Universal Logic III
Estoril 2010
Tutorial on Truth Values
Fabien Schang and Heinrich Wansing
Lecture 1

Heinrich Wansing

Abstract

In this lecture, it is explained how and why Frege's notion of a truth value has become part of the standard philosophical and logical terminology. Moreover, various applications of the notion of a truth value in logic and philosophy are presented and discussed. These topics include the famous slingshot argument, Suszko's Thesis, the notion of a many-valued logic and multi-lattices of generalized truth values.

The notion of a truth value has been introduced into logic by Gottlob Frege in 1891 and 1892 in the papers 'Function und Begriff' and 'Über Sinn und Bedeutung'. In the posthumously published 'Einleitung in die Logik' he explains that

> *A sentence proper is a proper name, and its Bedeutung, if it has one, is a truth-value: the True or the False.*

Being saturated expressions, for Frege declarative sentences refer to objects of special kind, namely truth values.

This new and revolutionary idea has had a far reaching impact on the development of modern logic. In the context of a functional analysis of languages, the notion of a function could be generalized by introducing a special kind of functions, namely propositional functions, whose range of values consists of the set of truth values.

Truth values clearly have something to do with the concept of truth. Therefore it may seem rather tempting to try to incorporate considerations on truth values into the broader context of traditional truth-theories, such as correspondence, coherence, or even anti-realistic, or pragmatist conceptions of truth.

Nevertheless, it is unlikely that attempts at incorporating considerations on truth values as objects into the broader context of traditional theories of truth can give rise to any considerable success.

Indeed, Frege's truth values do not commit us to any specific metaphysical doctrine of truth. However, the idea of truth values contravenes traditional approaches to truth by bringing to the fore the problem of its categorial classification.

Is truth an object or is it a property?

It seems natural and it is quite customary to assume that truth is a property of sentences, propositions, or beliefs.

But Frege himself thought that characterizing a sentence as "true" adds nothing new to its content, for "It is true that 5 is a prime number" says exactly the same as just "5 is a prime number". This idea gave an impetus to the deflationary conception of truth (advocated by Ramsey, Ayer, Quine, Horwich, and others).

Although Frege admits the redundancy of truth as a property, he emphasizes its importance and indispensable role in some other respect. Namely, truth, accompanying every act of judgment as its ultimate goal, secures an objective value of cognition by arranging for every assertive sentence a transition from the level of sense (the thought expressed by a sentence) to the level of denotation (its truth value).

This conception highlights the significance of taking truth as a particular object.

As Tyler Burge explains:

> *Normally, the point of using sentences, what "matters to us", is to claim truth for a thought. The object, in the sense of the point or objective, of sentence use was truth. It is illuminating therefore to see truth as an object (Burge 1986, p. 129).*

Another difficulty with interpreting truth as a property concerns the question of what kind exactly are the entities that can possess this property? Are these sentences, propositions, beliefs, thoughts or anything else?

If we treat truth as a property, then any concrete decision on this problem becomes crucial, for it would determine the very nature of truth, because properties of sentences clearly are of an essentially different conceptual nature than, say, properties of beliefs.

One advantage of thinking of truth as an object is that we escape the problem of addressing the problematic question of truth bearers, because it is hardly disputable that one and the same truth value can be correlated with different sorts of things—not only with a sentence, but also with the corresponding proposition or the belief with this propositional content.

Of course, it is always possible to maintain a truth predicate in a metalanguage (à la Tarski), and to keep using a more customary wording, by stipulating that, e.g., a sentence is true if and only if it designates the truth value "the true". Such a linguistic convention, understood as a mere abbreviation, would be rather harmless.

Moreover, it has been observed repeatedly in the literature that the stress Frege laid on the notion of a truth value was, to a great extent, "pragmatically" motivated.

On the one hand, Frege expected a gain for his system of "Basic Laws" reflected in enhanced technical clarity, simplicity, and unity.

On the other hand, Frege sought to substantiate in this way his view on logic as a normative discipline with truth as its main goal and primary subject-matter. Frege's "the true" is not just a functional value, it is also a value in the sense of being something to be attained.

G. Gabriel (1986) showed that Frege's ideas can be linked up with a value-theoretical tradition in German philosophy of the second half of the 19th century.

Wilhelm Windelband, the founder and the principal representative of the Southwest school of Neo-Kantianism, was actually the first who employed the term "truth value" ("Wahrheitswert") as early as in 1884, even if he was very far from treating a truth value as a value of a function.

Windelband also considered the triad of basic values: "true", "good", and "beautiful" which was later taken up by Frege (1918) when he defined the subject-matter of logic. This connection between logic and a value theory can be traced back to Hermann Lotze, whose seminars in Göttingen were attended by both Windelband and Frege.

The decisive move made by Frege was to bring together a philosophical and a mathematical understanding of values on the basis of a generalization of the notion of a function on numbers.

If predicates are construed as a kind of functional expressions which, being applied to singular terms as arguments, produce sentences, then the values of the corresponding functions must be references (denotations, designations) of sentences.

Taking into account that the range of any function typically consists of objects, one naturally concludes that references (denotations, designations) of sentences must be objects as well.

And if we assume that sentences refer to truth values, then it turns out that truth values are indeed objects.

But do sentences designate truth values?

There is a famous argument (more precisely, a family of arguments) that is designed to show that all true sentences designate (denote, refer to) the same thing, as well as all false sentences do. These things are precisely the truth values: the true and the false.

The argument is anticipated (implicitly at least) by Frege (1892), and it was first formulated explicitly by Alonzo Church in his 1943 review of Carnap's "Introduction to Semantics". In "Introduction to Mathematical Logic" (1956) Church reconstructs the point of his proof by means of a rather informal line of reasoning. Other remarkable versions of the argument are those by Kurt Gödel (1944) and Donald Davidson (1967), which make use of the formal apparatus of a theory of descriptions.

Jon Barwise and John Perry (1981) called this family of arguments "the slingshot", stressing thus its extraordinary simplicity and the minimality of presuppositions involved. Versions of the slingshot argument have been analyzed by many authors, in particular by Stephen Neale (2001).

Stated generally, the pattern of the argument goes as follows. One starts with a certain sentence, and then moves, step by step, to a completely different sentence. Every two sentences in any step designate presumably one and the same thing. Hence, the starting and the concluding sentences of the argument must have the same designation as well. But the only thing they have in common seems to be their truth value.

Thus, what any sentence designates is just its truth value.

Gödel's slingshot

Gödel (1944) hints at a "rigorous proof" of the claim that "all true sentences have the same signification (as well as all false ones)" by making use of some assumptions.

Let $(\iota x)(x = a \wedge Fx)$ stand for the definite description "the x such that x is identical to a and x is F", and let for any sentence X, $[X]$ stand for what X designates. Then Gödel's assumptions can be articulated as follows:

(A1) $[Fa] = [a = (\iota x)(x = a \wedge Fx)]$.

(A2) Every sentence can be transformed into an equivalent sentence of the form Fa. (This assumption allows Gödel to expand his argument beyond the atomic sentences).

We will reconstruct Gödel's proof essentially as in (Neale 1995), although our formulation is somewhat different. To do this, we will need an introduction rule for the ι-operator:

$$\iota\text{-INTR:} \quad \frac{\Sigma(x/\alpha)}{\alpha = (\iota x)(x = \alpha \wedge \Sigma(x))}$$

where α is a singular term, $\Sigma(x)$ is a sentence containing at least one free occurrence of the variable x, and $\Sigma(x/\alpha)$ is the result of replacing every occurrence of x in $\Sigma(x)$ by α.

Other inference rules, which allow substitution for definite descriptions, are also taken from (Neale 1995):

$$\iota\text{-SUB:}\quad \frac{\begin{array}{c}(\iota x)\phi = (\iota x)\psi \\ \Sigma((\iota x)\phi)\end{array}}{\Sigma((\iota x)\psi)} \qquad \frac{\begin{array}{c}(\iota x)\phi = \alpha \\ \Sigma((\iota x)\phi)\end{array}}{\Sigma(\alpha)} \qquad \frac{\begin{array}{c}(\iota x)\phi = \alpha \\ \Sigma(\alpha)\end{array}}{\Sigma((\iota x)\psi).}$$

Suppose now the sentences G1–G3 below are true.

G1. Fa

G2. $a \neq b$

G3. Gb

Then one can proceed as follows:

G4. $a = (\iota x)(x = a \wedge Fx)$ G1, ι-INTR

G5. $a = (\iota x)(x = a \wedge x \neq b)$ G2, ι-INTR

G6. $b = (\iota x)(x = b \wedge Gx)$ G3, ι-INTR

G7. $b = (\iota x)(x = b \wedge a \neq x)$ G2, ι-INTR

G8. $(\iota x)(x = a \wedge Fx) = (\iota x)(x = a \wedge x \neq b)$ G4, G5, ι-SUB

G9. $(\iota x)(x = b \wedge Gx) = (\iota x)(x = b \wedge a \neq x)$ G6, G7, ι-SUB

G10. $[Fa] = [a = (\iota x)(x = a \wedge Fx)]$ A1

G11. $[a \neq b] = [a = (\iota x)(x = a \wedge x \neq b)]$ A1

G12. $[Fa] = [a = (\iota x)(x = a \wedge x \neq b)]$ G8, G10, ι-SUB

G13. $[Fa] = [a \neq b]$ G11, G12, *Transitivity of Identity*

G14. $[Gb] = [b = (\iota x)(x = b \wedge Gx)]$ A1

G15. $[a \neq b] = [b = (\iota x)(x = b \wedge a \neq x)]$ A1

G16. $[Gb] = [b = (\iota x)(x = b \wedge a \neq x)]$ G9, G14, ι-SUB

G17. $[Gb] = [a \neq b]$ G15, G16, *Transitivity of Identity*

G18. $[Fa] = [Gb]$ G13, G17, *Transitivity of Identity*

Instead of G2 we may also assume $a = b$. In any case Fa and Gb must have the same designation. But Fa and Gb may be completely different sentences having nothing (of semantic relevance) in common, except that they are both true, as assumed.

Russell held that any true sentence stands for a fact. In this case the argument above would say that all true sentences stand for one and the same fact, reducing thus Russell's view to an absurdity.

The argument can be equally used as a collapsing argument showing that sentences do not designate situations, states of affairs or anything of the sort, leading any attempt to assume so to a breakdown of the class of supposed designata "into a class of just two entities (which might as well be called "Truth" and "Falsity")" (Neale 1995). Another famous argument of this kind is the one by W.V. Quine (1953, 1960), by which he intended to demonstrate that quantifying into modal contexts leads to a collapse of modality.

By and large there are two ways of dismissing the slingshot: either to challenge at least one of its assumptions, or to examine the underlying theory of descriptions.

The assumptions most often called into question are the principle of co-referentiality of logically equivalent sentences (in Davidson's version of the argument) and the principle of substitutivity of co-referential singular terms.

However, objections against these assumptions typically look hardly more evident than the assumptions they question, and thus it is not so obvious why we should reject these principles rather than their denial.

Concerning descriptions, Gödel (1944) already drew the conclusion that if one assumes Russell's theory, in which uses of definite descriptions in a sentences are analyzed as giving rise to existential claims, then the argument can be blocked.

Yet, if we wish to maintain descriptions as denoting terms, the slingshot apparently hits its target.

Anna Wójtowicz (2005) attempts to reassess the slingshot argument by translating it into the formal language of Roman Suszko's (1975) non-Fregean logic extended by the ι-operator (or the λ-operator, depending on the version of the slingshot argument under consideration).

In addition to the vocabulary of classical first-order logic, the language of non-Fregean logic contains a binary connective $\equiv$. Intuitively, a formula $A \equiv B$ states that the sentences A and B denote (or describe) the same situation. (Using the notation above this can be expressed as $A \equiv B$ iff $[A] = [B]$.)

The so-called Frege Axiom then takes the form $(A \leftrightarrow B) \rightarrow (A \equiv B)$ (if A and B are materially equivalent, then they describe the same situation). Wójtowicz claims to show that any suitably extended non-Fregean predicate logic, which allows one to formalize the slingshot argument and its underlying assumptions, validates Frege's Axiom. This might be taken, Wójtowicz concludes, as evidence for the circularity of the slingshot argument.

To sum up, despite a manifold and occasionally quite sophisticated criticism of the argument developed by Church, Gödel, Davidson, and others, it seems to present a powerful and clear rationale for the view that truth values do exist, acting as designata for sentences.

Nevertheless, this does not at all mean that an ontology and semantics of situations (facts, states of affairs, etc.) is not worthy of investigation or even technically infeasible.

If we accept truth values and regard them as a special kind of entities, the question as to the nature of these entities arises.

The above characterization of truth values as objects is far too general, and one way of being slightly more specific is to qualify truth values as abstract objects.

Note that Frege himself never used the word "abstract" when describing truth values. Instead, he has a conception of so called "logical objects", truth values being the most fundamental (and primary) of them (Frege 1976). Among the other logical objects Frege pays particular attention to are sets and numbers, emphasizing thus their logical nature (in accordance with his logicist view).

Church (1956), when considering truth values, explicitly attributes to them the property of being "abstract". Since then it is customary to label truth values as abstract objects, thus allocating them into the same category of entities as mathematical objects (numbers, classes, geometrical figures) and propositions.

Finding an adequate definition of abstract objects is a matter of considerable controversy. According to a common view, abstract entities lack spatiotemporal properties and relations, as opposed to concrete objects which exist in space and time (Lowe 1995).

This view is frequently accompanied with a typical objection to the effect that some apparently abstract entities, such as languages or the game of chess, seem to possess at least temporal characteristics because they allegedly are liable to change in time.

Answering to this objection Jonathan Lowe (1997) discriminates between a "language" conceived as a universal and a "language" conceived as a social practice. A language in the first sense, Lowe notes, is timeless, whereas a language in the second sense is not. But only the former is an abstract entity unlike the latter, which is concrete. The same distinction can well be drawn in other analogous cases, such as that of the game of chess.

An alternative reaction would be to insist that proper abstract entities yet *must* be non-spatiotemporal, and thus, non-timeless abstract entities should be considered in a way as defective. In this respect truth values may be regarded as perfect abstract objects as they clearly have nothing to do with physical spacetime.

In a similar way truth values fulfill another requirement often imposed upon abstract objects, namely the one of a causal inefficacy (see, e.g., (Grossmann 1992)). Here again, truth values are very much like numbers and geometrical figures: they have no causal power and "make nothing happen".

Finally, let us consider how truth values can be introduced by applying so-called abstraction principles which are used for supplying abstract objects with *criteria of identity*. The idea of this method of characterizing abstract objects is also largely due to Frege (1888) who wrote:

> *If the symbol* a *is to designate an object for us, then we must have a criterion that decides in all cases whether* b *is the same as* a*. . . .*

More precisely, one obtains a new object by abstracting it from some given kind of entities, in virtue of certain criteria of identity for this new (abstract) object. This abstraction is in terms of an equivalence relation defined on the given entities (see (Wrigley 2006)).

The celebrated slogan by Quine "No entity without identity" is intended to express essentially the same understanding of an (abstract) object as an "item falling under a sortal concept which supplies a well-defined criterion of identity for its instances" (Lowe 1997).

For truth values such a criterion has been suggested in (Anderson & Zalta 2004), stating that for any two sentences p and q, the truth value of p is identical with the truth value of q if and only if p is equivalent with q.

We may formalize this idea following the style of presentation in (Lowe 1997) as follows:

$$\forall p \forall q (Sentence(p) \wedge Sentence(q)) \rightarrow (tv(p) = tv(q) \leftrightarrow (p \leftrightarrow q)).$$

Carnap (1947), when introducing truth-values as extensions of sentences, is guided by essentially the same idea. Applying the well-known technique of interpreting sentences as predicators of degree 0, he generalizes the fact that two predicators of degree n (say, P and Q) have the same extension if and only if $\forall x_1 \forall x_2 \ldots \forall x_n (P x_1 x_2 \ldots x_n \leftrightarrow Q x_1 x_2 \ldots x_n)$ holds. Two sentences (say, p and q), being interpreted as zero-degree predicators, must then have the same extension if and only if $p \leftrightarrow q$ holds. And then, Carnap remarks, it seems quite natural to take truth values as extensions for sentences.

Note, that here we deal with a so called "two-level criterion" (see (Dummett 1981), (Lowe 1997)), which employs a *functional dependency* between an introduced abstract object (in this case a truth value) and some other objects (sentences). More specifically we consider the *truth value of* a sentence (or proposition). The criterion of identity for truth values is formulated then by means of the relation of equivalence holding between these other objects (sentences, propositions), with an explicit quantification over them.

In contrast with this, a "one-level criterion" would involve a quantification over the introduced abstract objects themselves, defining identity by means of some other equivalence relation obtaining again between them.

Anderson and Zalta (2004), making use of an object theory from (Zalta 1983), propose the following definition of "the truth value of proposition p" ($tv(p)$, notation adjusted):

$$tv(p) =_{df} \iota x(A!x \wedge \forall F(xF \leftrightarrow \exists q(q \leftrightarrow p \wedge F = [\lambda yq]))),$$

where $A!$ stands for a primitive theoretical predicate "being abstract", xF is to be read as "x encodes F" and $[\lambda yq]$ is a propositional property ("being such a y that q").

That is, according to this definition, "the extension of p is the abstract object that encodes all and only the properties of the form $[\lambda yq]$ which are constructed out of propositions q materially equivalent to p" (Anderson & Zalta 2004).

A truth value can then be defined as an object which is the truth value of some proposition:

$$TV(x) =_{df} \exists p(x = tv(p)).$$

Using this apparatus, it is possible to define the Fregean truth values *the true* ($\top$) and *the false* ($\bot$):

$$\top =_{df} \iota x(A!x \wedge \forall F(xF \leftrightarrow \exists p(p \wedge F = [\lambda y p]))).$$

$$\bot =_{df} \iota x(A!x \wedge \forall F(xF \leftrightarrow \exists p(\neg p \wedge F = [\lambda y p]))).$$

Anderson and Zalta prove then that $\top$ and $\bot$ are indeed truth values, and moreover, that there are exactly two such objects. The latter result is quite expectable, if we keep in mind that what the definitions above introduce are the *classical* truth values (as the underlying logic is classical).

As is well-known, it was Łukasiewicz, who in 1918 proposed to take seriously other values in addition to truth and falsehood. He introduced a third truth value and interpreted it as "possible". By generalizing this idea of a many-valued logic, one naturally arrives at the representation of particular logical systems as valuation systems, alias logical matrices.

Let $\mathcal{L}$ be a propositional language defined on a set P of atomic sentences and a finite set $\mathcal{C} = \{c_1, \ldots, c_m\}$ of finitary connectives. Then a *valuation system* $\mathbf{V}$ for the language $\mathcal{L}$ is a triple $\langle \mathcal{V}, \mathcal{D}, \mathcal{F} \rangle$, where $\mathcal{V}$ is a non-empty set with at least two elements, $\mathcal{D} \subset \mathcal{V}$, $\mathcal{D} \neq \varnothing$ and $\mathcal{F} = \{f_{c_1}, \ldots f_{c_m}\}$ is a set of functions such that each $f_{c_i} : \mathcal{V}^{n_i} \longrightarrow \mathcal{V}$, where n_i is the arity of c_i.

To a matrix one may add an assignment which maps P into $\mathcal{V}$. Each assignment a relative to a valuation system **V** is then extended to a valuation function v_a from formulas to $\mathcal{V}$. Normally, it is required that for every $p \in P$, $v_a(p) = a(p)$, and for every $c_i \in \mathcal{C}$, $v_a(c_i(A_1, \dots A_n)) = f_{c_i}(v_a(A_1), \dots v_a(A_n))$.

The set $\mathcal{D}$ of designated values represents a generalization of the classical value T (*the true*). Why?

A sentence A is a tautology in a valuation system **V** iff for every assignment a relative to **V**, $v_a(A) \in \mathcal{D}$.

For a given valuation system **V** an entailment relation ($\models_V$) is usually defined by postulating the preservation of designated values from the premises to the conclusion:

$$\Delta \models_V A \text{ iff } \forall v_a : (\forall B \in \Delta : v_a(B) \in \mathcal{D}) \Rightarrow v_a(A) \in \mathcal{D}.$$

A pair $\mathcal{M} = \langle \mathbf{V}, v_a \rangle$, where $\mathbf{V}$ is an (n-valued) valuation system and v_a a valuation in $\mathbf{V}$, may be called an (n-valued) *model* based on $\mathbf{V}$.

The notion of entailment can, of course, also be restricted to models: $\models_{\mathcal{M}}$ iff $\Delta \models_{\mathcal{M}} A$ iff $(\forall B \in \Delta : v_a(B) \in \mathcal{D}) \Rightarrow v_a(A) \in \mathcal{D}$.

If we have a syntactically defined logical system $\mathfrak{L}$ with a consequence relation $\vdash_{\mathfrak{L}}$ between sets of $\mathcal{L}$-formulas and single formulas, then a valuational system $\mathbf{V}$ is said to be strictly characteristic for $\mathfrak{L}$ just in case $\Delta \models_{V} A$ iff $\Delta \vdash_{\mathfrak{L}} A$. Conversely, $\mathfrak{L}$ is characterized by $\mathbf{V}$.

$\mathfrak{L}$ is characterized by a class $\mathfrak{K}$ of n-valued models iff $\vdash_{\mathfrak{L}} = \bigcap\{\models_{\mathcal{M}} \mid \mathcal{M} \in \mathfrak{K}\}$.

If we restrict ourselves to a non-implicational language, then Kleene's (strong) "logic of indeterminacy" K_3 is specified by the Kleene matrix $K_3 = \langle\{T, I, F\}, \{T\}, \{f_c : c \in \{\sim, \wedge, \vee\}\}\rangle$, where the functions f_c are defined as follows:

$f_\sim$	
T	F
I	I
F	T

$f_\wedge$	T	I	F
T	T	I	F
I	I	I	F
F	F	F	F

$f_\vee$	T	I	F
T	T	T	T
I	T	I	I
F	T	I	F

The Priest matrix $\mathbf{P}_3$ differs from $\mathbf{K}_3$ only in that $\mathcal{D} = \{T, I\}$.

There are intuitive interpretations of I in $\mathbf{K}_3$ and in $\mathbf{P}_3$ as the underdetermined and the overdetermined value respectively (a truth-value gap and a truth-value glut).

These readings can be modeled by considering $\mathbf{4} = \mathcal{P}(\{T, F\})$. Then $\mathbf{T} = \{T\}$ is understood as "true only", $\mathbf{F} = \{F\}$ as "false only", $\mathbf{N} = \varnothing$ as "*neither* true nor false" (I in K_3), and $\mathbf{B} = \{T, F\}$ as "*both* true and false" (I in Priest's "logic of paradox" P_3).

$\mathcal{P}(\{T, F\})$ is the set of values of the four-valued logic B_4 introduced by Dunn (1976) and Belnap (1977). This logic is determined by the matrix $\mathbf{B}_4 = \langle\{\mathbf{N}, \mathbf{T}, \mathbf{F}, \mathbf{B}\}, \{\mathbf{T}, \mathbf{B}\}, \{f_c : c \in \{\sim, \wedge, \vee\}\}\rangle$, where the functions f_c are defined as follows:

$f_\sim$	
T	**F**
B	**B**
N	**N**
F	**T**

$f_\wedge$	**T**	**B**	**N**	**F**
T	**T**	**B**	**N**	**F**
B	**B**	**B**	**F**	**F**
N	**N**	**F**	**N**	**F**
F	**F**	**F**	**F**	**F**

$f_\vee$	**T**	**B**	**N**	**F**
T	**T**	**T**	**T**	**T**
B	**T**	**B**	**T**	**B**
N	**T**	**T**	**N**	**N**
F	**T**	**B**	**N**	**F**

One might, perhaps, think that the mere existence of many-valued logics shows that there exist infinitely, in fact, uncountably many truth values. However, this is not at all clear.

First, in the literature on many-valued logics, one can sometimes find a cautious use of terminology, namely, a distinction is drawn between the classical truth values *the true* and *the false* on the one hand and (additional) quasi truth values on the other hand, see, for example, (Gottwald 1989).

Moreover, Roman Suszko declared that many-valued logic is "a magnificent conceptual deceit" (Suszko 1977) and claimed that "there are but two logical values, true and false", a statement now called Suszko's Thesis.

For Suszko, the set of truth values assumed in a logical matrix for a many-valued logic is a set of "admissible referents" (called "algebraic values") of formulas but not a set of logical values. Whereas the algebraic values are referents of formulas, the logical value *true* is used to define valid consequence: If every premise is true, then so is (at least one of) the conclusion(s). The other logical value, *false*, is preserved in the opposite direction: If the (every) conclusion is false, then so is at least one of the premises.

The logical values are thus represented by a bi-partition of the set of algebraic values into a set of designated values (truth) and its complement (falsity).

Suszko's thesis is substantiated by a proof (the Suszko Reduction) showing that every structural Tarskian consequence relation and therefore also every structural Tarskian many-valued propositional logic is characterized by a bivalent semantics.

Note also that R. Routley (1975) has shown that every logic based on a λ-categorical language has a sound and complete bivalent possible worlds semantics.

The dichotomy between designated values and values which are not designated and its use in the definition of entailment plays a crucial role in the Suszko Reduction.

Ryszard Wójcicki (1970) showed that every structural Tarskian logic (every reflexive, transitive consequence relation satisfying (Cut)) is characterized by its so-called Lindenbaum bundle:

$$\{\langle\langle\mathcal{L}, \{A \in \mathcal{L} \mid \Delta \vdash A\}, \mathcal{C}\rangle, v\rangle \mid \Delta \subseteq \mathcal{L}, v \text{ is a uniform substitution on } \mathcal{L}\}.$$

Theorem (Wójcicki)

Every structural Tarskian logic is characterized by a class of n-valued models, for some $n \leq \aleph_0$.

Theorem (Suszko 1977, Malinowski 1993)

Every structural Tarskian logic is characterized by a class of two-valued models.

Proof. Let $\Lambda = \langle \mathcal{L}, \vdash \rangle$ be a structural Tarskian logic. By Wójcicki's Theorem, the logic Λ is characterized by a class $\mathfrak{C}_\Lambda$ of n-valued models. For $\langle \langle \mathcal{V}, \mathcal{D}, \{f_c : c \in \mathcal{C}\} \rangle, v \rangle \in \mathfrak{C}_\Lambda$, the function t_v from $\mathcal{L}$ into $\{0, 1\}$ is defined as follows:

$$t_v(A) = \begin{cases} 1 & \text{if} \quad v(A) \in \mathcal{D} \\ 0 & \text{if} \quad v(A) \notin \mathcal{D} \end{cases}$$

The class
$\{\langle \langle \{0, 1\}, \{1\}, \{f_c : c \in \mathcal{C}\} \rangle, t_v \rangle \mid \langle \langle \mathcal{V}, \mathcal{D}, \{f_c : c \in \mathcal{C}\} \rangle, v \rangle \in \mathfrak{C}_\Lambda\}$ of 2-valued models characterizes Λ. q.e.d.

Whereas the algebraic values that are not designated are already given with the set of designated values $\mathcal{D}$ as its complement, the treatment of *true* and *false* as values that are independent of each other leads to distinguishing between sets $\mathcal{D}^+$ and $\mathcal{D}^-$ of designated and antidesignated values.

Let $\mathcal{L}$ be as above. An n-valued q-matrix (quasi-matrix) based on $\mathcal{L}$ is defined by Malinowski as a structure $\mathfrak{M} = \langle \mathcal{V}, \mathcal{D}^+, \mathcal{D}^-, \{f_c : c \in \mathcal{C}\} \rangle$, where $\mathcal{V}$ is a non-empty set of cardinality $n \geq 2$, $\mathcal{D}^+$ and $\mathcal{D}^-$ are distinct non-empty proper subsets of $\mathcal{V}$ such that $\mathcal{D}^+ \cap \mathcal{D}^- = \varnothing$, and every f_c is a function on $\mathcal{V}$ with the same arity as c.

To obtain a kind of entailment relation that does not admit of a reduction to a bivalent semantics, Malinowski defines *q*-entailment as depending on *both* sets $\mathcal{D}^+$ and $\mathcal{D}^-$.

A *q*-matrix $\mathfrak{M} = \langle \mathcal{V}, \mathcal{D}^+, \mathcal{D}^-, \{f_c : c \in \mathcal{C}\} \rangle$ determines a relation $\models_{\mathfrak{M}} \subseteq \mathcal{P}(\mathcal{L}) \times \mathcal{L}$ (*q*-entailment) by defining $\Delta \models_{\mathfrak{M}} A$ iff for every valuation v in $\mathfrak{M}$, $v(\Delta) \cap \mathcal{D}^- = \varnothing$ implies $v(A) \in \mathcal{D}^+$.

A pair $\mathcal{M} = \langle \langle \mathcal{V}, \mathcal{D}^+, \mathcal{D}^-, \{f_c : c \in \mathcal{C}\} \rangle, v \rangle$, where $\mathfrak{M} = \langle \mathcal{V}, \mathcal{D}^+, \mathcal{D}^-, \{f_c : c \in \mathcal{C}\} \rangle$ is an *n*-valued *q*-matrix and v a valuation in $\mathfrak{M}$, may be called an *n*-valued *q*-model based on $\mathfrak{M}$. The relation $\models_{\mathcal{M}} \subseteq \mathcal{P}(\mathcal{L}) \times \mathcal{L}$ determined by such a model is defined by the following equivalence: $\Delta \models_{\mathcal{M}} A$ iff $v(\Delta) \cap \mathcal{D}^- = \varnothing$ implies $v(A) \in \mathcal{D}^+$.

If $\mathfrak{M} = \langle \mathcal{V}, \mathcal{D}^+, \mathcal{D}^-, \{f_c : c \in \mathcal{C}\}\rangle$ is a q-matrix and $\mathcal{D}^+$ is not the complement of $\mathcal{D}^-$, there is no class of functions v from $\mathcal{L}$ into $\{1, 0\}$ such that $\Delta \models_{\mathfrak{M}} A$ iff for every function from that class, $v(\Delta) \subseteq \{1\}$ implies $v(A) = 1$.

Let $\mathcal{M} = \langle\langle \mathcal{V}, \mathcal{D}^+, \mathcal{D}^-, \{f_c : c \in \mathcal{C}\}\rangle, v\rangle$ be an n-valued q-model. Malinowski pointed out that an equivalent three-valued q-model $\mathcal{M}' = \langle\langle\{0, \frac{1}{2}, 1\}, \{1\}, \{0\}, \{f_c : c \in \mathcal{C}\}\rangle, t_v\rangle$ can be defined as follows:

$$t_v(A) = \begin{cases} 1 & \text{if} \quad v(A) \in \mathcal{D}^+ \\ \frac{1}{2} & \text{iff} \quad v(A) \in \mathcal{V} \setminus (\mathcal{D}^+ \cup \mathcal{D}^-) \\ 0 & \text{if} \quad v(A) \in \mathcal{D}^- \end{cases}$$

A q-entailment relation $\models_{\mathfrak{M}}$ is a special case of what Malinowski calls a q-consequence relation. A q-consequence relation on $\mathcal{L}$ is a relation $\Vdash \subseteq \mathcal{P}(\mathcal{L}) \times \mathcal{L}$ such that for every $A \in \mathcal{L}$ and every Δ, $\Gamma \subseteq \mathcal{L}$:

$$\text{If } \Delta \Vdash A \text{ then } \Delta \cup \Gamma \Vdash A \;\; \text{(Monotonicity)} \tag{1}$$

$$\Delta \cup \{B \mid \Delta \Vdash B\} \Vdash A \text{ iff } \Delta \Vdash A \;\; \text{(Quasi-closure)} \tag{2}$$

A q-consequence relation on $\mathcal{L}$ is called *structural* iff for every $A \in \mathcal{L}$, every $\Delta \subseteq \mathcal{L}$, and every uniform substitution function σ on $\mathcal{L}$ we have

$$\Delta \Vdash A \text{ iff } \sigma(\Delta) \Vdash \sigma(A) \;\; \text{(Structurality).} \tag{3}$$

A pair $(\mathcal{L}, \Vdash)$ is said to be a *q-logic*, it is structural iff $\Vdash$ is structural. A q-logic $\langle \mathcal{L}, \Vdash \rangle$ is said to be characterized by an n-valued q-model $\mathcal{M}$ iff $\Vdash = \models_{\mathcal{M}}$, and $\langle \mathcal{L}, \Vdash \rangle$ is characterized by a class $\mathfrak{K}$ of n-valued q-models iff $\Vdash = \bigcap \{ \models_{\mathcal{M}} \mid \mathcal{M} \in \mathfrak{K} \}$.

If $\Delta \subseteq \mathcal{L}$, let $\mathcal{D}^{+}_{\Delta} = \{ A \in \mathcal{L} \mid \Delta \vdash A \}$ and $\mathcal{D}^{-}_{\Delta} = \mathcal{L} \setminus (\Delta \cup \{ A \in \mathcal{L} \mid \Delta \vdash A \})$.

Malinowski (1990) showed that every structural q-logic is characterized by the following Lindenbaum bundle:

$$\{ \langle \langle \mathcal{L}, \mathcal{D}^{+}_{\Delta}, \mathcal{D}^{-}_{\Delta}, \mathcal{C} \rangle, v \rangle \mid \Delta \subseteq \mathcal{L}, v \text{ is a uniform substitution on } \mathcal{L} \}.$$

Theorem (Malinowski)

Every structural q-logic is characterized by a class of n-valued q-models, for some $n \leq \aleph_0$.

By the above definition of three-valued q-models $\mathcal{M}'$ and by the Suszko Reduction for the case that $\mathcal{V} \setminus (\mathcal{D}^+ \cup \mathcal{D}^-) = \varnothing$, it follows that q-logics are logically two-valued or three-valued.

Theorem (Malinowski)

Every structural q-logic is characterized by a class of two-valued q-models or by a class of three-valued q-models.

The Suszko Reduction can be proved also without making use of matrices. Such a proof has been given by Gabbay (1981) for Tarski-Scott multiple conclusion consequence relations.

A relation $\models \subseteq \mathcal{P}(\mathcal{L}) \times \mathcal{P}(\mathcal{L})$ is a Tarski-Scott multiple conclusion consequence relation iff it satisfies the following conditions:

- For every $\Delta \subseteq \mathcal{L}$, $\Delta \models \Delta$ (reflexivity);
- If $\Delta \models \Gamma \cup \{A\}$ and $\{A\} \cup \Theta \models \Sigma$, then $\Delta \cup \Theta \models \Gamma \cup \Sigma$ (transitivity);
- If $\Delta \subseteq \Theta$, $\Gamma \subseteq \Sigma$, and $\Delta \models \Gamma$, then $\Theta \models \Sigma$ (monotony).

Let $\mathcal{L}$ be a propositional language, let $\Vdash$ be any Tarski-Scott multiple conclusion consequence relation, and let Δ, Γ be sets of $\mathcal{L}$-formulas. The pair $\langle\Delta,\Gamma\rangle$ is said to be consistent iff $\Delta \nVdash \Gamma$, and it is said to be complete if $\Delta \cup \Gamma$ is the set of all $\mathcal{L}$-formulas. Note that the entailment relation $\models_X$ with respect to a set X of bivaluations from the set of $\mathcal{L}$-formulas into $\{0,1\}$ is a Tarski-Scott multiple conclusion consequence relation.

The relation $\models_X$ is defined by

$$\Delta \models_X \Gamma \text{ iff } (\forall v \in X)\,((\forall B \in \Delta)\, v(B) = 1) \Rightarrow ((\exists C \in \Gamma)\, v(C) = 1)).$$

Lemma

Every consistent pair of sets of $\mathcal{L}$-formulas can be extended to a consistent and complete pair of sets of $\mathcal{L}$-formulas.

The crucial idea for the reduction is to define a suitable set of bivaluations.

Definition

Let $\langle \Delta, \Gamma \rangle$ be a consistent and complete pair of sets of $\mathcal{L}$-formulas. The bivaluation $v_{\langle \Delta, \Gamma \rangle}$ from the set of all $\mathcal{L}$-formulas into $\{0, 1\}$ is defined by

$$v_{\langle \Delta, \Gamma \rangle}(A) = 1 \text{ iff } A \in \Delta.$$

The set S of bivaluations is defined by

$$S := \{v_{\langle \Delta, \Gamma \rangle} \mid \langle \Delta, \Gamma \rangle \text{ is consistent and complete}\}.$$

The following theorem shows that every Tarski-Scott multiple conclusion consequence relation is characterized by a set of bivaluations.

Theorem

$\Vdash \, = \, \models_S$.

Proof. Soundness: If $\Theta \Vdash \Sigma$, then $\langle \Theta, \Sigma \rangle$ is not consistent. Therefore, there exists no consistent and complete pair $\langle \Delta^*, \Gamma^* \rangle$ that extends $\langle \Theta, \Sigma \rangle$, i.e., such that $\Theta \subseteq \Delta^*$ and $\Sigma \subseteq \Gamma^*$. Suppose that $v_{\langle \Delta, \Gamma \rangle} \in S$. Then $\Delta \nVdash \Gamma$. If for every $A \in \Theta$, $v_{\langle \Delta, \Gamma \rangle}(A) = 1$, then $\Theta \subseteq \Delta$. Suppose furthermore that for every $B \in \Sigma$, $v_{\langle \Delta, \Gamma \rangle}(B)$ $= 0$. Then no $B \in \Sigma$ belongs to Θ. By the completeness of $\langle \Delta, \Gamma \rangle$, every $B \in \Sigma$ belongs to Γ, i.e., $\Sigma \subseteq \Gamma$. But then the complete pair $\langle \Delta, \Gamma \rangle$ extends the pair $\langle \Theta, \Sigma \rangle$, a contradiction. Thus, $\Theta \models_S \Sigma$.
Completeness: If $\Theta \nVdash \Sigma$, this pair has some consistent and complete extension $\langle \Delta, \Gamma \rangle$. The valuation $v_{\langle \Delta, \Gamma \rangle}$ shows that $\Theta \not\models_S \Sigma$.

Suszko (1977) does not define the notion of a logical value except for stating that *true* and *false* are the only logical values, but he claims that "any multiplication of logical values is a mad idea".

One may ask by virtue of which properties *true* and *false* are to be considered as *logical* values.

Truth is what is preserved in a valid inference from the premises to the conclusion. The value *true* is given with a non-empty proper subset $\mathcal{D}$ of the set of algebraic values and the corresponding notion of entailment, understood as the preservation of membership in $\mathcal{D}$ from the premises to the conclusion(s).

Let us refer to this notion of entailment as *t*-entailment. A formula A is logically true iff A is t-entailed by $\varnothing$ (iff for every assignment v of algebraic values to the formulas of the language under consideration, $v(A)$ is designated), and A is logically false iff A t-entails $\varnothing$ (iff for every assignment v, $v(A)$ is not designated).

One might wish to consider a notion of f-entailment understood as the backward preservation of values associated with falsity. Obviously, membership in the complement of $\mathcal{D}$ is preserved from the conclusion(s) to the premises, but this gives the same relation as t-entailment.

Since $\mathcal{D}$ is uniquely determined by its complement, and vice versa, logical two-valuedness is, in fact, reduced to logical mono-valuedness if there is just one entailment relation defined as truth preservation form the premises to the conclusion. Thus, classical propositional logic is *not* logically two-valued, because t-entailment and f-entailment coincide.

The notion of a q-matrix allows further generalization. One obtains generalized q-matrices by giving up the condition $\mathcal{D}^+ \cap \mathcal{D}^- = \varnothing$.

Also, in addition to the preservation of truth from the premises to the conclusion(s) and the preservation of falsity in the backward direction, there come to mind at least two other notions of entailment based on an interplay between $\mathcal{D}^+$ and $\mathcal{D}^-$. The framework of generalized q-matrices allows four basic entailment relations (t-entailment, f-entailment, q-entailment and p-entailment):

$$\Delta \models_t B \text{ iff } \forall v_a : (\forall A \in \Delta : v_a(A) \in \mathcal{D}^+) \Rightarrow v_a(B) \in \mathcal{D}^+ \quad (4)$$
$$\Delta \models_f B \text{ iff } \forall v_a : (\forall A \in \Delta : v_a(A) \notin \mathcal{D}^-) \Rightarrow v_a(B) \notin \mathcal{D}^- \quad (5)$$
$$\Delta \models_q B \text{ iff } \forall v_a : (\forall A \in \Delta : v_a(A) \notin \mathcal{D}^-) \Rightarrow v_a(B) \in \mathcal{D}^+ \quad (6)$$
$$\Delta \models_p B \text{ iff } \forall v_a : (\forall A \in \Delta : v_a(A) \in \mathcal{D}^+) \Rightarrow v_a(B) \notin \mathcal{D}^- \quad (7)$$

Whereas *t*-entailment is a generalization of the standard truth-preserving relation, *f*-entailment incorporates the idea of non-falsity preservation. Malinowski's relation of *q*-entailment can be seen as reflecting a reasoning from hypotheses (understood as statements that merely are taken to be non-refuted).

p-entailment has been studied by Frankowski (2004), who tried to explicate "reasonings wherein the degree of strength of the conclusion (i.e. the conviction it is true) is smaller th[a]n that of the premisses".

By a canonical definition of entailment one may understand a definition of entailment as a relation that preserves membership in a certain set of algebraic values either from the premises to the conclusion(s) of inferences, or from the conclusion(s) to the premises.

Every such set may be associated with some logical value, and the corresponding entailment relations is Tarskian.

Two logical values are independent of each other iff the canonically defined entailment relations determined by these values are distinct.

This idea of multiple logical truth values is developed in (Wansing and Shramko 2008). It is neither Malinowski's nor Suszko's understanding of logical truth values, but may be expressed by using Malinowski's term "inferential many-valuedness".

One may consider the notion of a *generalized valuation system* (an n-valued k-dimensional matrix, or just k-matrix) which is a structure $\mathfrak{M} = \langle \mathcal{V}, \mathcal{D}_1, \ldots, \mathcal{D}_k, \{f_c : c \in \mathcal{C}\} \rangle$, where $\mathcal{V}$ is a non-empty set of cardinality n $(2 \leq n)$, $2 \leq k$, every $\mathcal{D}_i$ $(1 \leq i \leq k)$ is a non-empty proper subset of $\mathcal{V}$, the sets $\mathcal{D}_i$ are pairwise distinct, and every f_c is a function on $\mathcal{V}$ with the same arity as c (cf. (Wójcicki 1988), (Czelakowski 2001)).

A logic may then be said to be logically (or inferentially) k-valued if it is a language with k canonically defined and pairwise distinct entailment relations on (the set of formulas of) this language.

Interpretations of distinguished sets of algebraic values need not appeal to truth or falsity. In a series of papers, Jennings, Schotch, and Brown, for example, have argued that paraconsistent logic can be developed as a logic that preserves a degree of incoherence from the premises to the conclusion of a valid inference.

Can one provide evidence for a multiplicity of logical values? More concretely, *is* there more than one logical value, each of which may be taken to determine its own (independent) entailment relation?

A positive answer to this question emerges from considerations on truth values as structured entities which, by virtue of their internal structure, give rise to natural partial orderings on the set of values.

The truth values of both Kleene's and Priest's logic can be ordered to form a lattice, *THREE*:

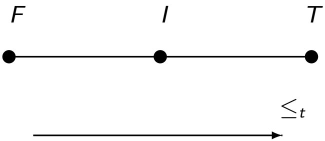


Figure: Lattice *THREE*

Here $\leq_t$ (the "logical order") orders T, I and F so that the intermediate value I is "more true" than F, but "less true" than T.

We can consider an entailment relation as expressing agreement with the truth order, that is:

$$\Delta \models B \text{ iff } \forall v_a : \bigsqcap_t \{v_a(A) \mid A \in \Delta\} \leq_t v_a(B), \qquad (8)$$

where $\sqcap_t$ is the lattice meet.

Another prominent partially ordered valuation system is the matrix $\mathbf{B}_4$ considered above. The set of truth values $\{\mathbf{N}, \mathbf{T}, \mathbf{F}, \mathbf{B}\}$ from $\mathbf{B}_4$ constitutes the bilattice $FOUR_2$ (see, e.g., (Ginsberg 1988), (Arieli and Avron 1996), Fitting (2006).

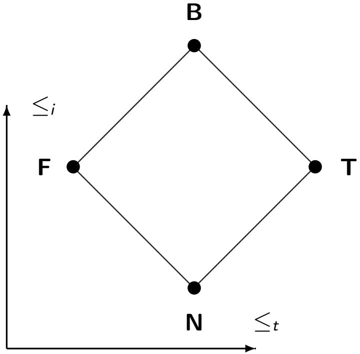


Figure: The bilattice $FOUR_2$

In addition to a truth order, there is an information order ($\leq_i$) which is said to order the values under consideration according to the information they give concerning a formula to which they are assigned. Lattice meet and join with respect to $\leq_t$ coincide with the functions $f_\wedge$ and $f_\vee$ in $\mathbf{B}_4$, $f_\sim$ turns out to be the truth order inversion, and an entailment relation, which happens to coincide with the matrix entailment, is defined as above.

Frege (1892) points out the possibility of "distinctions of parts within truth values". Although he immediately specifies that the word "part" is used here "in a special sense", the basic idea seems nevertheless to be that truth values are not something amorphous, but have an internal structure. That is, truth may well be interpreted as complex, structured entities that can be divided into "parts".

This idea nicely conforms with the modeling of truth values as subsets of the set of classical truth values. The latter approach stems essentially from (Dunn 1976), where a generalization of the notion of a classical truth-value function has been proposed to obtain so-called "underdetermined" and "overdetermined" valuations, see also (Dunn 2000).

By developing this idea, we arrive at the concept of generalized truth value functions, which are functions from sentences into the *subsets* of some *basic set of truth values* (Shramko and Wansing 2005). The values of generalized truth value functions can be called generalized truth values.

In (Shramko and Wansing 2005, 2006) it is argued that if we move from the set **4** of generalized truth values to its powerset $\mathcal{P}(\mathbf{4}) = \mathbf{16}$, we may not only obtain a *truth* order on the underlying set of values that is better justified than the truth order of $FOUR_2$ but in addition also an independent and equally well-justified *falsity* order.

1.	$\mathbf{N} = \varnothing$	9.	$\mathbf{FT} = \{\{F\}, \{T\}\}$
2.	$\boldsymbol{N} = \{\varnothing\}$	10.	$\mathbf{FB} = \{\{F\}, \{F, T\}\}$
3.	$\boldsymbol{F} = \{\{F\}\}$	11.	$\mathbf{TB} = \{\{T\}\}, \{F, T\}\}$
4.	$\boldsymbol{T} = \{\{T\}\}$	12.	$\mathbf{NFT} = \{\varnothing, \{F\}, \{T\}\}$
5.	$\boldsymbol{B} = \{\{F, T\}\}$	13.	$\mathbf{NFB} = \{\varnothing, \{F\}, \{F, T\}\}$
6.	$\mathbf{NF} = \{\varnothing, \{F\}\}$	14.	$\mathbf{NTB} = \{\varnothing, \{T\}, \{F, T\}\}$
7.	$\mathbf{NT} = \{\varnothing, \{T\}\}$	15.	$\mathbf{FTB} = \{\{F\}, \{T\}, \{F, T\}\}$
8.	$\mathbf{NB} = \{\varnothing, \{F, T\}\}$	16.	$\mathbf{A} = \{\varnothing, \{T\}, \{F\}, \{F, T\}\}.$

In $FOUR_2$ truth and falsity are not dealt with as independent of each other. Is assumed that F by itself is less true than T, and hence, in $FOUR_2$ $\{T, F\}$ is taken to be less true than $\{T\}$.

In **16** truth and falsity are treated as independent notions. For every x in **16** we define the sets x^t, x^{-t}, x^f, and x^{-f} as follows:

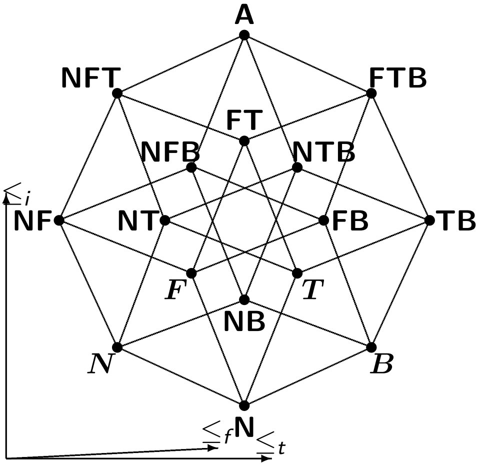
$$x^t := \{y \in x \mid T \in y\}\,; \;\; x^{-t} := \{y \in x \mid T \notin y\}\,;$$
$$x^f := \{y \in x \mid F \in y\}\,; \;\; x^{-f} := \{y \in x \mid F \notin y\}\,.$$

For every x, y in **16** we then put:

- $x \leq_i y$ iff $x \subseteq y$;
- $x \leq_t y$ iff $x^t \subseteq y^t$ and $y^{-t} \subseteq x^{-t}$;
- $x \leq_f y$ iff $x^f \subseteq y^f$ and $y^{-f} \subseteq x^{-f}$.

Note that $\leq_f$ is not the inversion of $\leq_t$.

We obtain a structure that combines the three (complete) lattices $(\mathbf{16}, \leq_i)$, $(\mathbf{16}, \leq_t)$, and $(\mathbf{16}, \leq_f)$ into the *trilattice SIXTEEN*$_3$= $(\mathbf{16}, \leq_i, \leq_t, \leq_f)$, cf. also (Shramko, Dunn, Takenaka 2001).



Meets and joints exist in $SIXTEEN_3$ for all three partial orders. We will use $\sqcap$ and $\sqcup$ with the appropriate subscripts for these operations under the corresponding ordering relations. Since from the operations one can recover the relations, $SIXTEEN_3$ may also be represented as the structure $(\mathbf{16}, \sqcap_i, \sqcup_i, \sqcap_t, \sqcup_t, \sqcap_f, \sqcup_f)$.

In what follows we shall focus on the "logical" operations $\sqcap_t, \sqcup_t, \sqcap_f$ and $\sqcup_f$. Since the relations $\leq_t$ and $\leq_f$ are treated on a par, the operations $\sqcap_t$ and $\sqcup_t$ are not privileged as interpretations of conjunction and disjunction. The operation $\sqcup_f$ may as well be regarded as a conjunction and $\sqcap_f$ as a disjunction. In other words, the logical vocabulary may be naturally considered to comprise a positive truth vocabulary and a negative falsity vocabulary.

Also certain unary truth and falsity operations with natural negation-like properties are available in $SIXTEEN_3$. A unary operation $-_t$ ($-_f$) on $SIXTEEN_3$ is said to be a t-inversion (f-inversion) iff the following conditions are satisfied:

1. $t-inversion(-_t)$:
(a) $a \leq_t b \Rightarrow -_t b \leq_t -_t a$;
(b) $a \leq_f b \Rightarrow -_t a \leq_f -_t b$;
(c) $a \leq_i b \Rightarrow -_t a \leq_i -_t b$;
(d) $-_t -_t a = a$.

2. $f-inversion(-_f)$:
(a) $a \leq_t b \Rightarrow -_f a \leq_t -_f b$;
(b) $a \leq_f b \Rightarrow -_f b \leq_f -_f a$;
(c) $a \leq_i b \Rightarrow -_f a \leq_i -_f b$;
(d) $-_f -_f a = a$.

A t-inversion (f-inversion) inverts the truth (falsity) order, leaves the other orders untouched, and is period-two. In $SIXTEEN_3$ such operations are definable.

Thus, also negation emerges in two versions, because $-_t$ and $-_f$ are both natural interpretations for a negation connective. Moreover, since $x \sqcap_t y \neq x \sqcup_f y$, $x \sqcup_t y \neq x \sqcap_f y$ and $-_t x \neq -_f x$, the two logical orderings $\leq_t$ and $\leq_f$ indeed give rise to pairs of *distinct* logical operations with the same arity.

We may define three propositional languages $\mathcal{L}_t$, $\mathcal{L}_f$, and $\mathcal{L}_{tf}$ as follows:

$$\begin{array}{rll} \mathcal{L}_t: & A ::= & p \mid \sim_t A \mid A \wedge_t A \mid A \vee_t A \\ \mathcal{L}_f: & A ::= & p \mid \sim_f A \mid A \wedge_f A \mid A \vee_f A \\ \mathcal{L}_{tf}: & A ::= & p \mid \sim_t A \mid \sim_f A \mid A \wedge_t A \mid A \vee_t A \mid A \wedge_f A \mid A \vee_f A \end{array}$$

Following the lattice approach to defining a many-valued logic, the logic of $SIXTEEN_3$ is semantically presented as a *bi-consequence system*, namely the structure $(\mathcal{L}_{tf}, \models_t, \models_f)$, where the two entailment relations $\models_t$ and $\models_f$ are defined with respect to the truth order $\leq_t$ and the falsity order $\leq_f$, respectively.

Let v be a map from the set of propositional variables into $\mathbf{16}$. The function v is recursively extended to a function from the set of all $\mathcal{L}_{tf}$-formulas into $\mathbf{16}$ as follows:

1. $v(A \wedge_t B) = v(A) \sqcap_t v(B)$;
2. $v(A \vee_t B) = v(A) \sqcup_t v(B)$;
3. $v(\sim_t A) = -_t v(A)$;
4. $v(A \wedge_f B) = v(A) \sqcup_f v(B)$;
5. $v(A \vee_f B) = v(A) \sqcap_f v(B)$;
6. $v(\sim_f A) = -_f v(A)$.

The relations $\models_t \subseteq \mathcal{P}(\mathcal{L}_{tf}) \times \mathcal{P}(\mathcal{L}_{tf})$ and $\models_f \subseteq \mathcal{P}(\mathcal{L}_{tf}) \times \mathcal{P}(\mathcal{L}_{tf})$ are defined by the following equivalences:

$$\Delta \models_t \Gamma \quad \text{iff} \quad \forall v \ \textstyle\bigsqcap_t \{v(A) \mid A \in \Delta\} \leq_t \textstyle\bigsqcup_t \{v(A) \mid A \in \Gamma\};$$
$$\Delta \models_f \Gamma \quad \text{iff} \quad \forall v \ \textstyle\bigsqcup_f \{v(A) \mid A \in \Gamma\} \leq_f \textstyle\bigsqcap_f \{v(A) \mid A \in \Delta\}.$$

Note, that in $SIXTEEN_3$ $\models_t$ and $\models_f$ are distinct relations.

In (Shramko and Wansing 2005) we have shown that in the language $\mathcal{L}_t$ ($\mathcal{L}_f$) the relation $\models_t$ ($\models_f$) for *SIXTEEN*$_3$ can be axiomatized as the system of first degree entailment. In (Shramko and Wansing 2006) we have shown that this holds true also for the trilattices based on $\mathcal{P}^n(\mathbf{4})$, $1 < n \in \mathbb{N}$ (for any Belnap trilattices).

We were also able to axiomatize $\models_t$ ($\models_f$) for *SIXTEEN*$_3$ in the language based on $\{\wedge_t, \vee_t, \sim_t, \sim_f\}$ ($\{\wedge_f, \vee_f, \sim_f, \sim_t\}$).

We did not obtain an axiomatization of $\models_t$ ($\models_f$) for *SIXTEEN*$_3$ in the full language $\mathcal{L}_{tf}$.

Sergei Odintsov (2009) found an axiomatization of $\models_t$ for $SIXTEEN_3$ in $\mathcal{L}_{tf}$ extended by $\rightarrow_t$, the residuum of $\leq_t$:

$$v(A \rightarrow_t B) = \bigsqcup\{x \mid x \sqcap v(A) \leq_t v(B)\}.$$