applied mathematics and factual science.36 David Miller37 has established the untenability of Bunge's earlier account.38 After briefly reviewing Miller's criticism of Bunge's earlier position, his more recent theory of partial truth will be reviewed, criticized, and rejected.

Bunge writes \(V(p) = r\) for the degree of truth of the proposition \(p\); if \(p\) is true, the \(V(p) = 1\), if \(p\) is false, \(V(p) = -1\), and if \(p\) is either meaningless or undecidable, \(V(p) = 0\).39 \(V\) is, we shall suppose, a real valued function satisfying the following axioms

\[(A1) \quad -1 \leq V(p) \leq 1.\]
On this account if \( t \) is a tautology, \( v(t) = 1 \). Consider now \( V(p \land t) \), which should equal \( V(p) \). If \( V(p) \neq -1 \), then

\[
(B1) \quad V(p \land t) = V(p) = V^3(p) + 1
\]

\[
-----------------
\]

\[
V^2(p) + 1
\]

\( (B1) \) is satisfied only if \( V(p) = 1 \). Hence if \( V(p) \neq -1 \), then \( V(p) = 1 \). Hence Bunge's theory of partial truth only permits at most two degrees of truth.

Bunge also sets out some presystematic ideas, which play the role of desiderata to be fulfilled by the theory of partial truth. Among these are:

1. If \( p \implies q \) then \( V(p \land \neg q) = V(p \lor q) = V(p) \).
2. \( V(p \land \neg p) = -1, V(p \lor \neg p) = 1 \).

Consider the following contradiction, \( p \implies \neg p \), which should receive value \(-1\). Then by (b) above \( V(p \land \neg p) = V(p \lor \neg p) = V(p) \). But \( V(p) \) may well be in the range \( 0 < V(p) < 1 \) or \( -1 < V(p) < 0 \). Consider also a conditional, \( p \implies p \) such that \( V(p \implies p) = 1 \). Then by (b) above, \( V(p \implies p) = V(p \land p) = V(p \lor p) = V(p) \). Now \( V(p \implies p) \) is surely 1, but \( V(p) \) need not be 1 at all. Compare these results with the results which can be obtained from Bunge's Theorem 3, which asserts that:

\[
(T3) \quad V(p \implies q) = \begin{cases} 1 & \text{if } V(p) = V(q) \\ 0 & \text{otherwise} \end{cases}
\]

Here \( V(p \implies \neg p) = 0 \), which is the claim that a genuine contradiction is meaningless or undecidable. This is also strongly counter-intuitive.

We shall now review Bunge's more recent theory of partial truth. Consider the structure \( B = \langle S, S_b, [S], \boxminus, \boxplus, \bowtie, \neg, V \rangle \), where \( S \) is a non-empty set, \( S_b \) a subset of \( S \), \([S]\) the quotient of \( S \) by the relation \( \implies \) of logical equivalence, \( \boxminus \) and \( \boxplus \) distinguished elements of \( [S] \), \( \bowtie \) and \( \neg \) binary operations on \( [S] \), \( \neg \) an unary operation on \( [S] \) and \( V \) a value function on \( S_b \). \( B \) is a metric Boolean algebra of statements if and only if \( B \) is a Boolean algebra with null element \( \boxminus \), universal element \( \boxplus \), and \( V \) is a real valued function on \( S_b \subset S \), such that for any elements \( p \) and \( q \) of \( S_b \):

\[
(15) \quad (a) \quad V(p \land q) + V(p \lor q) = V(p) + V(q)
\]

\[
(b) \quad V(p) = 0 \text{ for all } p \in \boxminus
\]
WHAT IS WRONG WITH VERISIMILITUDE? 525

Further, for any p, q, r in S:

\[(16) [q] = [p] \text{ iff } \neg p \leftrightarrow q \text{ is a tautology}\]
\[ [q] \cup [r] = [p] \text{ iff } \neg p \leftrightarrow q \vee r \text{ is a tautology}\]
\[ [q] \cap [r] = [p] \text{ iff } \neg p \leftrightarrow q \land r \text{ is a tautology}\]

and:

\[(17) \square = \{(p \in S) \& (q \in S) \mid p \leftrightarrow q \& \neg q\}\]
\[\square = \{(p \in S) \& (q \in S) \mid p \leftrightarrow q \vee \neg q\}.

The function \( \delta : S_0 \times S_0 \rightarrow [0,1] \) assigns each pair of propositions \( p, q \in S_0 \) a real number between 0 and 1, such that \( \delta (p,q) = |V(p) - V(q)| \); this is to be known as the horizontal distance. The distance function \( \delta \) satisfies the following axioms:

\[(18) (1) \quad \delta (p,q) = \delta (q,p)\]
\[(2) \quad \delta (p,q) + \delta (q,r) \geq \delta (p,r)\]
\[(3) \quad \delta (p,q) = 0 \text{ iff } V(p) = V(q), \text{ for any } p, q, \text{ and } r \in S_0\]

\( \delta \) defines a topology in the Space \( S_0 \). An open \( \varepsilon \)-neighborhood of \( p \in S_0 \) is the set:

\[(19) \mathcal{U}_\varepsilon (p) = \{ q \in S_0 \mid |V(p) - V(q)| < \varepsilon \} \text{ for } 0 \leq \varepsilon \leq 1.

All statements that agree with \( p \) to within the tolerance error \( \varepsilon \) are in \( \mathcal{U}_\varepsilon (p) \). As the distance between equivalent statements is nil, equivalent statements agree with one another.

A second distance function \( \delta_1 : S_0 \times S_0 \rightarrow [0,1] \), which to each pair of propositions \( p, q \in S_0 \) is assigned a real number between 0 and 1 such that \( \delta_1 (p,q) = |V(p \vee q) - V(p \& q)| \) is known as the vertical distance. A second topology in \( S_0 \) is defined such that an open \( \varepsilon \)-neighborhood of \( p \in S_0 \) is now:

\[(20) \mathcal{V}_\varepsilon (p) = \{ q \in S_0 \mid \delta_1 (p,q) < \varepsilon \} \text{ for } 0 \leq \varepsilon \leq 1.

The two truth spaces \( <S_0, \delta> \) and \( <S_0, \delta_1> \) are separable, that is, for any two propositions \( p, q \in S_0 \), there are open sets \( G \) and \( H \) in \( S_0 \) such that \( p \) is in \( G \) and \( q \) is in \( H \) and \( G \) and \( H \) are disjoint. Further:

\[(T1) \text{ If } p, q \in S_0, \text{ then } \delta_1 (p,q) \geq \delta (p,q).

A number of consequences may be derived from these assumptions: (1) for any \( p, q \in S_0 \), \( V(p \vee q) \geq V(p \& q) \); (2) for any \( p \in S_0 \), \( V(p) = 1 - V(p) \); (3) for any \( p, q \in S_0 \), if \( V(p \leftrightarrow q) = 1 \), then \( V(p) = V(q) \). Result (3) is not, Bunge claims, restricted to formally true biconditionals.
The following theorem is also of importance:

\[ (T2) \text{ For any } p, q \in S_0, \text{ if } V(p \rightarrow q) = 1, \text{ then:} \]

1. \( V(p \& q) = V(p) \)
2. \( V(p \lor q) = V(q) \).

Let \( p, q \in S_0 \) with \( V(p) \neq 0 \). Then the truth-value of \( q \) relative to \( p \) is defined as:

\[ q^p = \frac{V(p \& q)}{V(p)} \]

It is said that \( p \) is alethically independent of \( q \) if and only if \( V(q/p) = V(p) \), and is alethically dependent upon \( q \) otherwise. Alethic dependence subsumes logical dependence. Note that since \( V(p) \neq 0 \), the truth value of \( q \) relative to a contradiction cannot be made; this is inevitable given the mathematical form of Bunge's definition.

A further theorem is of importance:

\[ (T3) \text{ If } p \text{ and } q \text{ are alethically independent statements in } S_0 \text{ then:} \]

1. \( V(p \& q) = V(p) \cdot V(q) \)
2. \( V(p \lor q) = V(p) + V(q) - V(p) \cdot V(q) \)

With this logic machinery Bunge proceeds to define the degree of truth of a scientific theory: the degree of truth of a scientific theory equals the product of the truth-values of its initial assumptions, provided that these are independent. If \( T \) is a scientific theory with \( n \) independent assumptions \( A_i \), then (1) the degree of truth of the axiom base equals the product of the partial degrees of truth:

\[ V(\bigwedge_{i=1}^{n} A_i) = \prod_{i=1}^{n} V(A_i) \]

(2) the degree of truth of an assumption conjoined with any of its logical consequences equals the former:

\[ (23) \text{ If } A_i \vdash t, \text{ then } V(A_i \& t) = V(A_i). \]

On this basis, the concept of verisimilitude may be explicated as follows: theory \( T_2 \) has a greater verisimilitude than the theory \( T_1 \) if and only if the degree of truth of \( T_2 \) is greater than the degree of truth of \( T_1 \).

I will now outline some logical defects in Bunge's position. First, however, let us note that Bunge claims that \( V(p) = 0 \) for all \( p \in \Box \), such that \( p \) is a contradiction, and \( V(p) = 1 \) for all \( p \in \top \), such that \( p \) is a tautology. However, truth-values \( V(p) \) are in the real interval \([0,1]\). This leaves Bunge with the option of either stating that both truths and tautologies have value 1 and that both falsehoods and contradictions have value 0, or else to arbitrarily assign complete truth and complete falsity to some values in the interval \([0,1]\). In either case, in what fol-
Bunge claims that if \( V(p \leftrightarrow q) = 1 \), then \( V(p) \equiv V(q) \). Consider \( V(p \leftrightarrow p) = 1 \), and \( V(p) = 1/2 \) as \( p \) is a contingent partial truth. But also \( V(p \leftrightarrow (p \leftrightarrow p)) = 1 \). Now \( V(p) = 1/2 \) from assumption and \( V(p \leftrightarrow p) \) is \( 1/2 \neq 1 \). It is difficult to see how this result could be avoided since Bunge claims that his account is not merely restricted to formally true biconditionals.

A second defect is with T. Consider the claim:

\[(24) \text{For any } p, q \in S_0 \text{ if } V(p \to q) = 1, \text{ then } V(p \& q) = V(p) .\]

For a counter-example consider the classical tautology, \( p \& \neg p \to q \). Then \( V(p \& \neg p \to q) = 1 \). Consider \( V(p) = 1/2 \), so that \( V(\neg p) = 1/2 \). Now \( p \) and \( \neg p \) are alethically independent statements if the logic of the system containing them is consistent. Hence \( V(p \& \neg p) = V(p) \& V(\neg p) = 1/2 \). \( 1/2 = 1/4 \). But, since \( p \& \neg p \) is a classical contradiction, \( V(\neg p) = 0 \). Hence \( 0 = 1/4 \), which is absurd. Therefore Bunge's principle is absurd.

A parallel difficulty arises with the claim that for any \( p, q \in S_0 \) if \( V(p \to q) = 1 \), then \( V(p \text{ v } q) = V(q) \). Consider the classical tautology \( q \to p \text{ v } \neg p \), such that \( V(q \to \neg p) = 1 \). Then \( V(q \to (p \text{ v } \neg p)) = V(p \text{ v } \neg p) \). We consider now \( V(p \text{ v } \neg p) \) where \( p \) and \( \neg p \) are alethically independent and \( V(p) = 1/2 \) and \( V(\neg p) = 1/2 \). Then: \( V(p \text{ v } \neg p) = V(p) + V(\neg p) - V(p) \). \( V(\neg p) = 1/2 + 1/2 - 1/2 \). \( 1/2 = 1 - 1/4 = 3/4 \neq 1 \). Once more we obtain a contradiction.

According to Bunge, if \( p \) and \( q \) are alethically independent statements in \( S_0 \), then:

\[(25) V(p \& q) = V(p) \cdot V(q) .\]

Substitute \( \neg p \) for \( q \) in (25) to obtain:

\[(26) V(p \& \neg p) = V(p) \cdot V(q) .\]

Since \( V(p) = 1 - V(p) \), we obtain:

\[(27) V(p \& \neg p) = V(p) \cdot (1 - V(p)) = V(p) - V^2(p) .\]

Now \( V(p \& \neg p) \) we assume to be 0. However, consider \( V(p) = 1/2 \). Then by (27) \( V(p \& \neg p) = 1/2 - (1/2)^2 = 1/4 \neq 0 \). Finally, according to Bunge, if \( p \) and \( q \) are alethically independent statements in \( S_0 \), then

\[(28) V(p \text{ v } q) = V(p) + V(q) - V(p) \cdot V(q) .\]

Consider a tautology \( V(p \text{ v } p) = 1 \). Suppose \( V(p) = 1/2 \). Then, as \( V(\neg p) = 1 - V(p) \), \( V(\neg p) = 12 \). By (28), \( V(p \text{ v } \neg p) = V(p) + V(\neg p) = 1 - 1/4 = 3/4 \).

These results indicate that Bunge's theory of partial truth is badly inconsistent. Therefore the position is severely defective.
VIII. Krajewski and Rosenkrantz on Verisimilitude

Krajewski distinguishes between relative and absolute truths. Qualitative facts include event-facts (e.g., World War II occurred), facts about states of affairs (e.g., people die without food), and relational facts (e.g., a left hand is on the left side of a human body). Qualitative factual statements are not relatively true or false, but are either absolutely true or absolutely false. Krajewski also recognizes other absolutely true statement types, including existential statements (e.g., there is more than one object in the world) and qualitative law statements (e.g., all metals are good electrical conductors). By contrast many other statements such as the statement of the gas laws and cosmological statements about recession velocities are only ‘approximate’, holding only within some margin of error. Krajewski attempts an explication of the notion of approximate or relative truth by means of the notion of relative errors. The mechanics of this account will now be examined.

Let \( p \) be a quantitative fact-statement, and the truth-content of \( p \) be written \( \text{Tr}(p) \), and let the relative error made by \( p \) be \( E(p) \). Then

\[
(\text{K1}) \text{Tr}(p) = 1 - E(p).
\]

\( E(p) \) is never known exactly, but the possible maximal error on the basis of a given measurement usually is. If \( a_1 \) is the result of such measurement and \( \Delta a \) is a maximal absolute error, then \( E(p) = \frac{\Delta a}{a_1} \).

The case of defining \( \text{Tr}(L) \) for a quantitative law \( L \) is more complex. The degree of inadequateness (DI) of a law \( L \) with respect to a parameter \( B \) contained in it is equal to the Supremum of relative errors made by using \( L \) to predict values \( b_i \) of \( B \).

Let \( '\text{DI}_B(L)' \) designate the DI of \( L \) with respect to \( B \)

and \( '\text{E}_B(L)' \) designate the relative error made in the prediction \( b_i \) of \( B \)

then:

\[
(\text{K2}) \text{DI}_B(L) = \text{Sup}_{i} [\text{E}_B(L)]
\]

and the truth-content of \( L \) with respect to a parameter \( A \), \( \text{Tr}_A(L) \), is:

\[
(\text{K3}) \text{Tr}_A(L) = 1 - \text{DI}_A(L)
\]

Let \( 'j' \) designate any of the parameters contained in \( L \), then:

\[
(\text{K4}) \text{Tr}_C(L) = \text{Min}_{j} [\text{Tr}_j(L)] = 1 - \text{Max}_{j} [\text{Sup}_{i} [\text{E}_j(L)]]
\]

Finally, if a theory is held to be a conjunction of law-statements, then the truth-content of a theory \( T \) may be defined as the minimum of the truth-content of all laws \( L_i \) contained in it:
Krajewski takes it for granted in his account that all the inaccuracies of scientific theories are due to experimental errors. This project has some rather unacceptable counter-intuitive consequences if his proposals are taken to present a general theory of verisimilitude. These will now be detailed.

Whilst Krajewski only defines the concept of truth-content of a single factual-statement p and law L, a natural extension of this idea is to suppose that if some theory $T_1 = \{p_1, p_2, \ldots p_n\}$ where 'p1', 'p2', 'pn' denote the logical consequences of $T_1$, then the truth-content of $T_1$ is as follows:

$$(K5)^* \text{TrC}(T_1) = \text{Min}_{i=1}^{n} \text{TrC}(L_i).$$

Now if $T_2 = \{q_1, q_2, \ldots q_r\}$, then $\text{TrC}(T_2) = \sum_{i=1}^{r} (1 - E(q_i))$.

Suppose that $r > n$. Then it is possible for a theory $T_1$, which had only absolutely true consequences, to nevertheless be of less truth-content than a theory $T_2$. This would mean that the truth-content of $T_2$ could be much higher than that of $T_1$, even if each of $q_1, q_2, \ldots q_r$ had quite high relative errors. $T_1$ may say 'more' than $T_2$, but what it says is quite inaccurate. Intuitively, however, $T_2$ seems closer to the truth than $T_1$.

To avoid the rejoinder that the previous definition puts words into Krajewski’s mouth, consider now (K5). A theory $T_1$ is taken to be the conjunction of laws contained in it, i.e., $T_1 = L_1 \& L_2 \& \ldots \& L_n$. The truth-content of $T_1$ is defined as the minimum of the truth-content of all the laws of $T_1$. If we interpret this statement to mean that the truth-content of $T_1$ is only as good as its weakest law, then a theory $T_1$ with only one completely false law is no better than a totally inadequate theory with all its laws false. Alternatively, we may interpret Krajewski’s requirement to be this:

$$(K7)^* \text{TrC}(T_1) = \text{Min} \{ (1 - \text{Max}_j [D_j(L_1)]) + (1 - \text{Max}_j [D_j(L_2)]) + \ldots + (1 - \text{Max}_j [D_j(L_n)]) \}$$

Suppose $\text{TrC}(T_2)$ is such that $n$ is quite small relative to $r$, but that its laws are highly accurate. Then $\text{TrC}(T_1) > \text{TrC}(T_2)$ even if the laws of $T_1$ are extremely inaccurate. For example, $\text{TrC}(T_2) = \text{Min} \{(1 - 0.1) + (1 - 0.1) + \ldots + (1 - \text{Max}_j [D_j(L_1)]) = 0.9\}$ = 9). On the other hand, $\text{TrC}(T_1) = \text{Min} \{(1 - 0.9) + (1 - 0.9) + \ldots + (1 - \text{Max}_j [D_j(M_{100})]) = 100\}$. But 100 > 9.

R.D. Rosenkrantz has offered a probabilistic analysis of the notion of verisimilitude. We shall say that the support which an observation $E$ accords a hypothesis $H$ is measured by the likelihood $P(E|H)$, i.e., by the probability that $H$ accords $E$. For $K = K_1 \vee \ldots \vee K_a$ and $P(K|E) = \sum P(K_i|E)$, we have by Bayes' Theorem:
(29) \[ P(K|E) = \frac{P(K)}{P(E)} \left\{ \sum_{i=1}^{n} \frac{P(E|K_i)}{P(K)} \right\} \]

where \( \sum_{i=1}^{n} \frac{P(E|K_i)}{P(K)} \) is the average likelihood with \( P(K_i)/P(K) \) as a weight factor. The expected weight of evidence of a true hypothesis \( H^* \) with respect to \( H \) is for outcome \( x \) of an experiment \( X \) is:

(30) \[ I(H^*, H) = \sum_{x} \frac{P(x|H^*) \cdot \log_2 \left( \frac{P(x|H^*)}{P(x|H)} \right)}{P(K)} \]

and verisimilitude is defined as follows:

(31) \[ \text{Ver}(H) < \text{Ver}(K) \text{ if } I(H^*, H) > I(H^*, K). \]

This account is open to a very basic objection. The application concept of likelihood presupposes that an hypothesis may be true. For a theory which is actually refuted, the weight factor in (30) may render \( I(P,H) \) undefined as \( P(x|H) = 0 \) for \( 0 \leq P(x|H^*) \leq 1 \). The very point, however, of a theory of verisimilitude is to be able to make truthlikeness comparisons between actually false theories or actually false hypotheses. The model of Rosenkrantz is not designed to do this, and consequently it fails to be a satisfactory general account of verisimilitude.

IX. \( \text{Wójcicki's Account of Approximate Truth} \)

Ryszard Wójcicki\(^{41} \) has produced a definition of approximate truth of a set \( A \) of sentences of a first-order language \( L \), as part of his project of developing a formal methodology of the empirical sciences. Before we can examine this definition, a number of other formal concepts must be discussed.

First, a set-theoretical model of an empirical theory is an ordered set:

\[ (S) < L, \vdash, A_0, K> \]

such that \( K \) is a set of strictly similar operational empirical systems, \( L \) is a language conformed with the set of all quantitative systems similar to idealizations of the systems in \( K \), \( A_0 \) is a set of sentences of \( L \), and \( \vdash \) is a derivability relation defined on the set of sentences of \( L \). A number of further concepts now require explication, beginning with the concept of a quantitative system.

By \( \langle t, a \rangle \) we denote that denoted by the sentence 'the object \( a \) taken at the time \( t \)', and shall view every ordered pair of the form \( \langle t, a \rangle \) as a thing-slice. The set \( \text{Ob}(U) \) is the set of empirical objects and the set \( U \) is the set of thing-slices of objects in \( \text{Ob}(U) \). We say that \( a \) exists at time \( t \) if and only if \( \langle t, a \rangle \in U \). The interval \( i_0(a) = \{ t: \langle t, a \rangle \in U \} \), is called the period of existence of the object \( a \) in \( U \) and
the union, \( \iota(U) = \bigcup \{ \iota(a) : a \in \text{Ob}(U) \} \) will be called the period of existence of \( U \), the universe. If \( U \) and \( V \) are two universes, and if \( V \subseteq U \), then \( V \) is a subuniverse of \( U \). The symbol \( 'U(n)' \), where \( n \) is a natural number \( \geq 1 \), denotes the set of \( n \)-th limited Cartesian powers of \( U \):

\[
(32) \quad U^{(n)} = \{ \langle t, a_1, a_2, \ldots, a_n \rangle : \langle t, a_1 \rangle, \ldots, \langle t, a_n \rangle \in U \}
\]

and the mapping \( F : U(n) \to \mathbb{R} \) for real numbers is a \( n \)-ary numerical parameter on \( U \): the number \( n \) is the arity of \( F \). By the symbolization '\( F(t, a_1, \ldots, a_n) = x' \) it is meant that the magnitude \( F \) measured on the objects \( a_1, a_2, \ldots, a_n \) at time \( t \), takes the value \( x \). On this basis a quantitative structure defined on \( U \) is a structure \( \mathfrak{X} = \langle X, F_1, \ldots, F_n \rangle \) if and only if \( U \) is a universe, \( X \) is a subuniverse of \( U \) and \( F_1, \ldots, F_n \) are numerical parameters defined on \( U \). Any two structures, \( \mathfrak{X} \) and \( \mathfrak{Y} \), such that \( \mathfrak{X} = \langle X, F_1, \ldots, F_n \rangle \) and \( \mathfrak{Y} = \langle Y, G_1, \ldots, G_n \rangle \), not necessarily defined on the same universe, are similar if and only if for every \( i, 1 \leq i \leq n \), the parameters \( F_i \) and \( G_i \) are of the same arity. If \( F_1, \ldots, F_n \) are \( S_1, \ldots, S_n \)-ary parameters respectively, then the similarity type of the structure \( \mathfrak{X} \) is \( S = \langle S_1, S_2, \ldots, S_n \rangle \). Two structures \( \mathfrak{X} \) and \( \mathfrak{Y} \) are of the same similarity type if and only if the similarity type of both structures is \( S = \langle S_1, S_2, \ldots, S_n \rangle \).

If for every \( i, F_i = G_i \), then \( \mathfrak{X} \) and \( \mathfrak{Y} \) are said to be strictly similar.

The second concept requiring explication is that of an operational structure. Let \( F \) be a \( k \)-ary quantity defined on a universe \( U \). An operational measurement of \( F \) is an operation \( p \) which transforms \( F \) into a function \( pF \) such that:

\[
(33) \quad (1) \quad pF : U(k) \to \mathbb{R}, \text{ and } (2) \quad (\forall \langle t, a_1, \ldots, a_k \rangle) [(\langle t, a_1, \ldots, a_k \rangle \in U(k)) \rightarrow F(t, a_1, \ldots, a_k) \in pF(t, a_1, \ldots, a_k)].
\]

Consider a quantitative structure \( \mathfrak{X} = \langle X, F_1, \ldots, F_n \rangle \) defined on a universe \( U \) with \( p = \langle p_1, \ldots, p_n \rangle \), where '\( p_i \)', '\( p_n \)' denote operational measures of \( F_1, \ldots, F_n \) respectively. Consider a set \( \mathfrak{Y} = \langle X, \mathfrak{p}_1, \ldots, \mathfrak{p}_n \rangle \). If there exists a \( p = \langle p_1, \ldots, p_n \rangle \) such that for every \( \mathfrak{p}_1, \mathfrak{p}_j = p_1 F_1, \text{ then } \mathfrak{X} \) is said to be an operational structure defined on \( U \) and \( p\mathfrak{X} = \langle X, p_1 F_1, \ldots, p_n F_n \rangle \) is an operational system corresponding to \( \mathfrak{X} \). If an operational structure \( \mathfrak{X} \) corresponds to a quantitative structure \( \mathfrak{Y} \) (i.e., they are strictly similar), then \( \mathfrak{Y} \) is an idealization of \( \mathfrak{X} \).

Consider now a language \( \mathcal{L} \) conformed to a set \( \mathfrak{K} \) of all quantitative systems of a given similarity type, and let \( \mathfrak{X} \) be an operational system corresponding to a structure \( \mathfrak{X} \) in \( \mathfrak{K} \). Then a sentence \( \phi \) of the language \( \mathcal{L} \) is true if and only if it is true in every idealization of \( \mathfrak{X} \). A sentence \( \phi \) of \( \mathfrak{L} \) is false in \( \mathfrak{X} \) if and only if it is false in every idealization of \( \mathfrak{X} \). A sentence \( \phi \) of \( \mathcal{L} \) is indeterminate in \( \mathfrak{X} \) if and only if it is neither true nor false for every idealization of \( \mathfrak{X} \). A set of sentences \( A \) of \( \mathcal{L} \) is approximately true in \( \mathfrak{X} \) if and only if there is an idealization \( \mathfrak{Y} \) of \( \mathfrak{X} \) such that every sentence \( \phi \) of \( A \) is true in \( \mathfrak{Y} \).

This account of approximate truth has little to offer as an account of verisimilitude. Note that if it is the case that not every sentence \( \phi \) of \( A \) is true in some idealization \( \mathfrak{F} \), then the set of sentences of \( A \) of \( \mathcal{L} \) is not approximately true. For the operational structure \( \mathfrak{X} \) to have an idealization \( \mathfrak{F} \), it is sufficient that \( \mathfrak{F} \) be a quantitative structure, and that \( \mathfrak{X} \) and \( \mathfrak{F} \) be strictly similar. Let \( \mathfrak{X} = \langle X, \mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_n \rangle \) and \( \mathfrak{F} = \langle X, F_1, F_2, \ldots, F_n \rangle \).
\( \langle Y, G_1, \ldots, G_a \rangle \). Then if \( X \) and \( \mathcal{F} \) are strictly similar, then \( \tilde{\phi}_1 \equiv G_1 \) & \( \tilde{\phi}_2 \equiv G_2 \) & \( \ldots \) & \( \tilde{\phi}_a \equiv G_a \) is true. Now since strictly similar structures may differ only as to the sets of objects they involve, it may be assumed, given the condition that every sentence \( \phi \) of \( A \) is true in \( \mathcal{F} \), that \( X \subseteq Y \). Hence it is possible that there is a \( y \) such that \( y \in Y \) but \( y \notin X \). Thus \( G_1 \gamma \) and \( \tilde{\phi}_1 \gamma \) may differ in truth-value, in particular, that \( G_1 \gamma \) may be true, whilst \( \tilde{\phi}_1 \gamma \) is false. But if it is the case that \( \tilde{\phi}_1 \equiv G_1 \) is true, then if every sentence \( \phi \) of \( A \) is true in \( \mathcal{F} \), then \( \tilde{\phi}_1 \gamma \) and \( G_1 \gamma \) cannot differ in truth value. Hence \( X = Y \). But if \( X \subseteq Y \), then by the axiom of extensionality of sets, \( X = \mathcal{F} \). This means that \( X \) is approximately true if and only if every sentence of \( \mathcal{F} \) is true, that is, that \( \mathcal{F} \) is true. But this result is parallel to the Miller-Tichý-Harris Theorem and renders Wojcicki's concept of approximate truth theoretically unworkable.

X. Relevance Logic and Verisimilitude

Chris Mortensen in his paper entitled "A Theorem on Verisimilitude",42 argued that the Miller-Tichý-Harris Theorem is dependent upon the classical logical assumption \( \gamma : \alpha \in A \) and \( \gamma \beta \in A \), then \( \beta \in A \). This principle, however, fails for large classes of theories based on logics other than classical logic. Mortensen argued that there exist two RM3-theories, \( A, B \) (RM3 being but one system discussed by Anderson and Belnap in Entailment41) such that the verisimilitude of \( B \) is greater than the verisimilitude of \( A \), i.e., \( B \uparrow A \), and \( B \) has at least one false consequence, i.e., \( B \not\models \{ \beta \} \). A brief review of this result will be given and its limitations in turn outlined.

Consider a language \( L \) which has a denumerable number of constants, \( p_1, p_2, \ldots, \overline{p}_s, \ldots \) closed under negation and conjunction. The RM3 matrices are as follows:

<table>
<thead>
<tr>
<th>( \delta )</th>
<th>T</th>
<th>N</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma )</td>
<td>T</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>( \overline{\delta} )</td>
<td>N</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>( \overline{\gamma} )</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>( \overline{\delta} )</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

We take \( \mathcal{A} = (\{ \forall \mathcal{A}(\alpha) = T \) or \( \forall \mathcal{A}(\alpha) = N \}), where if \( \alpha \) is of the form \( \delta \& \overline{\gamma} \), then \( \forall \mathcal{A}(\alpha) \) is determined from the RM3-matrix for \( '\&' \) and if \( \alpha \) is of the form \( \neg\alpha (\delta \& \overline{\gamma}) \) for some \( n \geq 1 \), then \( \forall \mathcal{A}(\alpha) \) is also determined from the RM3 matrix for \( '\&' \). Further, for all \( n \geq 0 \), \( \forall \mathcal{A}(\neg \alpha \overline{p}_1) = \forall \mathcal{A}(\neg \alpha \overline{p}_2) = N \) and for all \( n_0 \geq 0 \) and \( n_2 \geq 3 \), \( \forall \mathcal{A}(\neg \alpha \overline{p}_m) = T \) and \( \forall \mathcal{A}(\neg \alpha \overline{p}_m) = F \). We also take \( \mathcal{B} = (\alpha : \forall \mathcal{B}(\alpha) = T \) or \( \forall \mathcal{B}(\alpha) = N \}), where \( \alpha \) is of the form \( \delta \& \overline{\gamma} \), then \( \forall \mathcal{B}(\alpha) \) is determined from the RM3-matrix for \( '\&' \) and if \( \alpha \) is of the form \( \neg\alpha (\delta \& \overline{\gamma}) \) for some \( n \geq 1 \), then \( \forall \mathcal{B}(\alpha) \) is also determined from the RM3-matrix for \( '\&' \). Further for all \( n_0 \geq 0 \), \( \forall \mathcal{B}(\neg \alpha \overline{p}_1) = N \) and for all \( n_0 \geq 0 \) and \( n_2 \geq 2 \), \( \forall \mathcal{B}(\neg \alpha \overline{p}_m) = T \) and \( \forall \mathcal{B}(\neg \alpha \overline{p}_m) = F \).

Now it is possible to construct a theory \( \mathcal{B}^* \), such that \( \mathcal{B}^* = (\alpha : \forall \mathcal{B}^*(\alpha) = T \) or \( \forall \mathcal{B}^*(\alpha) = N \}), and for all \( n_0 \geq 0 \) and \( n_2 \geq 1 \), \( \forall \mathcal{B}^*(\neg \alpha \overline{p}_m) = T \) and \( \forall \mathcal{B}^*(\neg \alpha \overline{p}_m) = F \); if \( \alpha \) is of the form \( \delta \& \overline{\gamma} \), then \( \forall \mathcal{B}^*(\alpha) \) is determined from the RM3-matrix for \( '\&' \) and if \( \alpha \) is of the form \( \neg\alpha (\delta \& \overline{\gamma}) \) for some \( n \geq 1 \), then \( \forall \mathcal{B}^*(\alpha) \) is also determined from the RM3-matrix for \( '\&' \). The principal theorem follows from the following five lemmas:

(34) Lemma 1. \( A, B, B^* \) are RM3 theories.
WHAT IS WRONG WITH VERISIMILITUDE?

Lemma 2. \( B^c \subseteq B \subseteq A \).

Lemma 3. \( A \) is nontrivial.

Lemma 4. \( B^c \) is a classical theory, which is negation consistent and complete in \( L^c \).

Lemma 5. Let \( T = B^c \), then \( A \cap T = B \cap T = T \), and \( B \cap F \neq \emptyset \).

Since a large class of relevant logics are weaker than RM3, E-, R-, EM-, RM-, etc. are also suitable logics for the establishment of the results of the principle theorem.

The rejection of the principle \( \gamma : \text{if } \alpha \in A \text{ and } \neg \gamma \vee \beta \in A \), then \( \beta \in A \) only serves to block one part of the Miller-Tichý-Harris Theorem, as Mortensen is well aware. Lemma 1 of the proof of (MTHT) given in section 2 of this paper, does not depend upon \( \gamma \), and the argument can only be blocked from a strictly logical point of view, by rejecting either: (1) \( f \in B \) and \( b \in B \), then \( f \& b \in B \); (2) \( f \in F \), then \( F \& b \in F \), or (3) \( f \& b \in A \), then \( b \in A \). None of these principles are rejected by any standard relevance logic. Hence Mortensen's hypothesis that the Miller-Tichý-Harris Theorem can be avoided by a change to relevance logic is falsified. Relevance logic gives no general solution to this problem.

More recently Mortensen has established that the Miller-Tichý-Harris Theorem cannot be avoided by using a relevance logic as a logic for scientific theories, since a severely limiting theorem can be proved for Popper's definition in even weak relevance logics. Consider a theory \( A \). Then \( A \) is prime if and only if, whenever \( a \vee b \in A \), then at least one of \( a, b \in A \). Mortensen has shown that the Miller-Tichý-Harris Theorem holds for all consistent prime theories. In addition he has established the following limitation theorems:

(35) (MT1) If \( A, B \) are \( L \)-theories, and \( L \) is prime, then if \( B \subseteq A \) and \( A \subseteq B \), then \( T \subseteq A \).

(MT2) If \( A, B \) are \( L \)-theories, and \( L \) is prime, then if \( A \forall B \), then \( A \subseteq T \) or \( T \subseteq A \).

(MT3) If \( B \subseteq A \) and \( A \subseteq B \) and \( A \) is complete, then \( A = T \).

(MT4) If \( A, B \) are classical theories and \( A \) is consistent and complete, then if \( A \forall B \), then \( A = T \).

It is the case that if an \( L_\alpha \)-theory is consistent and prime, then \( \gamma \) holds for it, for suppose \( a \in A \) and \( \neg \alpha \vee b \in A \), then since \( A \) is consistent \( \neg \alpha \notin A \), but if \( A \) is prime at least one of \( \alpha, b \in A \), so \( b \in A \). Thus most of Mortensen's limitation theorems are only directed against classical \( L_\alpha \)-theories. However, (MT1) is applicable to theories which may be paraconsistent (although non-trivial), as will now be demonstrated. Assume that \( B \subseteq A \) and \( A \subseteq B \) and that \( T \subseteq A \) for reductio ad absurdum. Let \( t \in T \) and \( t \notin A \) and \( b \in B \& \neg A \) so that \( b \in B, b \in F \), and \( b \notin A \). Then since \( b \in B \), \( t \vee b \in B \) and since \( t \in T \), then \( t \vee b \in T \). Hence \( t \vee b \in B \). But \( t \notin A \) and \( b \notin A \) and since \( A \) is prime, \( t \vee b \notin A \). Hence \( t \vee b \notin A \), thus contradicting \( B \subseteq A \).
Mortensen has also considered the possibility of avoiding the Miller-Tichy-Harris Theorem by intensionalizing the metalanguage. He has shown, however, that if the intensional verisimilitude relation holds between sets, then so does Popper's extensional relation, since it is a theorem of all the standard relevance logics that: \((\forall x) (Fx \to Gx) \to (\forall x) (Fx \supset Gx)\).

These results establish that the relevance program contributes nothing towards avoiding the Miller-Tichy-Harris Theorem. Indeed, as I pointed out earlier, Mortensen's program was doomed from the outset, since one part of the limitation theorem went through making use only of \((a \land b) \to a\) and some plausible set-theoretical principles. It is of course true that \(\vdash (X \land Y) \to X\) is rejected in connexive logic. To produce a unified solution to the verisimilitude limitation theorems by a change of logic will then require a more radical regimentation of logic than has yet been anticipated. The loss of provability power may prove to be too great a price to pay in restricting the logic of science, merely to save Popper's theory of verisimilitude. A solution should first be looked for elsewhere.

XI. An Alternative Explication of 'Qualitative Verisimilitude'

It is a merit of any account of verisimilitude that it not only is intuitively plausible, but that it is as well, formally simple and hence easy to operate. Most of the existing accounts of verisimilitude seem to add increasingly complex epicycles rather than increasingly deep insights. In this section I shall present and defend an intuitively plausible and formally simple account of verisimilitude which also avoids the Miller-Tichy-Harris Theorem. Further, my account will meet David Miller's requirement that any adequate account of verisimilitude will not permit any proposed ordering by the relation of verisimilitude to be reversed by a simple linguistic reformulation. The paper will conclude with a response to the problems which Miller raised in his paper, "The Accuracy of Predictions."45

The 'truth' may be viewed as a superset \(\sum_T\) which contains all true statements, and the superset \(\sum_F\) which contains all false statements. In my opinion truthlikeness or verisimilitude is best viewed as the "closeness" in size of theories to \(\sum_T\) and "smallness" in size to \(\sum_F\). I take the metaphor of 'size' to be best explicated by the notion of the cardinality of a set. Rather than define cardinal numbers as certain types of ordinal numbers it is proposed that one simply add to some 'logically regimented' set-theory, such as Zermelo-Fraenkel set-theory, the following axiom for cardinal numbers:

\[(ACN) \operatorname{Card}_n(A) = \operatorname{Card}_n(B) \iff A \approx B.\]
We shall say that a set $A$ has a greater cardinal number than a set $B$ if and only if there is a 1-1 correspondence between $B$ and a proper subset of $A$, but no 1-1 correspondence between $B$ and the whole of $A$. The set $A$ will be said to have a small cardinal number than a set $B$ if and only if $B$ has a greater cardinal number than $A$. It is hardly necessary to present a fully formal theory of cardinal numbers to supplement this gloss, since the reader can readily find this in the standard texts on set-theory. Without further ado, our definition of verisimilitude is given:

$$(SDV) A \triangleright v B = \sigma \{ (\text{Card}_N(\mathcal{A}_T) > \text{Card}_N(\mathcal{B}_T))$$
$$\land (\text{Card}_N(\mathcal{A}_F) \preceq \text{Card}_N(\mathcal{B}_T))$$
$$\lor (\text{Card}_N(\mathcal{A}_T) \lor \text{Card}_N(\mathcal{B}_T))$$
$$\land (\text{Card}_N(\mathcal{A}_F) < \text{Card}_N(\mathcal{B}_F)) \}$$

This definition is similar to Popper's original qualitative definition of verisimilitude. It has, however, added advantages, ignoring for the moment the problem of the Miller-Tichý-Harris Theorem. First, on Popper's definition when dealing with infinite sets we get no satisfactory indication of the size of the respective sets of truths and falsehoods. The theory of cardinal numbers is especially devised to speak of the sizes of both finite and infinite sets, so any definition of verisimilitude based upon it will be quite general. Second, on Popper's definition, we are committed to the claim that if $A \triangleright v B$, then if $\mathcal{B}_T \subseteq \mathcal{A}_T$, and if $\varnothing \in \mathcal{B}_T$, $\varnothing \notin \mathcal{A}_T$. It is impossible to speak of verisimilitude relationships holding between theories where some sentence $\varnothing \in \mathcal{B}_T$ is not also an element of $\mathcal{A}_T$. That the more comprehensive theory must contain all of the truths of the less comprehensive theory is an assumption which I have no use for, and further it is directly responsible for generating the various limitation results, as is obvious from my proofs given in section II above.

There is no doubt that some analogue of the Miller-Tichý-Harris Theorem might be provable for $(SDV)$, so this issue should be addressed and disarmed. The following limitation theorems might be suggested as suitable analogues:

$$(36) \quad (LT1) \quad A \triangleright v B \rightarrow \text{Card}_N(\mathcal{A}) \preceq \text{Card}_N(\bigcup \mathcal{T})$$

$$(LT2) \quad A \triangleright v B \rightarrow \text{Card}_N(\mathcal{A}_T) \preceq \text{Card}_N(\bigcup \mathcal{T})$$

$$(LT3) \quad A \triangleright v B \rightarrow \text{Card}_N(\mathcal{A}_T) \succ \text{Card}_N(\bigcup \mathcal{T})$$

$$(LT4) \quad A \triangleright v B \rightarrow \text{Card}_N(\mathcal{A}_T) \preceq \text{Card}_N(\mathcal{B}_T)$$

Suppose $\vdash (LT1)$. All this would mean is that a verisimilitude relationship would not exist between $A$ and $B$ if

$\text{Card}_N(\mathcal{A}_T) > \text{Card}_N(\bigcup \mathcal{T})$ which is quite true. Indeed, the latter proposition is self-contradictory, since there is no $\varnothing \in \mathcal{A}_T$
which is not also in \( \sum^\tau \), by definition of \( \sum^\tau \). If on the other hand \( \vdash \) (LT3), then one would not merely refute my account of verisimilitude, but one would have raised a major problem for the theory of cardinal numbers in general, for we may interpret \( T \) to be the set of truths of mathematics and reword \( \vdash \) (LT3) to produce a contradiction in the theory of cardinal numbers. In both cases, however, it is difficult to imagine how such proofs could be forthcoming.

If \( \vdash \) (LT1), this would only mean that our theories must always fall far short of being absolutely comprehensive in their set of truths. Since no one can know that there are no unknowable truths (by definition), it would hardly be much of a limitation if \( \vdash \) (LT1).

Suppose on the other hand that \( \vdash \) (LT4). This would completely refute my account of verisimilitude. However, I see no way of establishing this theorem without in the process raising once more major problems for the theory of cardinal numbers.

It may be objected, as Tichý has suggested,\(^47\) that the verisimilitude of any false theory \( A \) would be increased by adding to \( A \) an arbitrary sentence which does not follow from it. Against this it may be replied that verisimilitude comparisons, if they are to have a point at all, must be strictly between the logical consequences of theories. Tichý's arbitrary sentence does not follow from \( A \) and we take this as a conclusive reason for disregarding it in our verisimilitude assessment. We should further note that this argument of Tichý is undermined by John Harris' proof that we cannot increase the truth-content of a false theory by conjoining a logically independent sentence \( b \) to it, without also increasing the falsity content and vice versa, whether \( b \) is true or not.\(^48\)

It is also noted that (SDV) does not involve one in having to have an impossible comparison theory \( \sum^\tau \) to make verisimilitude estimates. \( \sum^\tau \) stands as a metaphor, and is not a part of the formal mechanics of (SDV). Verisimilitude comparisons are made without recourse to any third comparison theory. Hence John Harris' problems about verisimilitude comparisons between \( A \) and \( B \) with respect to a comparison theory \( C \) do not arise.\(^49\)

Perhaps (SDV) might be taken to fail for theories which are inconsistent within the framework of classical logic. This, however, is not so. The possibility of a classically inconsistent (and hence trivial theory) being closer to the truth than a non-trivial theory is excluded by cardinal number comparisons of the respective sets of false statements of the two theories. Since \( p \land \neg p \rightarrow q \) in classical logic, for a classically inconsistent theory \( A \), \( \text{Card}_n(\sum^\tau A) = \text{Card}_n(\sum^\tau) \) and also \( \text{Card}_n(\sum^\tau A^\tau) = \text{Card}_n(\sum^\tau) \).

I shall also add here that considerations of verisimilitude need not be made solely on the basis of the total number of truths and false-
hoods of two theories. If the set of truths and the set of falsehoods of both theories is infinite, then some filter might be placed upon these sets to decrease their sizes. Thus the set \( A_T \) of truths of A may be deflated by considering the resultant truths relative to a set of problems or some set of interests. This will determine a set \( SAT \subseteq A_T \) of the significant or relevant truths of A. Criteria for deciding this can hardly be expected to be presented in a discussion of merely the explanation of the concept of verisimilitude.

Tichý\(^50\) takes it as non-controversial that a person who maintains that there are exactly eight planets in the solar system is closer to the truth than his/her opponent who insists that there are only five. Agreed. This being granted, this situation does not constitute a counter-example to (SDV). It is one thing for a theory to be closer to the truth than another, another thing for a statement to be closer to the truth than another. The account of verisimilitude advanced here is a purely qualitative one, which assumes that the truth valuations of statements are not 'degrees of truth'. If there existed a satisfactory account of partial truth, then the theory here would need to be modified to incorporate its insights.

This completes the defense of (SDV) which will be offered here. Needless to say, there are still quite substantial epistemological and methodological questions about verisimilitude assessments which cannot be addressed here. The paper will conclude with a response to difficulties facing accounts of verisimilitude which David Miller presented in his paper, "The Accuracy of Predictions".\(^51\)

David Miller has claimed to show that, if theories A and B answer the same quantitative questions and evaluate the same quantities, then the constants that A predicts truly cannot all be constants that B predicts truly, unless B always predicts truly. Hence the constants that B predicts falsely cannot all be constants that A predicts falsely, unless B always predicts truly. Whilst Miller gives a general algebraic argument to this conclusion, it is more informative to review an example which he himself cites.

Let \( H \) and \( F \) be two physical constants, and suppose that A asserts that \( H = 8 \) and \( F = 0 \), and B asserts that \( H = 7 \) and \( F = 2 \). Suppose that \( H \)'s true value is \( 0 \), whilst \( F \)'s is \( 2 \). The theory A is then incorrect with respect to both constants, whilst B is right about one of them. Miller asks us then to consider the physical constants \( P \) and \( X \), defined as \( P = H + F \) and \( X = H + 4F \). The predicted values and the true values of \( P \) and \( X \) are now such that A is correct about \( X \), but wrong about \( P \), whilst B is wrong about both. Indeed, as long as A is false for both \( H \) and \( F \) and B for only one, there will always be a linear combination of \( H \) and \( F \) for which A predicts the true value whilst B is in error.

We can see that something is drastically wrong with Miller's argument if we allow \( P = f(H,F) \) and \( X = g(H,F) \). Suppose we take Miller's own values for \( H \) and \( F \) as given in the above example, and let \( P = H + F \) and \( X = H + 4F \). Then, for Miller's own values, A which was false for both \( H \) and \( F \) is now true for both \( P \) and \( X \) and B which was true about \( F \) is now false for both \( P \) and \( X \). As a second example, consider I.J. Good's proposal,\(^52\) that a switch between A and B can occur even when they refer to a one-dimensional parameter. Suppose that the values of
the parameters predicted by theories A and B are \( \alpha \) and \( \beta \) respectively, such that \( 0 < \alpha' < \beta' \), then in some cases, by taking the decimal representation of the respective numbers and interchanging the first two digits, then the third and fourth digits, and so on, it will be found that the transformed values \( \alpha' \) and \( \beta' \) are such that \( 0 < \beta' < \alpha' \). Thus, if \( \alpha = \alpha' \cdot a_3 a_2 ... \) and \( \beta = \beta' \cdot b_2 b_1 ... \) then \( \alpha' = \alpha' \cdot a_3 a_2 ... \) and \( \beta' = \beta' \cdot b_2 b_1 ... \) and \( a_3 > b_2 \).  

In reply, it is alleged that the strategies of Miller and Good are outrightly bogus. Miller hasn't shown how his result can be produced in any physical theory, and, until he does, the claim that \( P \) and \( X \) are 'physical constants' is unjustified. The physical significance of reversing decimal points is a game played by manipulating symbols which is of no philosophical and scientific significance.

ENDNOTES


3 P.M. Quay, "Progress as a Demarcation Criterion for the Sciences", *Philosophy of Science*, 41, 1974, 154-70. Citation p. 154.


6 The thesis of the progress of science through increasing increase in verisimilitude is to be distinguished from the first thesis of convergent realism as stated by L. Laudan ("A Refutation of Convergent Realism", in U.J. Jensen and T. Harre [eds], *The Philosophy of Evolution* [Sussex: Harvester Press, 1980], 232-68):

\[(R1) \] Scientific theories (at least in the 'mature' sciences) are typically approximately true and more recent theories are closer to the truth than older theories in the same domain. .. (pp. 233-4)

Thesis (R1) is quite problematic, and this situation has not been improved even by C.L. Hardin and A. Rosenberg's response to Laudan ("In Defense of Convergent Realism", *Philosophy of Science*, 49, 1982, 604-15). Thesis (R1), as it originally occurs in R. Boyd's "Scientific Realism and Naturalist Epistemology" (manuscript) makes use of the notion of a 'mature science' to rule out counter-examples of reference failure made by recourse to any 'arbitrarily chosen scientific theory'. Boyd takes 'mature scientific theories' to be those which have passed a take-off point. This concept, in turn, insofar as it is explicated at all, is explicated by reference to the concept of truth or approximate truth
of certain background theories. This, however, leads us straight into a vicious infinite regress, as the concept of 'mature science' was initially introduced to rule out counter-examples of reference failure, and in turn to defend the convergent realist's idea of approximate truth. Boyd's explication only serves to lead us back to the very ideas of truth and approximate truth.

We have argued elsewhere (B.C. Goodwin, G. Webster and J.W. Smith, "Neo-Darwinism and Constructional Biology" [submitted for publication]) that (R1) is empirically false: a 'mature science' such as the neo-Darwinist account of evolution is not closer to the truth than an older position such as Rational Morphology. The realist thesis defended in the present paper is the conditional claim that, if scientific progress occurs, then an increase in verisimilitude of the compared theories will occur.

7 Popper, op.cit., note 2, 235.


11 The present paper is primarily concerned with qualitative accounts of verisimilitude, and the difficulties in presenting a quantitative account of verisimilitude must be addressed elsewhere.


13 A source of possible confusion should be disarmed at this point. It might be argued, as is suggested by some remarks of R. Harré (cf. "The Conditions for Applying an Evolutionary Model to the Development of Science: Commentaries on Mittelstrasse, Laudan and Newton-Smith", in U.J. Jensen and R. Harré [eds], The Philosophy of Evolution [Sussex: Harvester Press], 290-94, esp. 292-3) that the present "logicist" conception of the nature of scientific theories is responsible for the Miller-Tichý-Harris Theorem. This theorem might be taken to constitute a reductio ad absurdum of such a conception of scientific theories. (For this style of argument, cf. R. Harré, The Principles of Scientific Thinking [London: Macmillan, 1970]). However, little of significance is at stake in our present use of the term 'theory'. It may be replaced by another term such as 'set', and the Miller-Tichý-Harris Theorem will still stand, as long as Popper's original qualitative definition of verisimilitude is upheld and as long as we can still meaningfully form the sets of logical consequences of a theory.

14 Tichý, op.cit., note 8.
16 The failure of verisimilitude orderings with respect to the relation of logical equivalence is a criticism which Graham Oddie ("Verisimilitude and Distance in Logical Space", *Acta Philosophica Fennica*, 30, 1976, 227-42) has successfully advanced against David Miller's own account in "On the Distance from the Truth as a True Distance", *Bulletin of the Section of Logic, Polish Academy of Sciences*, 6, 1977, 15-26, and "On the Truthlikeness of Generalizations", in R. Butts and J. Hintikka (eds), *Basic Problems in Methology and Linguistics* (Dordrecht: D. Reidel, 1977), 121-47.


18 It is strange that someone who accepts J. Hintikka's notion of depth and his account of distributive normal forms in first-order logic (cf. "Distributive Normal Forms in First-Order Logic", in *Logic, Language-Games and Information* [Oxford: Clarendon Press, 1973], 242-86) would choose to take this exit, since Hintikka is a sharp critic of the thesis that tautologies are "uninformative".


20 Ibid., 35.


24 Tichý, *op.cit.*, note 22, 189.

25 Ibid., 189.


29 Niiniluoto, "Verisimilitude and Scientific Progress", *op.cit.*., note 26, 255.

WHAT IS WRONG WITH VERISIMILITUDE?

31 Ibid., 608-9.


33 Cf. B. Sobociński, "On Well Constructed Axiom Systems", Rocznik Polskiego Towarzystwa Naukowego Na Obczyźnie, 6, 1955-6, 54-63. Citation is from 65.

34 Mott, op.cit., note 32, 256.

35 Ibid., 263.


42 Mortensen, op.cit., note 12.


44 Mortensen, "Relevance and Verisimilitude", op.cit., note 12.


47 Tichý, op.cit., note 8, fn. 2.

48 Harris, op.cit., note 9, 163.

49 Harris, op.cit., 163.

50 Tichý, op.cit., note 22, 175.

51 Miller, op.cit., note 45.

52 I.J. Good, "Comments on David Miller", Synthese, 30, 1975, 205-6.