Set-theoretical invariance criteria for logicality

Solomon Feferman

Abstract. This is a survey of work on set-theoretical invariance criteria for logicality. It begins with a review of the Tarski-Sher thesis in terms, first, of permutation invariance over a given domain and then of isomorphism invariance across domains, both characterized by McGee in terms of definability in the language $L_{\infty,\infty}$. It continues with a review of critiques of the Tarski-Sher thesis, and a proposal in response to one of those critiques via homomorphism invariance. That has quite divergent characterization results depending on its formulation, one in terms of FOL, the other by Bonnay in terms of $L_{\infty,\infty}$, both without equality. From that we move on to a survey of Bonnay's work on similarity relations between structures and his results that single out invariance with respect to potential isomorphism among all such. Turning to the critique that calls for sameness of meaning of a logical operation across domains, the paper continues with a result showing that the isomorphism invariant operations that are absolutely definable with respect to KPU−Inf are exactly those definable in full FOL; this makes use of an old theorem of Manders. The concluding section is devoted to a critical discussion of the arguments for set-theoretical criteria for logicality.

1. Introduction. This is a survey of work in terms of set-theoretical invariance criteria on the question: *Which truth-valued operations on one or more relations are to be regarded as logical?* It is a sequel to my article [11] that took for its point of departure Tarski's thesis [29], as modified by Sher [26]. Tarski had proposed to identify the logical operations on relations over a given domain D with those that are invariant with respect to arbitrary permutations of D. Sher generalized this to operations across domains that are invariant with respect to bijection between domains (equivalently, isomorphism of structures with these domains). McGee [21] characterized the logical operations in Tarski's sense as precisely those that are definable in the language $L_{\infty,\infty}$ with equality over a given domain, and from that he obtained a related characterization of the operations that are logical in Sher's sense.

I critiqued the Tarski-Sher thesis in [11] on three grounds, the first of which is that it assimilates logic to mathematics, the second that the notions involved are not set-theoretically robust, i.e. not absolute, and the third that no natural explanation is given by the thesis of what constitutes the *same* logical operation over arbitrary basic domains. In this last respect, as an example of a notion that could compare domains of different cardinality, I had proposed consideration in [11] of the homomorphism invariant operations (in a strong sense); it was shown op. cit. that the operations that are definable from monadic homomorphic invariant operations are exactly those expressible in the first-order predicate calculus $L_{\omega,\omega}$ without equality. However, Bonnay [6], [7] later characterized the operations that are outright homomorphism invariant as just those definable in the language $L_{\infty,\infty}$ without equality. Bonnay went on to consider operations that are invariant under other kinds of similarity relations. His main results distinguish potential isomorphism (Iso_p) among all such relations, and that has led him to propose Iso_p-invariance as the criterion for logicality; it turns out that the operations invariant under potential isomorphism go somewhere beyond those definable in the language $L_{\infty,\omega}$.

The Tarski-Sher thesis and McGee's results concerning it are reviewed in sec. 2, then my critiques of it and result for homomorphism invariant operations are reviewed in sec. 3, and Bonnay's work is described in sec. 4. Following that, I propose in sec. 5 an explanation of what constitutes the same operation across arbitrary domains in terms of those that are uniformly definable within the language of set theory. Moreover, in order to meet the second critique above, one should restrict to definitions that are absolute with respect to a system of set theory that makes no assumptions about the size of the universe. Specifically, I look at operations on relational structures that are definable in an absolute way relative to KPU−Inf, i.e. Kripke-Platek set theory with urelements and without the Axiom of Infinity. It is shown to follow from an old result of Manders [19] (reproved in Väänänen [30]) that the operations in question on structures whose domains consist of urelements are exactly those expressible in the ordinary first-order predicate calculus with equality. The aguments in favor of set-theoretical invariance criteria for logicality are discussed critically in the concluding section 6; despite the attraction of various of the results that have been obtained, my overall conclusion is that none of the set-theoretical invariance proposals on offer provide a sufficiently convincing criterion for logicality in their own right.[1]

2. The Tarski-Sher thesis; a review. Tarski's article, "What are logical notions?" [1986] was based on the text of a lecture that he had given for a general audience at Bedford College, London, in 1966. With Tarski's agreement, it was edited by John

Corcoran, but it did not appear until three years after his death in 1983. Tarski's answer to the question in his title is informal, but essentially it takes logical notions to be relations between individuals, classes and relations over an arbitrary non-empty domain D of individuals, and singles out the logical relations as exactly those that are invariant under arbitrary permutations of D. In his lecture, Tarski gave several simple examples of logical notions in this sense, as follows:

(i) The only classes of individuals which are logical are the empty class and the universal class.

(ii) The only binary relations between individuals which are logical are the empty relation, the universal relation, the identity relation and its complement.

(iii) At the next level, i.e. classes of classes of individuals, Tarski mentions as logical notions those given by cardinality properties of classes such as "that a class consists of three elements, or four elements...that it is finite, or infinite—these are logical notions, and are essentially the only logical notions on this level."

(iv) Finally, among relations between classes (of individuals) Tarski points to several which are "well known to those of you who have studied the elements of logic" such as "inclusion between classes, disjointness of two classes, overlapping of two classes," and so on.

Tarski did not attempt to give examples of logical notions in higher types than those in (iii) and (iv), though, as explained in [11], his proposal makes sense for objects in the finite relational type structure over D, where the objects at each level are relations of one or more arguments between objects of lower levels. Nor did Tarski raise the question of characterizing the logical notions, and more generally of the operations on members of the type structure that are invariant under arbitrary permutations. This is understandable in view of the general audience to which his lecture was addressed. The first such characterization was provided by McGee [21], who showed that an operation is logical according to Tarski's permutation-invariance criterion if and only if it is definable in the language $L_{\infty,\infty}$; this is the language defined in set theory which allows—in addition to the operation of negation—conjunctions and disjunctions of any cardinality, together with universal and existential quantification over a sequence of variables of any cardinality.

For simplicity in the following, and as is common in discussions of logicality, we shall restrict attention to the question of what are logical notions Q of type level 2, i.e. relations between relations R between individuals.[2] The letter 'Q' is used here for such because logical notions in Tarski's sense at this level over a given domain D are the restriction to D of a generalization of quantifiers due to Lindström [17]. Given $\underline{R} = (R_1,\ldots,R_n)$ with R_i a k_i-ary relation between elements of D (k_i a non-zero natural number) we write $Q_D(R_1,\ldots,R_n)$ or $Q_D(\underline{R})$ to express that the relation Q_D holds of $\underline{R}$ over D. For each permutation π of D, each R_i, and each k_i-ary sequence $\underline{a}$ of elements of D, let $\pi(R_i)$ be the relation that holds of $\pi(\underline{a})$ if and only if R_i holds of $\underline{a}$; then $\pi(Q_D)$ is defined to be the relation that holds of $(\pi(R_1),\ldots,\pi(R_n))$ if and only if Q_D holds of $(R_1,\ldots,R_n)$. In these terms we can now define:

> Q_D is a *logical notion in Tarski's sense over* D if and only if Q is invariant under all permutations of D, i.e. $\pi(Q_D) = Q_D$ for all permutations π of D.

In the following we also think of relations as operations to truth values, i.e. we take $Q_D(\underline{R}) = T$ if Q_D holds of $\underline{R}$ and $= F$ otherwise. In those terms we call Q_D a *logical operation* (in Tarski's sense) over D if it meets the permutation invariance criterion. In general we shall treat the Q_D as relations and as operations interchangeably.

Tarski's examples (iv) of logical notions over an arbitrary domain are the inclusion relation, the disjointness relation, and the overlapping relation; they are relations (or operations) of *monadic type*, i.e. have unary relations as arguments. The first holds between two classes A and B of individuals just in case $A \subseteq B$, the second holds just in case $A \cap B = \varnothing$ and the third just in case $A \cap B \neq \varnothing$; formal-logically speaking these are expressed in the first case by use of the universal quantifier together with implication, in the second case by the same with negation, and in the third case by use of the existential quantifier together with conjunction. The pure universal quantifier $\forall$ relative to D is the unary relation of monadic type that holds of A just in case A = D, while the pure existential quantifier $\exists$ relative to D holds of A just in case $A \neq \varnothing$. For each cardinal number κ, the cardinality quantifier $\exists!_\kappa$ is also of monadic type and consists of all subclasses A of D whose cardinality card(A) = κ, while the quantifier $\exists_{\geq\kappa}$ consists of all A for which card(A) $\geq \kappa$. In particular, the "infinitely many…exist" quantifier is given by $\exists_{\geq\omega}$. All these are logical notions in Tarski's sense. Not mentioned by Tarski are examples of permutation invariant notions of non-monadic type, such as being a linear ordering or well-ordering.

We are here taking Tarski's extensional, set-theoretical framework at face value for dealing with the question: what are logical notions?—and save any reconsideration of that until the end. Granted that framework, the permutation invariance criterion is a natural *necessary* condition for logicality if one agrees that what counts as a logical notion should be independent of the nature of the particular entities in a given domain of discourse and of the properties of those entities. Tarski himself motivated it in relation to the Klein *Erlanger Programm*, which identified the notions to be studied in various geometries such as Euclidean, affine and projective geometry according to the groups of (one-one and onto) transformations under which they are invariant; similarly the notions appropriate to topology are those invariant under all homeomorphisms of a topological space with itself. With logic thought of as the mathematics of structures of the most general sort, i.e. with no distinguished mathematical content, the transformations to be considered are simply all the permutations. Actually, the criterion was not original with Tarski; it was apparently first proposed by F. I. Mautner [20], though he pursued the idea in a somewhat different direction from the one taken by Tarski. But it had already been noted in an article by Lindenbaum and Tarski [16] that every relation definable in the simple theory of types is provably invariant under every permutation of the domain of individuals. It is thus surprising that he did not expressly have this in mind when he raised the issue of the division between logical and extra-logical notions in his article [27] on logical consequence, instead of saying that "...no objective grounds are known to me which permit us to draw a sharp boundary between the two groups of terms." And within the Tarski school itself, his former student Andrzej Mostowski [23] had already brought attention to those unary operations of monadic type that are invariant under all permutations of the domain of individuals, including the various cardinality quantifiers mentioned above.

Consider a relation Q_D over a domain D of signature $\sigma = (k_1,\ldots,k_n)$ $(k_i > 0)$. As already mentioned, McGee characterized which such Q_D are invariant under permutations of D in terms of the language $L_{\infty,\infty}$, whose formulas for the statement of his result are generated as follows.

(i) For each i = 1,..., n and k_i-ary sequence of variables $\underline{x}$, $P_i(\underline{x})$ is an atomic formula; also each equation between variables is an atomic formula;

(ii) if φ is a formula then $\neg\varphi$ is a formula;

(iii) if Φ is any non-empty set of formulas then $\bigvee\varphi\ [\ \varphi \in \Phi\ ]$ is a formula;

(iv) if φ is a formula and U is any non-empty set of variables then $(\exists U)\varphi$ is a formula.

Given a domain D, an interpretation $\underline{R} = (R_1, \ldots, R_n)$ in D of the predicate symbols $P_1, \ldots, P_n$, resp., a formula φ of $L_{\infty,\infty}$, and an assignment s to the free variables of φ in D, one inductively defines as usual

$$(D, \underline{R}) \models \varphi[s].$$

When φ is a sentence, this is simply written

$$(D, \underline{R}) \models \varphi.$$

φ is said to define Q_D over D if for any $\underline{R} = (R_1, \ldots, R_n)$ with R_i a k_i-ary relation in D, we have:

$$Q_D(\underline{R}) = T \text{ iff } (D, \underline{R}) \models \varphi.$$

THEOREM 1 (McGee [21]). Q_D is invariant under arbitrary permutations of the domain D of individuals if and only if Q_D is definable in $L_{\infty,\infty}$.

It is straightforward that every $L_{\infty,\infty}$ definable operation is invariant under arbitrary permutations of the domain of individuals. The idea of McGee's proof in the other direction is to lay out all possibilities for the operation Q_D as its arguments range over all possible $\underline{R} \in D[\sigma]$. This can be achieved using a set W of variables with card(W) = $\kappa + 1$. Enumerate D as $\{d_\alpha : \alpha < \kappa\}$, and W as $\{x_a : \alpha < \kappa\} \cup \{y\}$. The x_α act as formal surrogates of the d_α. Let $\psi_{\underline{R}}$ be the diagram of $\underline{R}$ under this association together with

$\neg(x_\alpha = x_\beta)$ for each $\alpha < \beta$, and then take $\chi_{\underline{R}}$ to be the formula which says that there exist x_a $(\alpha < \kappa)$ such that $\psi_{\underline{R}}$ holds and such that each y in the domain is one of the x_α. Finally, take φ to be the disjunction of all the $\chi_{\underline{R}}$ over all sequences $\underline{R}$ such that $Q_D(\underline{R})$ holds; note that this final disjunction may be of cardinality as large as 2^κ, and the longest quantifier sequences in φ are of length at least κ.

McGee says that this theorem "gives us good reason to believe that the logical operations on a particular domain are the operations invariant under permutations." I shall take strong issue with that below. But even if one accepts that, it is natural not to tie logical operations to specific domains. And, indeed, McGee goes on to consider operations *across domains* which for each non-empty set D of individuals gives a relation Q_D of type σ over D. Then he argues (rightly, in my view), that "(i)n order for an operation across domains to count as logical, it is not enough that its restriction to each particular domain be a logical operation." For example, McGee defines a relation of "wombat disjunction" Q_W across domains which acts like ordinary disjunction when there are wombats in the universe of discourse D and like conjunction when there are no wombats in D. Clearly wombat disjunction is not a logical notion, though on each domain it is invariant under permutations. Another example given is that of "affluent cylindrification" $(A), for A unary, which holds in a domain just in case some rich person belongs to the class A; again this is not a logical operation, but meets the permutation invariance condition on "upper-crust domains" in which every person is rich. However, on an equinumerous domain containing at least one rich and one poor person, the operation $ is not permutation invariant, by taking A to be a singleton of one of these. Thus McGee is led to consider an extension of the permutation invariance criterion for logicality as proposed by Sher [26]: by the *Tarski-Sher thesis*, McGee means the claim that *the logical operations across domains are just those invariant under bijections between them*. The following is then a corollary of Theorem 1.

THEOREM 2 (McGee [21]). An operation Q across domains is a logical operation according to the Tarski-Sher thesis iff for each cardinal $\kappa \neq 0$ there is a formula φ_κ of $L_{\infty,\infty}$ which describes the action of Q on domains of cardinality κ.

More specifically, one can take φ_κ to be the formula constructed for the proof of Theorem 1 for any domain D of cardinality κ. Whatever such φ_κ is taken, in order to obtain a *single* definition of the operation Q across arbitrary domains, one must take something like the disjunction—over the class of all non-zero cardinals κ—of φ_κ conjoined with the sentence expressing that there are exactly κ elements in the domain. This goes well beyond $L_{\infty,\infty}$ as ordinarily conceived.

3. Critiques of the Tarski-Sher thesis; homomorphism invariant operations.

McGee's results lay bare the character of logical operations according to the Tarski-Sher thesis. In my article "Logic, logics and logicism" [11], I raised three basic criticisms of it:

I. The thesis assimilates logic to mathematics, more specifically to set theory.

II. The set-theoretical notions involved in explaining the semantics of $L_{\infty,\infty}$ are not robust.

III. No natural explanation is given by it of what constitutes the *same* logical operation over arbitrary basic domains.

The first of these, also referred to as the "overgeneration problem", speaks for itself, given McGee's results, but it will evidently depend on one's gut feelings about the nature of logic as to whether this is considered objectionable or not. For Sher, to take one example, that is no problem. Indeed, she avers that:

> The bounds of logic, on my view, are the bounds of mathematical reasoning. Any higher-order mathematical predicate or relation can function as a logical term, provided it is introduced in *the right way* into the syntactic-semantic apparatus of first-order logic. (Sher [26], pp. xii-xiii, italics mine)[3]

What that "right way" is for Sher, is spelled out in a series of syntactic/semantic conditions A-E (op. cit. pp. 54-55), of which the crucial ones are the "first-order" condition A—that a logical operation be of type-level at most 2—and condition E, which is that for invariance under bijections. The paradigms of condition A are the cardinality quantifiers of Mostowski [23] and, more generally, the generalized quantifiers of Lindström [17], where the bound variables range over individuals of the domain. But note that despite the appearance of this being limited to first-order quantification, $L_{\infty,\infty}$ also accomodates second-order quantification as a logical operation across domains (in the Tarski-Sher sense). This is seen as follows. First, given formulas $\psi(X)$ and $\theta(x)$ of this language, where X is a second-order variable, by $\psi(\{x: \theta(x)\})$ is meant the result of substituting $\theta(t)$ for each occurrence of an atomic formula $t \in X$ in ψ. Thus, on a domain of cardinality κ, $(\forall X)\psi(X)$ is equivalent to the statement φ_κ that there exist κ elements x_α which are distinct and exhaustive of the domain, and are such that

$$\bigwedge_{S \subseteq \kappa} \psi(\{y : \bigvee y = x_\alpha [\alpha \in S]\})$$

holds. (Again, we require a conjunction of cardinality 2^κ in this formula.) So, from Theorem 2 above, the restriction to bound first order variables is only apparent, and Sher's condition A is not set-theoretically restrictive. By a trick similar to the preceding, we can quantify over arbitrary relations on the domain, and then say that they are functions, etc. In particular, we can express the Continuum Hypothesis and many other substantial mathematical propositions as logically determinate statements on the Tarski-Sher thesis. Of course, if one follows Tarski by allowing consideration of invariant notions in all finite types, the assimilation of logic to set theory is patent on his thesis, without needing to invoke infinite formulas at all. But insofar as one or the other version of the thesis requires the existence of set-theoretical entities of a special kind, or at least of their determinate properties, it is evident that we have thereby transcended logic as the arena of universal notions independent of "what there is".

The critique II is in a way subsidiary to that in I. The notion of "robustness" for set-theoretical concepts is vague, but the idea is that if logical notions are at all to be explicated set-theoretically, they should have the same meaning independent of the exact extent of the set-theoretical universe. For example, they should give equivalent results in the constructible sets and in forcing-generic extensions. Gödel's well known concept of *absoluteness* provides a necessary criterion for such notions, and when applied to operations defined in $L_{\infty,\infty}$, considerably restricts those that meet this test. For example, the quantifier "there exist uncountably many x" would not be logical according to this restriction, since the property of being uncountable is not absolute. My proposed alternative to the Tarski-Sher thesis in sec. 5 below will hinge directly on a restriction to absolute notions.

At first, critique III was for me perhaps the strongest reason for rejecting the Tarski-Sher thesis, at least as it stands. It seems to me there is a sense in which the usual operations of the first-order predicate calculus have the *same meaning* independent of the domain of individuals over which they are applied. This characteristic is *not* captured by invariance under bijections. As McGee puts it, "(t)he Tarski-Sher thesis does not require that there be any connections among the ways a logical operation acts on domains of different sizes. Thus, it would permit a logical connective which acts like disjunction when the size of the domain is an even successor cardinal, like conjunction when the size of the domain is an odd successor cardinal, and like a biconditional at limits." (McGee [21], p. 577)

In the end (though perhaps more for other reasons), McGee accepted the Tarski-Sher thesis as a necessary condition for an operation across domains to count as logical, but not a sufficient one. I agree completely, and believe that if there is to be an explication of the notion of a logical operation in set-theoretical/semantical terms, it has to be one which shows how the way an operation behaves when applied over one domain D connects naturally with how it behaves over any other domain D'. I made a first step in that direction in [11], where I proposed a notion of (strong) homomorphism invariance as a criterion for logicality of operations Q across domains. By a such a homomorphism $h: (D, \underline{R}) \rightarrow (D', \underline{R}')$ is meant one that is a map from D *onto* D' such that for each $i = 1,\ldots, n$ and each k_i-ary sequence $\underline{x}$ of individuals in D, and for $h(\underline{x})$ the corresponding sequence of h values in D', we have $R_i(\underline{x})$ iff $R_i'(h(\underline{x}))$. Immediately excluded by homomorphism invariance are the identity relation between individuals and all the cardinality quantifiers. This evidently brings us closer to first-order logic. Then a truth-valued operation Q across domains is said to be (strong) homomorphism invariant if whenever h is such a homomorphism then $Q(D, \underline{R}) = Q(D', \underline{R}')$. The paradigmatic homomorphism invariant operation is that of existential quantification, which is of monadic type. Note also that the truth-functional operations such as negation and conjunction preserve homomorphism invariance. In the following we shall write FOL for the first-order predicate calculus *with* equality, and FOL^- for the same *without* equality.

THEOREM 3. (Feferman [11]) The operations definable in FOL^- are exactly those λ-definable from homomorphism invariant operations of monadic type.

To explain the sense of λ-definability that is intended in this statement, consider for example the operation Q(P, R, S) defined for unary P and binary R and S in FOL^- by the sentence

$$\forall x[P(x) \rightarrow \exists y \exists z(R(x,y) \wedge S(x,z))],$$

which is equivalent to

$$\neg\exists x[P(x) \wedge \neg\exists y \exists z(R(x,y) \wedge S(x,z))].$$

Then its λ-definition is given in terms of the operations of negation (N), conjunction (C) and existential quantification (E) and the characteristic functions p, r, s of P, R, S respectively by $N(E(\lambda x[C(p(x),N(E(\lambda y E(\lambda z C(r(x,y),s(x,z)))))]))$.

The reader is referred to [11] for the proof of Theorem 3.

As is shown by the result of Denis Bonnay in the next section, homomorphism invariant operations in general go far beyond the first-order predicate calculus. For a simple example for the moment, consider the negation of the well-foundedness quantifier WF, i.e. the operation $Q_D(R)$ for binary R which holds in a given domain D just in case there exists a function $f: N \rightarrow D$ such that $\forall n[R(f(n+1), f(n))]$; that is homomorphism invariant.

Independently of such examples, one immediate criticism of the homomorphism-invariance criterion for logicality is that it excludes the identity relation, which is ordinarily counted as a part of FOL. Actually, that is a controversial matter. See, for example, the discussion by Quine of that question in his *Philosophy of Logic* (Quine [25], pp. 61 ff). On the one hand, he says that it "seems fitting" that the predicate of = is to be counted with predicates such as < and ∈ as part of mathematics and *not* of logic. On the other hand, he gives three arguments for counting = as part of logic. The first is the completeness of the logic of the first-order predicate calculus with equality , the second is the "universality" of =, and the third is the possibility of "simulating" = in a language L containing finitely many predicate symbols; by that he means its explicit definition from those predicates to satisfy the condition of identity of indiscernibles.

Finally, as pointed out to me by Bonnay, it is hard to see how identity could be determined to be logical or not by a set-theoretical invariance criterion of the sort considered here, since either it is presumed in the very notion of invariance itself that is employed—as is the case with invariance under isomorphism or one of the partial isomorphism relations considered in the next section—or it is eliminated from consideration as is the case with invariance under homomorphism.

4. Invariance with respect to similarity relations; Bonnay's work. As mentioned above, it has been shown by Bonnay that the operations in general that are homomorphism invariant go far beyond those definable in FOL. The result is stated in his paper [7], but a proof is not given there; instead the reader is referred back to his dissertation:

THEOREM 4. (Bonnay [6]) An operation Q across domains is invariant under homomorphisms iff it is definable in the language $L_{\infty,\infty}$ without equality.

The proof of this in [6] proceeds by a straightforward modification of McGee's proof of Theorem 1 using a detour via quotient structures. Moreover, for each choice of finitely many predicate symbols, this language is essentially of the same expressive power as full $L_{\infty,\infty}$, by means of Quine's method of simulating identity.

Bonnay has obtained further interesting results by consideration of a more general question: which operations across domains are S-invariant where S is a "similarity" relation $M \sim_S M'$ between structures $M = (D, \underline{R})$ and $M' = (D', \underline{R}')$ of the same signature? Basic examples of such are isomorphism and strong homomorphism as above. But

Bonnay also considers a number of others, including α-isomorphism and potential isomorphism, defined by Karp [15] as follows:

An *α-isomorphism* I from $M = (D, \underline{R})$ to $M' = (D', \underline{R}')$ is a sequence $I_0 \supseteq I_1 \supseteq \ldots \supseteq I_\beta \supseteq \ldots \supseteq I_\alpha$ such that (i) I_α is non-empty, (ii) for any $\beta \leq \alpha$, I_β is a set of partial isomorphisms f between these two structures with $dom(f) \subseteq D$ and $rng(f) \subseteq D'$, and (iii) if $\beta+1 \leq \alpha$ then for any $f \in I_{\beta+1}$ and x in D (resp. y in D′) there exists g in I_β with $f \subseteq g$ and $x \in dom(g)$ (resp. $y \in rng(g)$). We write $M \sim_\alpha M'$ if there exists such an α-isomorphism; the similarity relation in this case is denoted Iso_α.

A *potential isomorphism* I between $M = (D, \underline{R})$ and $M' = (D', \underline{R}')$ is a non-empty collection of partial isomorphisms such that for each $f \in I$ and $x \in D$ (resp. $y \in D$) there exists $g \in I$ with $f \subseteq g$ and $x \in dom(g)$ (resp. $y \in rng(g)$). We write $M \sim_p M'$ if there exists such an I, and the similarity relation in this case is denoted Iso_p.

The similarity relations are partially ordered by $S \leq S'$ iff $S' \subseteq S$. The smallest S w.r.t. $\leq$ is the universal relation Univ between structures of the same signature; where there are no constant symbols, this agrees with Iso_0. For any α, $Iso_0 \leq Iso_\alpha \leq Iso_p \leq Iso$, where Iso is the relation of being isomorphic; the strong homomorphism relation is incomparable with Iso_p. It is a familiar result due to Fraïssé [13] that two structures are elementarily equivalent in $L_{\omega,\omega}$ (= FOL) just in case they are in the Iso_ω relation. Karp [15] obtained analogous results for the languages $L_{\infty,\omega}$ whose formulas φ are generated by arbitrary conjunctions and disjunctions and closed under ordinary quantification, i.e. formation of $\forall x\varphi$ and $\exists x\varphi$ for any variable x. One defines the quantifier rank of φ, qr(φ), in a natural way. Then Karp's theorems are that for limit α, two structures are in the Iso_α relation if and only if they satisfy the same sentences φ for which $qr(\varphi) < \alpha$, and two structures are in the Iso_p relation if and only they satisfy the same sentences of $L_{\infty,\omega}$.

The class of operations Q across domains that are invariant under a given similarity relation S is denoted by Inv(S). Bonnay's main result characterizes the similarity relation Iso_p in two different ways in the $\leq$ relation. The first of these makes use of a natural additional criterion for logicality, namely that any operation definable from the operations in Inv(S) should already be invariant under S. We can explain this notion of definability by setting up a language L^S containing a generalized quantifier symbol **Q** for each Q invariant under S, with the semantics that interprets **Q** as Q in the way explained by Lindtröm [17]. Taking CInv(S) to consist of all the operations definable in L^S, Bonnay argues for the following:

Principle for Closure under Definability. CInv(S) = Inv(S).

This is a strong condition; for example S = Iso_ω fails to satisfy it. A counter-example is provided by the "infinitely many" quantifier: it is Iso_ω invariant but one construct a quantifier from it that is first order definable but which is not Iso_ω invariant. In fact, Bonnay's main result is the following.

THEOREM 5. (Bonnay [7]) Iso_p is the least S in the ≤ relation for which $Iso_1 \leq S$ and Cinv(S) = Inv(S).

Bonnay also defines an operation Sim dual to the operation Inv; the domain of Inv is the class of all similarity relations S between structures M and the domain of Sim is the class of all collections K of operations Q across structures. Inv maps the former to the latter, while Sim maps the latter to the former as follows: two structures M, M′ of the same signature are in the relation Sim(K) if for every Q in K, Q(M) = Q(M′). The classes K of operators across structures are ordered by K ≤ K′ iff K ⊆K′. It is shown that the class of all similarity relations S and the class of all classes of operators K form a Galois connection with respect to their respective orderings. In particular, for the class K of operators definable in $L_{\infty,\omega,}$, we have Sim(K) = Iso_p by Karp's theorem and CInv(Sim(K)) = Inv(Sim(K)) by Theorem 5. Note well that this does *not* tell us that Inv(Sim(K)) = K; for, the well-foundedness quantifier WF is Iso_p-invariant but not definable in $L_{\infty,\omega}$ ([15]).

In further favor of Iso_p as a distinguished similarity relation, Bonnay quotes the following characterization of it due to Barwise in terms of the notion of absoluteness, to be discussed at length in the next section.

THEOREM 6. (Barwise [4]) Iso_p is the greatest S in the ≤ relation that is absolute with respect to ZFC and for which ZFC proves that S ≤ Iso.[4]

Theorems 5 and 6 together lead Bonnay to state the following:

> Iso_p THESIS FOR LOGICALITY. An operator Q is logical iff Q is Iso_p-invariant. (Bonnay [7], p.61).

We shall discuss Bonnay's arguments for this thesis at length in the final section below. But I'd like here to look at his approach via similarity relations from a different angle. Historically speaking, one started with natural logics L like $L_{\omega,\omega}$ and $L_{\infty,\omega}$ and asked for a mathematical characterization of elementary equivalence with respect to such L, the results being given in these particular cases by the work of Fraïssé and Karp via the similarity relations Iso_ω and Iso_p, respectively. In each case, we could ask of the given L: *if each operation in L is to be counted as logical, what else ought to be counted as logical*? Let K_L be the class of operations defined in L. The first thought from a similarity invariance point of view is to count as logical all those operations invariant under $Sim(K_L)$, i.e. the operators in $Inv(Sim(K_L))$. But in each of the two specific cases, as we have seen, that takes us beyond the given logic. In the case of $L_{\omega,\omega}$ an example is provided by the "infinitely many" quantifier, and in the case of $L_{\infty,\omega}$ that is provided by the "well-foundedness" quantifier. But if one asks whether an *individual* operator Q ought to be counted as logical, given that *each* operator in L is counted as logical, we are asking that it pass too strong a test. For we could say of each operator defined in L that the reason *it* is logical is based on a more refined invariance condition than that of being invariant w.r.t. the associated similarity relation $Sim(K_L)$. And then we should not require of a new operation Q under consideration to pass anything stronger than such a refined condition. Specifically, in the case of FOL, the following theorem tells us that there is no reason to count as logical anything stronger than what is counted as logical anything that is not Iso_n-invariant for some n.

THEOREM 7. Q is definable in FOL if and only if there exists $n < \omega$ such that Q is Iso_n-invariant.

Proof. In the forward direction, one shows as usual that if Q is definable by a sentence φ of FOL with $qr(\varphi) < n$, then Q is Iso_n-invariant. The idea of the proof in the converse direction stems from Fraïssé [13]. First, one shows that for each n and M there is a sentence $\chi_n(M)$ of FOL such that for all M′, $M \sim_n M'$ iff $M' \vDash \chi_n(M)$. Moreover, the set Typ_n of all sentences $\chi_n(M)$ is finite. Let $Typ_n = \{\chi_n(M_1),\ldots, \chi_n(M_k)\}$ for suitable $M_1, \ldots, M_k$. Then given an operation Q and an n such that whenever $M \sim_n M'$ then $Q(M) = Q(M')$, we can take the disjunction of those (and only those) $\chi_n(M_i)$ for which $Q(M_i) = T$ as the sentence that defines Q in FOL.

A similar result can be stated for $L_{\infty,\omega}$ using the work of Karp [15].

5. Adding absoluteness criteria to isomorphism invariance. To return to the central question, let's look in more detail at the *absoluteness criterion* suggested by my critique II of the Tarski-Sher thesis. Let T be a set of axioms in the language of set theory. A formula φ of set theory is defined to be *absolute w.r.t. to* T if φ is invariant under end-extensions for models of T. It was proved some time ago by Kreisel and me, as strengthened in Feferman [9], that φ is absolute w.r.t. T iff it is Δ rel. to T, i.e. it is provably equivalent to both a Σ and a Π formula rel. to S where here by Σ (Π) is meant the class of formulas in prenex form in which all unbounded quantifiers are existential (universal). Note well that the notion of being absolute is relative to a system of axioms. For his proof of the consistency of AC and GCH relative to ZF, Gödel needed to show that a number of notions are absolute relative to that system. It was since established that all those notions are absolute relative to Kripke-Platek set theory KP; below it will be more useful to deal with the slightly weaker system KPU, which allows urelements, and if we speak of absoluteness without explicit reference to a system of axioms, one means relative to KPU. This system includes the Axiom of Infinity, Inf, in the form that guarantees the existence of ω; also, in it (even without Inf) every Σ (Π) formula is equivalent to a Σ_1 (Π_1) formula. Among the notions that are absolute w.r.t. KPU are being an ordinal, being ω, and being a formula of FOL true in a structure M. Among those that are not absolute are being an uncountable ordinal, being ω_1, and being the power set of ω.

For the determination of which operators across domains ought to be counted as logical on the basis of certain absoluteness invariance criteria, it turns out one can make use of results about absolute logics within the framework of abstract model-theory. The general background is explained in the chapters by Ebbinghaus [8] and Flum [12] of the volume *Model-Theoretic Logics* ([5]). For our purposes, an abstract logic is determined by specifying for each signature σ of type level 2 a set $Sent_\sigma$ of "sentences" and a relation $M \models \varphi$ between structures of signature σ and members of $Sent_\sigma$, satisfying certain regularity conditions. A class C of structures of a given signature σ is said to be an *elementary class for* L, if for some $\varphi \in Sent_\sigma$, C consists of all M for which $M \models \varphi$. Logics are ordered by the relation $L \leq L'$ which holds when every class C that is elementary for L is also elementary for L′. Examples of logics that we shall consider below are $L_{\omega,\omega}$ (= FOL), $L_{\infty,\omega}$, and $L_{\infty,\infty}$. The regularity conditions usually assumed on a logic L insure that $L_{\omega,\omega} \leq L$. Lindström's famous theorem [18] characterizes $L_{\omega,\omega}$ as the largest logic satisfying the compactness theorem and the Löwenheim-Skolem theorem; he also showed that it is the largest logic such that the set of L-valid sentences is recursively enumerable and that satisfies the Löwenheim-Skolem theorem. Another relevant theorem from the same paper characterizes first-order logic as the largest logic that satisfies the Löwenheim-Skolem-Tarski theorem, in other words no sentence of the logic can have models in just one infinite cardinal. That is a generalization of a result of Mostowski for his cardinality quantifiers.

A logic L is said to be absolute if the sets $Sent_\sigma$ and the $\models$ relation for L are absolute. Barwise [3] initiated the study of absolute logics with his proof that $L_{\infty,\omega}$ is the largest logic which is absolute for KP if no restriction is made as to the sets $Sent_\sigma$.[5] The subject of set-theoretic definability of logics and in particular of absolute logics was extensively surveyed and considerably advanced in the chapter by Väänänen [30] in [5]. A number of further results have been obtained in the 1995 thesis [1] of Väänänen's student Jyrki Akkanen. A natural question to ask after Barwise's result is whether FOL can be characterized by more refined absoluteness criteria than that of [3]. Indeed, in an unpublished manuscript dated 1979, Ken Manders proved the following:

THEOREM 9. (Manders [19]) $L_{\omega,\omega}$ is the largest logic L that is absolute relative to KPU–Inf, whose set of sentences is contained in the hereditarily finite sets HF and whose structures M = (D, R̲) have domains D consisting only of urelements.

A published proof of Theorem 9 is to be found in [30], pp. 620-622, though the result there (3.1.5) is incorrectly stated for KP–Inf instead of KPU–Inf.[6] Väänänen's proof of this theorem is different from Manders' in that it makes essential use of my notion of *adequacy to truth* of the notion of one logic L being adequate to truth for another language L′ [10]. Roughly speaking what this means is that the satisfaction relation for L′ for all subformulas of any given formula of L′ is, in a suitable sense, uniformly implicitly invariantly definable in L. L is said to be *truth maximal* if whenever it is adequate to truth in L′ we have L′ ≤ L. The main results in [10] for that notion were that a logic is truth-maximal iff it has the Δ-interpolation property, and that $L_{\omega,\omega}$ is truth-maximal among all logics whose sentences are represented in HF. The crucial step in Väänänen's proof is to push back being absolute w.r.t. KPU–Inf to the Δ-interpolation property.

Relative to any set S of axioms in the language of KPU, an operation Q across domains is said to be *absolute* if the relation between D and R̲ such that D is a set of urelements and $Q_D(\underline{R}) = T$ is absolute. When Q is preserved under isomorphism, it serves to determine a Lindström quantifier in the sense of [17]. Then we can formally extend the language of FOL by a symbol **Q** for Q, with its semantics determined by Q.

LEMMA 10. If Q is absolute w.r.t. an extension S of KPU–Inf then the logic $L = L_{\omega,\omega}(\mathbf{Q})$, obtained by adjoining **Q** to $L_{\omega,\omega}$, is also absolute w.r.t. S.

This is easily seen by the fact that the satisfaction relation for L among subformulas of any given formula is Δ in Q w.r.t. S, and that being Δ in Δ definable is equivalent to being Δ definable.

THEOREM 11. If an operation Q across domains is isomorphism invariant and is absolute w.r.t. KPU–Inf then Q is definable in $L_{\omega,\omega}$.

By Lemma 10, this is a corollary of Manders' Theorem 9.

CONJECTURE. If an operation Q across domains is homomorphism invariant and is absolute w.r.t. KPU–Inf then Q is definable in FOL^{-}.

6. Discussion. Bonnay [7] presents an interesting analysis of the informal arguments for various set-theoretical invariance criteria for logicality. He formulates the first such, for Tarski's thesis, in terms of the idea of levels of generality. In the Klein *Erlanger Programm*, levels of generality of a geometry are distinguished by the levels of generality of the associated groups of transformations. Thus, e.g., affine geometry is more general than Euclidean geometry since the affine transformations are more general than the isometric transformations (as well as the more general similarity transformations). Continuing in this vein leads one to explaining logic, which is the most general theory of all, in terms of the largest group of transformations, namely the class of permutations on any given domain, and to the identification of the logical notions with those invariant under permutations of the underlying universe. More explicitly as given by Bonnay, the *generality argument* for Tarski's thesis runs as follows.

> G.1 The distinctive feature of logic among other theories is that it is the most general theory one can think of.
> G.2 The bigger the group of transformations associated with a theory, the more general the theory.
> G.3 The biggest group of transformations is the class of all permutations.
> [Hence]

The logical notions are notions invariant under permutation. ([7] p.33)

By contrast, Bonnay analyzes the informal case made by Sher and others for the permutation invariance criterion in terms of what he calls the *formality argument*, which runs as follows:

F.1 Logic deals with formal notions, as opposed to non formal ones.
F.2 Formal notions are those which are insensitive to arbitrary switching of objects.
F.3 A notion is insensitive to arbitrary switching of objects iff it is invariant under permutation.
[Hence,]
The logical notions are the notions invariant under permutation. ([7] p.34)

Bonnay rejects the Tarski-Sher thesis on the grounds that it overgenerates, for reasons along the lines of my Critique I, and more specifically because it counts as logical any isomorphism invariant mathematical notion. At the conclusion of [7] he tries to make a case instead for his Iso_p thesis, via a pair of informal arguments modifying the preceding. The first is what he calls the *mild generality argument*, that runs as follows:

MG.1 Logic deals with very general notions, but not only with trivial notions.
MG.2 The truth-functions, functional application and first-order existential quantification are logical operators.
MG.3 The good notion of invariance for logicality is to be provided by a similarity relation S such that S is closed under definability.
MG.4 The good notion of invariance for logicality is to be provided by the lowest similarity relation compatible with MG.2 and MG.3.
[Hence}
The logical notions are the Iso_p-invariant notions. ([7], p. 59)

Bonnay's reasoning is that the conclusion follows from MG.1-MG.4 by means of his main result stated as Theorem 5 above. And in place of the formality argument, he proposes the following *lack of content argument*, to reflect the idea that "logical notions should not encapsulate any problematic set-theoretical content":

> LC.1 Logic deals with notions which are deprived of non formal content and of problematic set-theoretic contents.
> LC.2 The good notion of invariance for logicality is to be provided by a similarity relation S such that $S \leq Iso$.
> LC.3 The good notion of invariance for logicality is to be provided by a similarity relation S such that S is absolute with respect to ZFC.
> LC.4 The good notion of invariance for logicality is to be provided by the greatest similarity relation S satisfying LC.2 and LC.3.
> [Hence]
> The logical notions are the Iso_p-invariant notions. ([7], p. 60)

In this case, the reasoning is supported by Barwise's Theorem 6 above.

Bonnay returns to the overgeneration problem as a challenge to the Iso_p thesis for logicality in his final subsection (4.3). Though cardinality quantifiers like $\exists_{\geq\kappa}$ for κ an uncountable cardinal are not logical on this thesis, the quantifier "there exist infinitely many" is. Thence, as Bonnay acknowledges, all arithmetical truths count as logical truths, and "the overgeneration problem is at least eased, if not solved, by the shift from Iso invariance to Iso_p-invariance" ([7], p.65).[7]

I don't find either of the modified arguments—mild generality and lack of content—convincing even with the supporting theorems, and certainly not as compelling on the face of it as the generality and formality arguments for permutation invariance as the criterion for logicality. For one thing, the presumption in both arguments is that invariance is to be expressed in terms of a single, global (or "coarse-grained") similarity relation.[8] But Theorem 7 above uses invariance instead with respect to what might be called a collection of local similarity relations, and reaches the much different conclusion that the logical notions are just those definable in FOL. And even if one accepts that invariance is to be given by a single, global similarity relation, it seems to me equally plausible to substitute for LC. 3 the following:

> LC.3′ The good notion of invariance for logicality is to be provided by a similarity relation S such that S is absolute with respect to KPU−Inf,

since much more so than absoluteness w.r.t. ZFC (and even more so than w.r.t. KP), absoluteness w.r.t. KPU−Inf guarantees that one does "not encapsulate any problematic set-theoretical content." (Not that the Infinity Axiom is mathematically problematic; rather it is problematic as an assumption in the explanation of what counts as a logical notion.) My guess would be that if one substitutes LC.3′ for LC.3, one would be led (in analogy to Barwise's Theorem 6 above) to the conclusion that the logical notions are just those invariant under Iso_{ω}, thus bringing us closer to FOL as given by Theorem 11. But beyond meeting my Critiques I and II of the Tarski-Sher thesis, which have also been Bonnay's motivations, Theorem 11 was mainly designed to meet Critique III, namely that the criterion of isomorphism invariance does not explain what it means to be the same logical operation for domains of different size. To be sure, the result of Theorem 11 still does not insure sameness of meaning, since we can define an operator in FOL by means of a sentence which has one semantics on domains, say, of ≤ 5 elements and another on domains of ≥ 6 elements. Similar examples can be provided in FOL^{-}, so this is not an issue that depends in any essential way on whether identity is taken to be a logical notion. For either case, a better explanation is needed of what constitutes sameness of meaning across domains if Critique III is to be dealt with in any way beyond what is done here.

Coming back to the Critique II: by requiring of the definition of Q that it be absolute relative to a weak set theory without the axiom of infinity, we are insuring that its meaning does not depend on any special set-theoretical assumptions about what exists beyond the most elementary set-constructions that generate HF from any set of urelements. i.e. it rests on just what is needed for a theory of the syntax of any humanly manageable system of logical reasoning.

This last connects with the completely different program to characterize logical notions in terms of rules of inference that implicitly determine them; that was initiated by

Gerhard Gentzen and has subsequently been pursued by Dag Prawitz, Per Martin-Löf, Ian Hacking, Kosta Došen and Jeffery Zucker among others; cf. [11] sec. 6.5 for references. And that returns us to the traditional conception of logic as the study of the *forms of correct reasoning*, *of what invariably leads from truths to truths*. Despite the various appealing results above, and despite my personal feeling that the logical operations do not go beyond those represented in FOL, I do not find the various arguments for logicality based on any of the invariance notions considered here convincing in their own right. In my view, the semantical and syntactic (inferential-theoretic) approaches are complementary to each other, and a proper explanation of what are logical notions and of what is logic—if there is to be one—will have to take both into account. In the direction of a characterization of the logical notions that does just that, consider, by way of conclusion, the following result.

THEOREM 12. Suppose Q is an operation across domains that is

(i) isomorphism invariant

(ii) absolute w.r.t. KP

(iii) and is such that the set of valid sentences of $L_{\omega,\omega}(Q)$ is recursively enumerable.

Then Q is definable in FOL.

Proof. By Lemma 10 and (ii), the logic $L_{\omega,\omega}(Q)$ is absolute w.r.t. KP, and of course its sentences are representable in HF. Then by Theorem 3.2 of Barwise [3] (p.325), the logic $L_{\omega,\omega}(Q)$ is contained in L_A, where A is the least admissible set that contains ω, namely the constructible sets below the least non-recursive ordinal. That language L_A satisfies the Löwenheim-Skolem theorem, hence so also does $L_{\omega,\omega}(Q)$. But then by Lindström [18], $L_{\omega,\omega}(Q)$ is contained in $L_{\omega,\omega}$.

Note that condition (ii) is more robust on the set-theoretical side than absoluteness w.r.t. KPU−Inf as assumed in Theorem 11. Re condition (iii), it is plausible to assume of any system of human logical reasoning, that its sentences are represented in HF and that it makes use of some finite set of effective rules. It follows that the totality of sentences that can be shown to be valid in the given logic constitutes a recursively enumerable set.

Of course it does *not* follow from that that (iii) must hold, since there is no guarantee that any such system of rules for the semantics that is determined by the given Q is complete.

Acknowledgements. I wish to thank Denis Bonnay, Dag Westerståhl and Jouko Väänänen for their helpful comments on a draft of this paper.

REFERENCES

[1] Addison, J. W., L. A. Henkin and A. Tarski, eds., *The Theory of Models*, North-Holland Publishing Company, Amsterdam, 1965.

[2] Akkanen, J., "Absolute logics," *Annales Academiæ Scientiarum Fennicæ, Series A, I. Mathematica Dissertationes*, vol. 100, 1995.

[3] Barwise, J., "Absolute logics and $L_{\infty,\omega}$," *Annals of Pure and Applied Logic*, vol. 4 (1972), pp. 309-40.

[4] Barwise, J., "Back and forth through infinitary logic," pp. 5-34 in *Studies in Model Theory*, The Mathematical Association of America, 1973.

[5] Barwise, J. and S. Feferman, eds., *Model-theoretic Logics*, Springer-Verlag, Berlin, 1985.

[6] Bonnay, D., *Qu'est-ce qu'une constante logique?*, Ph. D. Dissertation, Université Paris I, 2006.

[7] Bonnay, D., "Logicality and invariance," *The Bulletin of Symbolic Logic*, vol. 14 (2008), 29-68.

[8] Ebbinghaus, H. D., "Extended logics: The general framework," pp. 25-76 in *Model-theoretic Logics*, Springer-Verlag, Berlin, 1985.

[9] Feferman, S., "Persistent and invariant formulas for outer extensions," *Compositio Mathematica*, vol. 20 (1968), pp. 29-52.

[10] Feferman, S., "Applications of many-sorted interpolation theorems," pp. 205-23 in *Proceedings of the Tarski Symposium*, American Mathematical Society, Providence, 1974.

[11] Feferman, S., "Logic, logics and logicism," *Notre Dame Journal of Formal Logic*, vol. 40 (1999), pp. 31-54.

[12] Flum, J. "Characterizing logics," pp. 77-120 in *Model-theoretic Logics*, Springer-Verlag, Berlin, 1985.

[13] Fraïssé, R., "Sur quelques classifications des systèmes de relations," *Université d'Alger, Publications Scientifique, Série A,* vol. 1 (1954), 35-182.

[14] Henkin, L., et al., eds., *Proceedings of the Tarski Symposium*, American Mathematical Society, Providence, 1974.

[15] Karp, C., "Finite quantifier equivalence," pp. 407-12 in *The Theory of Models*, North-Holland Publishing Company, Amsterdam, 1965.

[16] Lindenbaum, A. and A. Tarski, "On the limitation of the means of expression of deductive theories," English translation by J. H. Woodger, pp. 384-92 in *Logic, Semantics, Metamathematics*, 2d edition, edited by J. Corcoran, Hackett, Indianapolis, 1983.

[17] Lindström, P., "First order predicate logic with generalized quantifiers," *Theoria*, vol. 32 (1966), pp. 175-71.

[18] Lindstrom, P., "On extensions of elementary logic," *Theoria*, vol. 35 (1969), pp. 1-11.

[19] Manders, K., "First-order logical systems and set-theoretic definability", Manuscript, University of Pittsburgh, 1979.

[20] Mautner, F. I., "An extension of Klein's *Erlanger Programm*: logic as invariant theory," *American Journal of Mathematics*, vol. 68 (1946), pp. 345-84.

[21] McGee, V., "Logical operations," *Journal of Philosophical Logic*, vol. 25 (1996), pp. 567-80.

[22] Morley, M. D., ed., *Studies in Model Theory*, The Mathematical Association of America, 1973.

[23] Mostowski, A., "On a generalization of quantifiers," *Fundamenta Mathematicae*, vol. 44 (1957), pp. 12-36.

[24] Peters, S. and D. Westerståhl, *Quantifiers in Language and Logic*, Clarendon Press, Oxford, 2006.

[25] Quine, W. V. O., *Philosophy of Logic*, 2d edition, Harvard University Press, Cambridge, 1986.

[26] Sher, G., *The Bounds of Logic,* The MIT Press, Cambridge, 1991.

[27] Tarski, A., "On the concept of logical consequence," English translation by J. H. Woodger, pp. 409-420 in *Logic, Semantics, Metamathematics,* 2d edition, edited by J. Corcoran, Hackett, Indianapolis, 1983.

[28] Tarski, A., *Logic, Semantics, Metamathematics,* 2d edition, edited by J. Corcoran, Hackett, Indianapolis, 1983.

[29] Tarski, A., "What are logical notions?," *History and Philosophy of Logic*, vol. 7 (1986), pp. 143-154.

[30] Väänänen, J., "Set-theoretic definability of logics," pp. 599-643 in *Model-theoretic Logics*, Springer-Verlag, Berlin, 1985.

Department of Mathematics

Stanford University

Stanford, CA 94305-2125

Email: feferman@stanford.edu

---

[1] The material for this article is drawn from the second of three Tarski Lectures that I gave at the University of California at Berkeley during the week of April 3, 2006, this one under the title: "The 'logic' question." It has not previously been published.

[2] In [11] I also allowed individuals and truth values as arguments to an operation Q over any given domain; the restriction here to n-ary relations as arguments (n > 0) is taken for simplicity.

[3] Bonnay [7], 1.2, points out that any mathematical notion in the form of a class K of structures (D, $\underline{R}$) of a given signature that is closed under isomorphism determines a logical notion Q in the Tarski-Sher sense by Q(D, $\underline{R}$) = T iff (D, $\underline{R}$) $\in$ K.

[4] In fact, Iso_p is absolute w.r.t. KP by [3].

[5] There are larger logics that are absolute w.r.t. stronger systems such as $L_{\infty,\omega}$(WF) and its further extension by the "game quantifier."

[6] Akkanen [1] pointed out that the "infinitely many" quantifier is absolute w.r.t. KP–Inf. I recently asked Väänänen what the problem is with his proof of 3.1.5 in [30], and he replied that it only works if one is dealing with L-structures M = (D, $\underline{R}$) for which D is a set of urelements.

[7] In defense of the Iso_p thesis, Bonnay calls on natural language use to support the logicality of arithmetic notions, as well as the quantifiers "infinitely many" and "most" (the latter only for countable structures), [7] pp. 64-65. On the face of it this seems at odds with his questioning my appeal to natural language use in support of the homomorphism invariance criterion via reduction to monadic quantifiers ([7], p. 44); however, there are independent considerations for each. For a comprehensive treatment of quantifiers in natural language and logic see Peters and Westerstahl [24].

[8] Denis Bonnay has pointed out that this criticism also applies to the original generality argument, since it is a hidden assumption there that one is dealing with a global similarity relation, rather than a family of such relations.