A quantitative-informational approach to logical consequence

Alves, Marcos Antonio; Loffredo D'Otaviano, Itala M. A quantitative-informational approach to logical consequence. In: *The Road to Universal Logic*. Switzerland: Springer International Publishing, pp. 105-24 (Studies in Universal Logic), 2015

DOI: 10.1007/978-3-319-15368-1_3.
http://link.springer.com/chapter/10.1007/978-3-319-15368-1_3.

Marcos Antonio Alves
UNESP, Philosophy Department, State University of São Paulo, Marília/SP, Brazil
marcosalves@marilia.unesp.br

Itala M. Loffredo D'Ottaviano
Unicamp, Philosophy Department, Centre for Logic, Epistemology and the History of Science – CLE, University of Campinas, Campinas, Brazil
itala@cle.unicamp.br

Abstract
In this work, we propose a definition of logical consequence based on the relation between the quantity of information present in a particular set of formulae and a particular formula. As a starting point, we use Shannon's quantitative notion of information, founded on the concepts of logarithmic function and probability value. We first consider some of the basic elements of an axiomatic probability theory, and then construct a probabilistic semantics for languages of classical propositional logic. We define the quantity of information for the formulae of these languages and introduce the concept of informational logical consequence, identifying some important results, among them: certain arguments that have traditionally been considered valid, such as *modus ponens*, are not valid from the informational perspective; the logic underlying informational logical consequence is not classical, and is at the least paraconsistent *sensu lato*; informational logical consequence is not a Tarskian logical consequence.

Keywords: Logical consequence. Information. Probability. Semantics. Informational logical consequence. Non-classical logics. Paraconsistent logic

1 Introduction

Logical consequence can be considered in different ways. In intuitive terms, it is often understood as a relation established between a given set of statements of a language and a statement of the same language. In logic, it is common to define it in a way that is close to intuitive, for example, as a type of relation between elements of the power-set of a nonempty set, as done by Feitosa and D'Ottaviano (2004), amongst others. In this sense, logical consequência is as a relation between a given set (which could be empty or even infinite) of formulae of a language (usually formal), and a formula of the same language. The statements or formulae belonging to the given set are termed *premises*, and the other statement or formula is termed the *conclusion*. The role of the premises is to found, sustain, and support the conclusion.

In the thirties, Tarski (1956a; 1956b; 1956c; 1956d) introduced and improved his definition of consequence operator, or logical consequence, and proved the fundamental properties of it (cf. ALVES, 2012b). In Tarskian terms, given the language L of the (classical) propositional calculus, a *consequence operator* on the set of formulae of L, Form(L), is a function C: $\wp(\text{Form(L)}) \rightarrow \wp(\text{Form(L)})$ such that, for all $\Gamma, \Delta \subseteq$ Form(L), it satisfies the following properties:

(T1) $\Gamma \subseteq C(\Gamma)$;
(T2) If $\Gamma \subseteq \Delta$, then $C(\Gamma) \subseteq C(\Delta)$;
(T3) $C(C(\Gamma)) \subseteq C(\Gamma)$.

The first property above is named reflexivity; the second is named monotonicity; and the third property is named transitivity. A logical consequence operator that satisfies these properties is usually named a *Tarskian logical consequence*.

When suitably combined in a sequence, the set of premises and the conclusion constitute an *argument*. We can say that a conclusion is logical consequence of a set of premises if, and only if, the argument constituted by the union is logically valid, or, simply, valid.

In semantic terms, the relation of consequence is usually defined starting from the degree of truth of the premises and conclusion: a formula is a logical consequence of a given set of formulae if, and only if, it is true under all circumstances (for example, valuation, structure, interpretation, model), such that all the premises are true.

In this work, we propose a definition of logical consequence based on the quantity of information present in the set of premises and in the conclusion. As a starting point, we use the usual languages of classical propositional logic (CPL), as constructed by Shoenfield (1967), for example.

In our theoretical approach to logical consequence, we do not consider the qualitative semantic aspects of information, considered in works such as those of Dretske (1986) and Gonzalez (1996). In our case, the informational value of a message or a formula depends only on its probability of occurrence, established using probability space. In the next section, we consider some elements of a usual axiomatic theory of probabilities, indicating some of its definitions and basic results, which will be used later; these concepts include the notions of event and random experiment.

In the third section, we construct a probabilistic semantics for CPL; we establish a relation between the formulae of the CPL language and the events of a random experiment, from which we define a probability value for each formula of a given language. At the end of the section, we introduce the definitions of probabilistically valid and probabilistically equivalent formulae, and of probabilistic logical consequence.

In the fourth section, we discuss the notion of quantity of information present in a formula of a CPL language; we propose a quantitative-informational definition of logical consequence, which we call *informational logical consequence*, and we demonstrate some of the results and properties that follow from this definition. In particular, we show the existence of arguments which are considered valid according to the classical perspective, but which are invalid from the informational perspective. For example, *modus ponens* is informationally invalid, given the possibility that the conclusion of this argument could possess a greater quantity of information than its set of premises. Furthermore, we show that the logic underlying informational logical consequence is not classical, but is, at the least, paraconsistent *sensu lato*. In addition, we demonstrate that although it might satisfy

the property of transitivity, informational logical consequence is neither reflexive nor monotonic; in other words, it is not a Tarskian logical consequence.

In the final considerations, we summarize and analyze the main properties and results of informational logical consequence.

Our approach is based on the quantitative concept of information, developed in the Mathematical Theory of Communication. One of the pioneers in studies of the quantification, storage, and transmission of information was Hartley (1928). He describes the quantity of information present in a source in terms of its number of possible messages. Later, Shannon (1948) further developed this idea, including new factors such as the effect of noise in the channel, possible economy in the transmission of information, and the possibility that messages might possess distinct quantities of information.

From the quantitative perspective, there can only be information where there is doubt; this, in turn, requires the existence of alternatives, which presupposes the presence of choice, selection, and discrimination. For Hartley (1928), the information in a message is measured by the freedom of choice that someone has in selecting it, based on a source. According to Shannon and Weaver (1949, p. 8-9), "... information in communications theory relates not so much to what you *do* say, as to what you *could* say. That is, information is a measure of one's freedom of choice when one selects a message."

In an unbiased toss of a coin, for example, there are two equally probable possibilities: heads or tails. In an unbiased throw of a dice, there are six possibilities. The degree of freedom of choice in the first case is less than in the second. In the case of the dice, we could say many more things than would be possible in the case of the coin. Hence, from the present perspective, the quantity of information present in the throw of the dice is greater than that present in the toss of the coin.

For Hershberger (1955), information can be defined as a measure of the reduction of uncertainty. It is related to the unpredictability in a message or a source, with the emergence of an element that was absent prior to its occurrence. In the toss of the coin, the reduction of uncertainty is less than in the throw of the dice. The occurrence of an event in a source such as the example of the toss of a coin only eliminates one alternative, while in the throw of a dice five equally probable alternatives are eliminated. Information is also associated with notions such as those of order (or organization, in the terms of Bresciani Filho and D'Ottaviano (1990)) and entropy, as discussed by Alves (2012a).

The central interest in quantitative studies of information generally lies in measurement of the quantity of information in a source, rather than in particular messages. In a broad sense, a *source* can be characterized as a process that generates information. Its constituent elements can be understood as a finite set of events, of messages, of symbols that have a certain probability of occurrence. *Discrete ergodic sources* are those in which every element produced, in addition to being discrete, has the same statistical properties as any other, and its properties remain unaltered with time. Once the probabilities of occurrence of the elements are discovered, it becomes possible to predict its probability of occurrence at any moment, as occurs in the toss of a coin or the throw of a dice.

According to Shannon and Weaver (1949), the *quantity of information of the i-th message* of a source F, denoted by $I_{i(F)}$, is the numerical value defined by:

$$I_{i(F)} =_{df} - \log_2 p_{i(F)},$$

where "$p_{i(F)}$" denotes the probability of occurrence of the *i*-th message of F.

The *quantity of information in a source* F with *n* elements, denoted by H_F, is defined by:

$$H_F =_{df} \Sigma^{n}_{i=1} p_{i(F)} \times I_{i(F)}.$$

If H_C and H_D denote the quantities of information in the throw of the coin and dice, respectively, then $H_C = 1$ and $H_D \approx 2.58$. Since, in both cases, the events corresponding to the messages are equiprobable, it can be demonstrated that the quantity of information in each message in the source is equal to the quantity of information in the source itself.

The greater the freedom of choice and reduction of uncertainty in a source, the more informative it is. Information reaches its maximum value in a source when all its messages have an equal chance of being selected. The informational value is zero when only one of them can occur. It is within this framework that in the next sections we propose our informational perspective of logical consequence.

2 Elements of an axiomatic theory of probability

We employ probability in cases where two or more different results can occur in a given circumstance. This means that the result is not predictable (or is indeterminate), in the sense that it is not possible to previously determine which result might occur at a given moment.

Probability theory, henceforth denoted $\mathbb{P}$, studies random experiments. As a basis for $\mathbb{P}$, we shall use the Zermelo-Fraenkel set theory with the Choice Axiom (ZFC), with the usual elementary arithmetic theory (cf. ENDERTON, 1977). Hence, the language (alphabet and definitions) and theorems of ZFC will also be considered elements of $\mathbb{P}$. The only symbols belonging to the alphabet of $\mathbb{P}$ are the symbols A_i, for $0 \leq i \leq n$ and $i \in \mathbb{N}$, where "A" represents the primitive concept of $\mathbb{P}$ known as *happening*, *result* or *occurrence*, used for the definition of random experiment.

Definition 2.1 (Random experiment)

a: A *random experiment* or *random phenomenon*, denoted by "Σ", is one that, repeated various times, presents different results or occurrences, called *results of Σ* or *occurrences of Σ*, denoted by "$A_i(\Sigma)$".

b: The *sample space* of a random experiment Σ, denoted by "U(Σ)", is the set of all possible results of Σ.

c: The *number of elements* of the sample space, denoted by "n(U(Σ))", in which $n(U(\Sigma)) > 0$ is finite, is the quantity of elements of "U(Σ)".

d: A sample space is *equiprobable* when all its elements have the same chance of occurring.

The toss of a coin, the throw of a dice, and the removal of a card from a pack are examples of random experiments. Their results would be the fall of the coin with one or other face upwards, the fall of the dice with one of the numbers from one to six upwards, and the withdrawal of one of the cards of the pack, respectively.

Definition 2.2 (Event of a random experiment)

a: An *event* of a random experiment Σ, denoted by "E(Σ)", is any subset of the sample space U(Σ), in other words, $E(\Sigma) \subseteq U(\Sigma)$.

b: The *number of elements* of an event $E(\Sigma)$, denoted by "$n(E(\Sigma))$", is the quantity of elements of $U(\Sigma)$ belonging to $E(\Sigma)$.

c: An *elementary event E of* Σ is that where $n(E(\Sigma)) = 1$.

d: The *correct event E of* Σ is that where $n(E(\Sigma)) = n(U(\Sigma))$.

e: The *impossible event E of* Σ, denoted by "$\varnothing$", is that where $n(E(\Sigma)) = 0$.

f: A *contingent event E of* Σ is that where $n(\varnothing) < n(E(\Sigma)) < n(U(\Sigma))$.

g: The *event* E_i *of* Σ*, complementary to* $E(\Sigma)$, denoted by "$\bar{E}(\Sigma)$", is defined by:

$$\bar{E}(\Sigma) =_{df} \{A \in U(\Sigma) \mid A \notin E(\Sigma)\}.$$

h: The *event E of* Σ*, the union of* $E_i(\Sigma)$ *and* $E_j(\Sigma)$, denoted by "$(E_i(\Sigma) \cup E_j(\Sigma))$", is defined by:

$$(E_i(\Sigma) \cup E_j(\Sigma)) =_{df} \{A \in U(\Sigma) \mid A \in E_i(\Sigma) \text{ ou } A \in E_j(\Sigma)\}.$$

i: The *event E of* Σ*, the intersection of* $E_i(\Sigma)$ *and* $E_j(\Sigma)$, denoted by "$(E_i(\Sigma) \cap E_j(\Sigma))$", is defined by:

$$(E_i(\Sigma) \cap E_j(\Sigma)) =_{df} \{A \mid A \in E_i(\Sigma) \text{ e } A \in E_j(\Sigma)\}.$$

Henceforth, when there is no risk of ambiguity, we shall remove the references to Σ between parentheses of the notations. Hence, instead of $A(\Sigma)$, $U(\Sigma)$, or $n(U(\Sigma))$, we shall use only A, U, or n(U), respectively.

Example 2.3 (Random experiments)

$\boldsymbol{\Sigma_i}$	$\boldsymbol{U(\Sigma_i)}$	$\boldsymbol{E(\Sigma_i)}$	$\boldsymbol{n(E(\Sigma_i))}$
$\boldsymbol{\Sigma}_1$ (Toss of coin)	{H, T}	$E_1(\boldsymbol{\Sigma}_1)$: {H} (Fall head) $E_2(\boldsymbol{\Sigma}_1)$: {T} (Fall tail)	1 1
$\boldsymbol{\Sigma}_2$ (Toss of biased coin)	{H_1, H_2, H_3, T}	$E_1(\boldsymbol{\Sigma}_2)$: {$H_1$, H_2, H_3} (Fall head) $E_2(\boldsymbol{\Sigma}_2)$: {T} (Fall tail)	3 1
$\boldsymbol{\Sigma}_3$ (Throw of dice)	{1, 2, 3, 4, 5, 6}	$E_1(\boldsymbol{\Sigma}_3)$: {2, 4, 6} (Fall even) $E_2(\boldsymbol{\Sigma}_3)$: $\varnothing$ (Fall head)	3 0

In the above example, a "model" is proposed for each random experiment, its sample space, and some of its events, associating them with entities of a "world". In the third column, in parentheses, we give the common name for each event in order to express the results that comprise it.

Although our theoretical approach does not consider qualitative elements of information, the examples suggested throughout this paper involve content, in order to aid understanding. Meanwhile, the same results could have been obtained using a purely formal mode of construction.

The notion of random experiment can be compared with that of a discrete ergodic source of information, as indicated in the Introduction section of this work. In a way that is similar to this type of source, the elements belonging to a random experiment must be previously defined with precision. Furthermore, every sample space considered, in addition to being finite, must be equiprobable, and the probability values of the events must be given

and fixed. This will enable us to determine with precision the existence of the relation of logical consequence between formulae (as shown in Section 4).

Definition 2.4 (Probability of an event)
The *probability of occurrence* of an event E in the random experiment Σ with an equiprobable sample space U(Σ), denoted by "p(E(Σ))", is the numerical value defined by:

$$p(E(\Sigma)) =_{df} \frac{n(E(\Sigma))}{n(U(\Sigma))}.$$

The *probability function*, p, provides events of a random experiment with values between 0 and 1; in other words, p: $E(\Sigma) \rightarrow [0,1] \subseteq \mathbb{Q}$. The *probability value* of an event $E(\Sigma)$ is given by $p(E(\Sigma))$.

Example 2.5 (Probability of the events of Example 2.3)

E	$E_1(\Sigma_1)$	$E_2(\Sigma_1)$	$E_1(\Sigma_2)$	$E_2(\Sigma_2)$	$E_1(\Sigma_3)$	$E_2(\Sigma_3)$	$E_1(\Sigma_3) \cup E_2(\Sigma_3)$	$E_1(\Sigma_2) \cap E_2(\Sigma_2)$	$\bar{E}_1(\Sigma_2)$
p(E)	½	½	¾	¼	½	0	½	0	¼

Having constituted the basic elements of the language of ℙ, we now describe its axioms and some of its elementary results.

The axioms for ℙ are as follows:

(Ax$\mathbb{P}$1): $p(E) \geq 0$, for every $E \subseteq U$
(Ax$\mathbb{P}$2): $p(E_i \cup E_j) = p(E_i) + p(E_j) - p(E_i \cap E_j)$
(Ax$\mathbb{P}$3): $p(E \cup \bar{E}) = 1$

In the next theorem, some results concerning the probability of events are stated, being important to the continuation of this paper. For simplicity, we do not present the proof of this, and some other theorems. Their proofs can be found in Alves (2012b).

Theorem 2.6
a: $p(E) \leq 1$, for every $E \subseteq U$.
b: $p(U) = 1$.
c: $p(\varnothing) = 0$.
d: $n(E_i \cap E_j) = 0 \Rightarrow p(E_i \cup E_j) = p(E_i) + p(E_j)$.
e: $p(E \cap \bar{E}) = 0$.
f: $p(\bar{E}) = 1 - p(E)$.
g: $\sum^{m}_{i=1} p(E_i) = 1$, for $E_i = \{A_i\}$, for $1 \leq i \leq m$, with $U = \{A_1, ..., A_m\}$. ■

Based on ℙ, in what follows we develop a probabilistic semantics for classical propositional logic (CPL).

3 A probabilistic semantics for languages of CPL

We call this perspective the *probabilistic semantics for CPL* (henceforth, $S_{\mathbb{P}}$). As shown by Alves (2012b), the behavior of $S_{\mathbb{P}}$ is not strictly equivalent to the behavior of the

usual classical veritative-functional semantics (henceforth, $S_{\mathcal{V}}$), as developed by Tarski (1956a), Mendelson (1964), Shoenfield (1967), and Mates (1972), amongst others.

We associate the formulae of a CPL language, denoted by "*L*", with the events of a random experiment, Σ. We define some of the notions that are fundamental to the objectives of this paper, and describe some results characteristic of $S_{\mathbb{P}}$.

The expression "Form(*L*)" denotes the set of formulae of a language *L*, the letters "φ", "ψ", and "γ" are metalinguistic variables that represent elements of Form(*L*); "P_0", "P_1", "P_2" etc are the atomic formulae of *L*, and "Γ" represents any finite subset of Form(*L*). We adopt negation and disjunction as primitive logical connectives of *L*.

Definition 3.1 (Situation for a language *L*)
A function *f* is a *Σ-situation for L*, or simply a *situation*, denoted by "*f*(Σ)", if *f*(Σ): Form(*L*) $\rightarrow \wp(U(\Sigma))$, such that:

a: If φ is atomic, then $f(\Sigma)(\varphi) = E(\Sigma)$, defined by *f* itself;
b: If φ is of the form $\neg\psi$, then $f(\Sigma)(\varphi) = \overline{f(\Sigma)(\psi)}$;
c: If φ is of the form $\psi \vee \gamma$, then $f(\Sigma)(\varphi) = f(\Sigma)(\psi) \cup f(\Sigma)(\gamma)$.

A *situation* for *L* consists of an attribution of a single event of a given random experiment to each well-formed formula of *L*, according to a function *f*. The fact that a situation is defined using a function enables us to avoid ambiguities in the next definitions, especially in the case of informational logical consequence.

To say that *f is an Σ-situation for L* is the same as saying that the random experiment *Σ is an f-structure for L*.

Although each formula of *L* is associated with a single event in a given situation *f*(Σ), distinct formulae can be associated with the same event in *f*(Σ). This always occurs, given that the set of formulae of any *L* is infinite, in contrast to the number of events of a random experiment, which is always finite.

Definition 3.2 (Probability of formulae according to a situation)
The *probability function of a formula* φ according to f(Σ), denoted by "P(f(Σ))", is defined as follows, where "p" is the probability function concerning events as defined in ℙ:

a: If φ is atomic, $P(f(\Sigma))(\varphi) =_{df} p(f(\Sigma)(\varphi))$;
b: If φ is of the form $\neg\psi$, $P(f(\Sigma))(\varphi) =_{df} p(\overline{f(\Sigma)(\psi)})$;
c: If φ is of the form $\psi \vee \gamma$, $P(f(\Sigma))(\varphi) =_{df} p(f(\Sigma)(\psi) \cup f(\Sigma)(\gamma))$.

It can be seen that P(*f*(Σ)): Form(*L*) $\rightarrow [0,1] \subseteq \mathbb{Q}$, such that p(*f*(Σ))(φ)) is the image of φ according to P(*f*(Σ)). The probability value of the formula φ according to f(Σ) is given by P(*f*(Σ))(φ), a rational number between 0 and 1. It is the probability value of the event of Σ corresponding to φ, according to *f*(Σ).

It can be shown that P(*f*(Σ)) satisfies the properties described in Theorem 2.6, interpreted in the light of $S_{\mathbb{P}}$.

Henceforth, when we say "for every *f*", we mean "for every situation *f*(Σ), given Σ". In addition, when there is no possibility of ambiguity or imprecision, we shall use "P(φ)" as an abbreviation for "P(*f*(Σ))(φ)", "p(φ)" as an abbreviation for "p(*f*(Σ)(φ)", "*f*(φ)" as an abbreviation for "*f*(Σ)(φ)", and "*f*" as an abbreviation for "*f*(Σ)".

In accordance with the usual definitions in L, we have:

a: $f(\varphi \wedge \psi) = f(\neg(\neg\varphi \vee \neg\psi))$
b: $f(\varphi \rightarrow \psi) = f(\neg\varphi \vee \psi)$
c: $f(\varphi \leftrightarrow \psi) = f((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$
d: $P(\varphi \wedge \psi) = P(\neg(\neg\varphi \vee \neg\psi))$
e: $P(\varphi \rightarrow \psi) = P(\neg\varphi \vee \psi)$
f: $P(\varphi \leftrightarrow \psi) = P((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$.

Example 3.3:
(Probability of formulae in Σ_1 and Σ_2, for $U(\Sigma_1) = \{1,2,3,4,5,6\}$ and $U(\Sigma_2) = \{H, T\}$)

φ	$f(\Sigma_1)(\varphi)$	$P(f(\Sigma_1))(\varphi)$	$f(\Sigma_2)(\varphi)$	$P(f(\Sigma_2))(\varphi)$
P_1	{2,4,6}	½	{H}	½
P_2	{1,3,5}	½	{T}	½
P_3	{1,2,3,5}	2/3	∅	0
P_4	{1}	1/6	∅	0
P_5	∅	0	∅	0
$P_1 \wedge P_2$	∅	0	∅	0
$P_1 \wedge P_3$	{2}	1/6	∅	0
$P_2 \wedge P_3$	{1,3,5}	½	∅	0
$P_1 \vee P_3$	{U}	1	{H}	½
$\neg(P_1 \vee P_3)$	∅	0	{T}	½
$P_1 \rightarrow P_2$	{1,3,5}	½	{T}	½

Theorem 3.4: For every f, we have:
a: $P(\neg\varphi) = 1 - P(\varphi)$.
b: $f(\varphi \wedge \psi) = f(\varphi) \cap f(\psi)$.
c: $P(\varphi \wedge \psi) = p(f(\varphi) \cap f(\psi))$.
d: $P(\varphi \rightarrow \psi) = p(\overline{f(\varphi)} \cup f(\psi))$. ■

Definition 3.5 (Inconsistency, validity, and contingency of formulae of L in $S_{\mathbb{P}}$)
a: A formula φ is *probabilistically inconsistent* or *probabilistically contradictory*, which is denoted by "$\perp_{\mathbb{P}}$", if, for every $f(\Sigma)$, $P(f(\Sigma))(\varphi) = 0$.
b: A formula φ is *probabilistically valid* or *probabilistically tautological*, which is denoted by "$\top_{\mathbb{P}}$", if, for every $f(\Sigma)$, $P(f(\Sigma))(\varphi) = 1$.
c: A formula φ is *probabilistically contingent* if it is neither probabilistically contradictory nor probabilistically valid.

Definition 3.6 (Probabilistically equivalent formulae)
Two formulae, φ and ψ, are *probabilistically equivalent*, which is denoted by "$\varphi \equiv \psi$", if, for every situation $f(\Sigma)$, $P(f(\Sigma))(\varphi) = P(f(\Sigma))(\psi)$.

Definition 3.7 (Probability value of a set of formulae)
Let $\Gamma = \{\varphi_1, ..., \varphi_n\} \subseteq \text{Form}(L)$. The *probability value of Γ according to $f(\Sigma)$*, denoted by "$P^*(f(\Sigma))(\Gamma)$", is defined by:

$$P^*(f(\Sigma))(\Gamma) =_{df} P(f(\Sigma))(\varphi_1 \wedge ... \wedge \varphi_n).$$

If $\Gamma = \varnothing$, then $P^*(f(\Sigma))(\Gamma) = 1$. When Γ is infinite, $P^*(f(\Sigma))(\Gamma)$ is indefinite.

It can be seen that $P^*(f(\Sigma))$: $\wp(\text{Form}(L)) \rightarrow [0,1] \subseteq \mathbb{Q}$, such that $P^*(f(\Sigma))(\Gamma))$ is the image of Γ according to $P^*(f(\Sigma))$. The *probability value of the set of formulae Γ according to $f(\Sigma)$* is given by $P^*(f(\Sigma))(\Gamma)$, a rational number between 0 and 1. When there is no risk of ambiguity or imprecision, we shall write $P(\Gamma)$ instead of $P^*(f(\Sigma))(\Gamma)$.

According to Definition 3.7, the probability value of a given set of formulae of L in a situation f is, in the final analysis, defined from the probability value of the intersection of the events associated with the elements of Γ. In other words, by Definition 3.7, and Theorem 3.4c, we have that $P^*(f(\Sigma))(\Gamma) = P(f(\Sigma))(\varphi_1 \wedge ... \wedge \varphi_n) = p(f(\Sigma)(\varphi_1) \cap ... \cap f(\Sigma)(\varphi_n))$. Given the definitions in question, we can say that $f(\Sigma)(\Gamma) = f(\Sigma)(\varphi_1 \wedge ... \wedge \varphi_n)$.

Strictly speaking, it is not possible to define the probability value of an empty set of formulae. Definition 3.2, which forms the basis of Definition 3.7, only applies to the probability values of formulae of language L. But $\varnothing$ is not a formula of L. Although a formula might be associated with the empty event, with its probability value being defined as zero, it makes no sense to say that an empty formula can have any probability value.

In order not to leave the empty set without a definition of the probability value, it was decided to define it arbitrarily in the way described above. This decision can be justified as follows: in this case, to say that $P(\Gamma) \neq 1$ would signify the existence of some $\varphi_i \in \Gamma$ such that $P(\varphi_i) \neq 1$. Since there are no formulae in Γ with this value, given the inexistence of formulae in Γ, then $P(\Gamma) = 1$.

Definition 3.8 (Probabilistic logical consequence)
A formula φ is *probabilistic logical consequence* of a set Γ of formulae, which is denoted by "$\Gamma \vDash_{\mathbb{P}} \varphi$", if, for every situation $f(\Sigma)$, $P(f(\Sigma))(\Gamma) \leq P(f(\Sigma))(\varphi)$.

When $\Gamma = \varnothing$, instead of $\Gamma \vDash_{\mathbb{P}} \varphi$ we simply write $\vDash_{\mathbb{P}} \varphi$. The expression "$\Gamma \nvDash_{\mathbb{P}} \varphi$" denotes that φ is not probabilistic logical consequence of Γ. The formulae of Γ are called *premises* and φ is termed *conclusion*.

Despite possessing certain specific characteristics, $S_{\mathbb{P}}$ is a semantics for CPL. The formulae considered valid by $S_{\mathbb{P}}$ are exactly the same as those considered valid by $S_{\mathcal{V}}$. Furthermore, $\Gamma \vDash_{\mathbb{P}} \varphi$ if, and only if, $\Gamma \vDash_{\mathcal{V}} \varphi$. The similarities and differences between $S_{\mathbb{P}}$ and $S_{\mathcal{V}}$ are described by Alves (2012b). In the following discussion, we address the notion of informational logical consequence.

4 Informational logical consequence

In this section, we propose a quantitative-informational definition of logical consequence and present some of its properties. We first introduce notions such as the quantity of information present in a formula of L and in a set of formulae, based on the notion of quantity of information developed by Shannon and Weaver (1949).

Definition 4.1 (Quantity of information of a formula according to a situation)
The *quantity of information* or *informational value of a formula φ of L according to a situation $f(\Sigma)$*, denoted by "$I(f(\Sigma))(\varphi)$", is the numerical value defined by:

$$I(f(\Sigma))(\varphi) =_{df} - \log_2 P(f(\Sigma))(\varphi).$$

When $P(f(\Sigma))(\varphi) = 0$, we define that $\log_2 0 = 0$, in other words, $I(f(\Sigma))(\varphi) = 0$. When there is no risk of ambiguity or imprecision, we shall use "$I(\varphi)$" instead of "$I(f(\Sigma))\varphi$". It can be shown that $I(f(\Sigma))$: $\text{Form}(L) \to \mathbb{Q}_+$.

Example 4.2 (Quantity of information in formulae based on the two situations of Example 3.3)

φ	$f(\Sigma_1)(\varphi)$	$I(f(\Sigma_1))\varphi$	$f(\Sigma_2)(\varphi)$	$I(f(\Sigma_2))\varphi$
P_1	{2,4,6}	1	{H}	1
P_2	{1,3,5}	1	{T}	1
P_3	{1,2,3,5}	0.58	$\varnothing$	0
P_4	{1}	2.58	$\varnothing$	0
$\neg P_1$	{1,3,5}	1	{T}	1
$P_1 \wedge P_3$	{2}	2.45	$\varnothing$	0
$P_1 \vee P_3$	U	0	{H}	1
$P_1 \to P_2$	{1,3,5}	1	{T}	1
$P_1 \to P_3$	{1,2,3,5}	0.58	{T}	1
$P_2 \to P_3$	U	0	{H}	1

Definition 4.3 (Quantity of information in a set of formulae)
Let $\Gamma = \{\varphi_1, ..., \varphi_n\} \subseteq \text{Form}(L)$. The *informational value of* Γ (*quantity of information in* Γ), *according to* $f(\Sigma)$, denoted by "$I^*(f(\Sigma))(\Gamma)$", is defined by:

$$I^*(f(\Sigma))(\Gamma) =_{df} -\log_2 P^*(f(\Sigma))(\Gamma).$$

When there is no risk of ambiguity or imprecision, we shall use $I(\Gamma)$ instead of $I^*(f(\Sigma))(\Gamma)$. Although the domain of this function may be different from that of the information function concerning formulae, its image set is the same; in other words, $I^*(f(\Sigma))$: $\wp(\text{Form}(L)) \to \mathbb{Q}_+$.

Definition 4.4 (Informational logical consequence)
A formula φ is *informational logical consequence* of a set Γ of sentences, which is denoted by "$\Gamma \models \varphi$", if, for every $f(\Sigma)$, $I(f(\Sigma))(\Gamma) \geq I(f(\Sigma))(\varphi)$.

When $\Gamma = \varnothing$, instead of "$\Gamma \models \varphi$", we simply write "$\models \varphi$"; "$\Gamma \not\models \varphi$" denotes that φ is not informational logical consequence of Γ. The formulae of Γ are called *premises* and φ is termed *conclusion*.

According to the above definition, a formula is informational logical consequence of a given set of formulae if, and only if, the quantity of information present in the conclusion is never greater than the quantity of information in the premises. In probabilistic logical consequence, the relation is inverse.

Theorem 4.5

a: $\varphi \to \psi \models \neg\varphi \vee \psi$.
b: $\varphi \vee \psi \models \neg(\neg\varphi \wedge \neg\psi)$.
c: $\varphi \wedge \psi \models \neg(\neg\varphi \vee \neg\psi)$.
d: $\varphi \wedge \psi \models \psi \wedge \varphi$.
e: $\varphi \vee \psi \models \psi \vee \varphi$.
f: $\neg\neg\varphi \models \varphi$.

g: $\varphi \wedge \varphi \Vdash \varphi$.
h: $\varphi \vee \varphi \Vdash \varphi$.
i: $\top_{\mathbb{P}} \wedge \varphi \Vdash \varphi$.
j: $\top_{\mathbb{P}} \vee \varphi \Vdash \top_{\mathbb{P}}$.
k: $\bot_{\mathbb{P}} \wedge \varphi \Vdash \bot_{\mathbb{P}}$.
l: $\bot_{\mathbb{P}} \vee \varphi \Vdash \varphi$.
m: $\top_{\mathbb{P}} \Vdash \bot_{\mathbb{P}}$.
n: $\varphi \Vdash \varphi$. ■

Since the quantity of information in the premise and conclusion of each one of the items of the above theorem is the same for each given situation, the reciprocal of each one of the items is also valid, such that one formula is informational logical consequence of the other.

The first three items of the above theorem show that the definitions of one connective, obtained from others, are maintained in informational logical consequence. The first item shows that the notion of implication presupposed here is that of material implication. The fourth and sixth items show, respectively, that the logic underlying informational logical consequence is neither temporal logic nor intuitionistic logic.

Theorem 4.6

a: $\Vdash \varphi \Leftrightarrow I(\varphi) = 0$, for every *f*.
b: $\Vdash\varphi \Leftrightarrow \Gamma \Vdash \varphi$, for every Γ.
c: $I(\Gamma) = 0$, for every *f*, and $\Gamma \Vdash \varphi \Rightarrow \Vdash \varphi$.
d: $\vDash_{\mathbb{P}} \varphi \Rightarrow \Vdash \varphi$. ■

Proposition 4.7

a: $\varphi \wedge \psi \nVdash \varphi$.
b: $\varphi \rightarrow \psi, \varphi \nVdash \psi$.
c: $\varphi \rightarrow \psi, \neg\psi \nVdash \neg\varphi$.
d: $\varphi \rightarrow \psi, \psi \rightarrow \gamma \nVdash \varphi \rightarrow \gamma$.
e: $\varphi \nVdash \psi \rightarrow \varphi$.
f: $\psi \nVdash \varphi \vee \psi$.
g: $\gamma \vee \varphi, \neg\gamma \vee \psi \nVdash \varphi \vee \psi$. ■

The above proposition shows that the rules of inference of a large part of formal logical systems are not valid in informational terms. The second item is the rule of *modus ponens*, adopted in systems such as that of Mendelson (1967); the last two items are the expansion rule and the cut rule, adopted by Shoenfield (1969). The same can be said when the above items are treated as arguments. Arguments that are traditionally considered valid can present more information in the conclusion than in the set of premises, which means that, according to the perspective in question, they act to amplify information.

As shown by Alves (2012b), the invalidity of these rules of inference or arguments is generally due to the possibility of the set of premises being informationally empty in a given situation. In the case of the expansion rule, when, in a given situation Σ, $P(\psi) = 0$ and $I(\varphi) \neq 0$, then $I(\psi) < I(\varphi \vee \psi)$. Hence, $\psi \nVdash \varphi \vee \psi$. While the premise is informationally empty, the conclusion produces novelty and reduces uncertainty. For example, in the game of dice, the sentence "it fell on number seven", for which the probability of occurrence is zero, would not reduce the uncertainty about what occurred in the game. Meanwhile, the sentence "it fell on an even number or on number seven" possesses a quantity of information that is greater than zero, since it reduces uncertainty: the dice could have fallen on numbers two, four, or six, eliminating the possibility of having fallen on an odd number. Thus, in the conclusion, there is something that did not exist in the premise. There was an informational gain, given that the quantity of information in the premise was null.

In the case of *modus ponens*, the situation $f(\Sigma_1)$ of Example 4.2 above provides an example in which the informational value of the premises, interpreted as "if it falls on evens, then it falls on odds" and "it falls on evens", is less than the informational value of the conclusion, interpreted as "it falls on odds". Here, the informational value of the premises is null. Since the events "fall on odds" and "fall on evens" are mutually exclusive, the sentence "if it falls on evens, then it falls on odds" is equivalent to "it falls on odds or it falls on odds", which is equivalent to "it falls on odds". The probability of the set of premises is therefore defined from the union of "it falls on odds" and "it falls on evens", equivalent to "it falls on odds and does not fall on odds", for which the probability is zero. Meanwhile, the conclusion possesses a quantity of information that is greater than zero, since it possesses an informational value that is greater than the value of the set of premises.

The example discussed in the preceding paragraph, which is an individual case of Theorem 4.8b, outlined below, seems to fit an intuitive notion of informational logical consequence. Intuitively, "fall on odd and not fall on odd" provides no information concerning a circumstance. Hence, any probabilistically contingent conclusion can contain more information than is contained by the premises.

Several steps of the demonstrations of the following theorems have been omitted. These steps, indicated by "TPP", refer to theorems previously proved in Alves (2012b).

Theorem 4.8: Let $\Gamma = \{\varphi_1, ..., \varphi_n\}$. Then:

a: $\Gamma \Vdash \varphi$ if and only if $I(\varphi) = 0$, for every f, or $(\varphi_1 \wedge ... \wedge \varphi_n) \equiv \varphi$;

b: If $I(\varphi) > 0$, for a given f, then $\bot_{\mathbb{P}} \nVdash \varphi$;

c: $\Gamma \Vdash \top_{\mathbb{P}}$;

d: $\Gamma \Vdash \bot_{\mathbb{P}}$;

e: If $I(\varphi) > 0$, for a given f, then $\top_{\mathbb{P}} \nVdash \varphi$.

Proof

a: **(⇒):** Let $\Gamma \Vdash \varphi$. Suppose that $I(\varphi) \neq 0$, for a given f, and $(\varphi_1 \wedge ... \wedge \varphi_n) \not\equiv \varphi$. It is then possible to show the existence of f' such that $I(f')(\Gamma) < I(f')(\varphi)$, contradicting the initial hypothesis: $I(f')(\Gamma) = 0$ and $I(f')(\varphi) = I(f)(\varphi)$. Hence, when $\Gamma \Vdash \varphi$ we have that if $I(\varphi) \neq 0$, for a given f, then $(\varphi_1 \wedge ... \wedge \varphi_n) \equiv \varphi$, in other words, $I(\varphi) = 0$, for every f, or $(\varphi_1 \wedge ... \wedge \varphi_n) \equiv \varphi$.

(⇐): Case 1: Let $I(\varphi) = 0$, for every f. Then, by TPP, $I(\Gamma) \geq I(\varphi)$, for every f and every Γ. So, by Definition 4.4, $\Gamma \Vdash \varphi$.

Case 2: By Definition 3.6, $(\varphi_1 \wedge ... \wedge \varphi_n) \equiv \varphi$ if and only if $P(\varphi_1 \wedge ... \wedge \varphi_n) = P(\varphi)$, for every f. By Definition 3.7, $P(\varphi_1 \wedge ... \wedge \varphi_n) = P(\varphi)$, for every f if and only if $P(\Gamma) = P(\varphi)$, for every f. Then, by TPP, $I(\Gamma) = I(\varphi)$, for every f. So, by Definition 4.4, $\Gamma \Vdash \varphi$.

b: Let $P(\varphi) = ½$. Then $I(\bot_{\mathbb{P}}) = 0 < I(\varphi) = 1$.

c: Since $I(\top_{\mathbb{P}}) = 0$, then $I(\Gamma) \geq I(\top_{\mathbb{P}})$.

d: Since $I(\bot_{\mathbb{P}}) = 0$, then $I(\Gamma) \geq I(\bot_{\mathbb{P}})$.

e: Let $P(\varphi) = ½$. Then $I(\top_{\mathbb{P}}) = 0 < I(\varphi) = 1$. ■

The first item above describes the arguments that are valid according to the informational perspective of logical consequence.

The second item explains that a probabilistically contradictory formula cannot lead informationally to any informative formula. In fact, only formulae that are probabilistically

valid or invalid, in other words not informative, are informational logical consequence of contradictory formulae.

Presupposing the distinction between classical and non-classical systems, as suggested by Da Costa (1993; 1997), or by Haack (1978), we can conclude from the second item of the above theorem that *classical logic is not the logic underlying informational logical consequence*. This is because, in classical formal logical systems, a contradiction generates any formula. Considering that complementary logics, such as modal logic, retains the same principles of classical logic, we can also conclude that *no complementary logic underlies informational logical consequence*. The remaining candidates are heterodox logics, such as intuitionistic and paraconsistent systems, as described by D'Ottaviano (1992).

In intuitionistic logics, negation possesses certain particular characteristics. Such characteristics do not permit, for example, recourse to proofs employing reduction to the absurd, given that formulae such as $\varphi \leftrightarrow \neg\neg\varphi$ are not valid in these systems. Meanwhile, it can be shown that $\neg\neg\varphi \Vdash \varphi$ and $\varphi \Vdash \neg\neg\varphi$ and, from Theorem 4.9b below, we have $\Vdash \varphi \leftrightarrow \neg\neg\varphi$. Therefore, *intuitionistic logic cannot provide a basis for informational logical consequence*.

Theorem 4.8, especially the second item, indicates that the logic underlying informational logical consequence is, at least, *paraconsistent sensu lato*. This is because this notion of consequence does not permit, for example, the *Ex Falso Quodlibet*, also known as the *Explosion Principle*, such that from a contradiction does not follow any formula.

The third item of the above theorem is shared by both probabilistic and veritative-functional logical consequence, but for different reasons. Informationally, a probabilistically valid formula is logical consequence of any formula, because its value is minimal; hence, it cannot provide more information than any other formula. In probabilistic and veritative-functional terms, it is logical consequence of any formula because its value is maximal; in other words, it possesses the probabilistic value '1' or the true value 'V'.

The fourth item, like the second, is not valid in the other two notions of logical consequence considered here. In these notions, an inconsistent formula is logical consequence solely of a contradictory set of formulae. In the informational version, it is logical consequence of any set of formulae. In informational terms, there is no distinction between valid and inconsistent formulae.

The fifth result expresses a similarity between informational logical consequence and the other notions of consequence: contingent formulae are not logical consequence of valid formulae.

Theorem 4.9

a: If $\varphi \Vdash \psi$, and $\psi \Vdash \gamma$, then $\varphi \Vdash \gamma$.

b: If $\varphi \equiv \psi$, then $\Vdash \varphi \leftrightarrow \psi$.

c: Is not the case that: $\Gamma, \varphi \Vdash \psi$ if and only if $\Gamma \Vdash \varphi \rightarrow \psi$.

d: Is not the case that:$\Vdash \varphi$ e $\Vdash \psi$ if and only if $\varphi \equiv \psi$.

e: Is not the case that:$\Gamma \Vdash \varphi$ if and only if $f(\Gamma) \subseteq f(\varphi)$, for every f.

f: Is not the case that:$\Gamma \Vdash \varphi$ if and only if $\Gamma \vDash_{\mathcal{V}} \varphi$.

Proof

a: Let $\varphi \Vdash \psi$, and $\psi \Vdash \gamma$. By Definition 4.4, $I(\varphi) \geq I(\psi) \geq I(\gamma)$, for every *f*. Then, by TPP, $I(\varphi) \geq I(\gamma)$, for every *f*. So, by Definition 4.4, $\varphi \Vdash \gamma$.

b: By TPP, if $\varphi \equiv \psi$, then $\varphi \leftrightarrow \psi$ is $\top_{\mathbb{P}}$. And, by TPP, if $\varphi \leftrightarrow \psi$ is $\top_{\mathbb{P}}$, then $\vDash_{\mathbb{P}} \varphi \leftrightarrow \psi$. So, by Theorem 4.6d, $\Vdash \varphi \leftrightarrow \psi$.

c: ($\nRightarrow$): Let $\Gamma = \varnothing$. $\bot_{\mathbb{P}} \nVdash \psi$, but $\Vdash \bot_{\mathbb{P}} \rightarrow \psi$.
($\nLeftarrow$): $\nVdash \psi \rightarrow \bot_{\mathbb{P}}$, but $\psi \Vdash \bot_{\mathbb{P}}$.

d: ($\nRightarrow$): $\Vdash \bot_{\mathbb{P}}$ e $\Vdash \top_{\mathbb{P}}$, but $\bot_{\mathbb{P}} \not\equiv \top_{\mathbb{P}}$.
($\nLeftarrow$): $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi$, but $\vDash_{\mathbb{P}} \varphi \rightarrow \psi$ and $\vDash_{\mathbb{P}} \neg\varphi \vee \psi$ do not occur.

e: ($\nRightarrow$): $\top_{\mathbb{P}} \Vdash \bot_{\mathbb{P}}$, but $f(\top_{\mathbb{P}}) = U \not\subseteq f(\bot_{\mathbb{P}}) = \varnothing$.
($\nLeftarrow$): Let $f(\varphi) \neq \varnothing$. Then $f(\bot_{\mathbb{P}}) \subseteq f(\varphi)$, but $\bot_{\mathbb{P}} \nVdash \varphi$.

f: ($\nRightarrow$): $\varphi \vee \neg\varphi \Vdash \varphi \wedge \neg\varphi$, but $\varphi \vee \neg\varphi \nVdash_{\mathcal{V}} \varphi \wedge \neg\varphi$.
($\nLeftarrow$): $\psi \vDash_{\mathcal{V}} \varphi \vee \psi$, but $\psi \nVdash \varphi \vee \psi$. ■

The above theorem expresses some of the characteristic properties of informational logical consequence, when compared to the probabilistic and veritative-functional versions. In contrast to the latter two versions, the reciprocal of Theorem 4.9b, especially, cannot be shown for the informational version; in other words, is not the case that if $\Vdash \varphi \leftrightarrow \psi$, then $\varphi \equiv \psi$, given that, for example, $\Vdash \bot_{\mathbb{P}} \leftrightarrow \top_{\mathbb{P}}$, but $\bot_{\mathbb{P}} \not\equiv \top_{\mathbb{P}}$.

Theorem 4.9c shows the invalidity of the corresponding semantics of the Deduction Theorem. The fourth and fifth items illustrate some of the innate characteristics of informational logical consequence, as discussed in the final considerations (below). Theorem 4.9f expresses a distinction between the relations of informational and veritative-functional logical consequence.

We show below that informational logical consequence is not a Tarskian logical consequence, according to the definition set out in the introduction of this paper.

Theorem 4.10

a: Is not the case that: if $\varphi \in \Gamma$, then $\Gamma \Vdash \varphi$.

b: Is not the case that: if $\Gamma \subseteq \Delta$ and $\Gamma \Vdash \varphi$, then $\Delta \Vdash \varphi$.

c: If $\Delta \Vdash \psi$, for each $\psi \in \Gamma$, and $\Gamma \Vdash \varphi$, then $\Delta \Vdash \varphi$.

Proof: Let *f* be any situation.

a: Let $P(\Gamma) = 0$ and $P(\varphi) = ½$. Then, by TPP, $I(\Gamma) < I(\varphi)$.

b: Let $\Gamma = \{\varphi\}$, $\Delta = \{\varphi, \neg\varphi\}$ and $I(\varphi) > 0$. Then, by TPP, $\Gamma \Vdash \varphi$ and $\Delta \nVdash \varphi$.

c: Case 1: Let $I(\Delta) = 0$. Then, by Hypothesis, $I(\psi) = 0$, for each $\psi \in \Gamma$. By TPP, $I(\Gamma) = 0$. Thus, by Hypothesis, $I(\varphi) = 0$. So, $I(\Delta) = I(\varphi)$.
Case 2: Let $I(\Delta) \neq 0$. By TPP, and by Hypothesis, $P(\Delta) \leq P(\psi)$, for each $\psi \in \Gamma$. In this case, it can be shown that $f'(\Delta) \subseteq f'(\Gamma) \subseteq f'(\varphi)$, for every f'. Then, by TPP and Definition 3.8, $P(\Delta) \leq P(\Gamma) \leq P(\varphi)$. So, by TPP, $I(\Delta) \geq I(\varphi)$. ■

Thus, informational logical consequence is neither reflexive nor monotonic, although it may satisfy the condition of transitivity.

According to our objectives in this paper, we finish our presentation of informational logical consequence. Others results and properties can be found in Alves (2012b).

Final considerations

The shift of perspective in the analysis of the notion of logical consequence, from the true value to the quantity of information of formulae, enables characteristics to emerge that are uniquely attributable to this approach. Alves (2012b) presents some of the main similarities and differences between this concept and the usual perspective whereby logical consequence is defined in terms of maintenance of the truth of the premises for the conclusion of an argument. In what follows, we highlight the main results obtained from the elements presented in this paper.

(FC1) Informationally empty formulae are informational logical consequence of the empty set (Theorem 4.6a). This means that probabilistically valid and contradictory formulae are self-sustained, which illustrates a first difference between, on one hand, the veritative-functional and probabilistic versions and, on the other hand, the informational version of logical consequence. In these two versions, a contradiction is generally not a logical consequence of a given set of premises.

(FC2) The formulae that are informational logical consequence of a given set of formulae whose quantity of information is always null are informational logical consequence of the empty set (Theorem 4.6c). This result also illustrates an inherent characteristic of the informational perspective of logical consequence. From this, it follows that if a formula is logical consequence of a contradictory set of formulae, it is informational logical consequence of the empty set. This does not generally hold in the case of the veritative-functional and probabilistic perspectives of logical consequence.

(FC3) Some of the rules of inference of classical formal logical systems, and some of the arguments traditionally considered valid, do not possess general validity in the informational perspective of logical consequence (Proposition 4.7). The cases in which the set of premises possesses null information provide examples showing that the conclusion can be more informative than the set of premises in these rules or arguments. Meanwhile, it can be shown that when the quantity of information in the set of premises is greater than zero, the quantity of information in the conclusion is always smaller than the quantity of information in the set of premises.

This result indicates that according to the veritative-functional perspective, the conclusion of a valid argument can possess more information than its set of premises. This seems to support the conception that in a valid argument, the information in the conclusion is already implicitly or explicitly given in the premises. We leave it open for future work to investigate the nature of the information underlying the veritative-functional perspective, as well as the association between the true and probability values and the sentences involved in a logical consequence relation. We believe that this analysis should resolve, at least partially, the strangeness indicated in this paragraph.

In the informational and veritative-functional perspectives, amplifying inductive arguments, where the conclusion is more informative than the set of premises, are invalid. Meanwhile, such arguments may be considered interesting in some areas of knowledge, given that they amplify information.

The theorem in question also shows that the informational perspective of logical consequence is not equivalent to the veritative-functional and probabilistic perspectives. There are formulae that are the veritative-functional and probabilistic logical consequence of a given set of formulae, but are not informational logical consequence of it. On the other hand, as already shown from Theorems 4.6d or 4.8d, some formulae can be informational logical consequence of a given set of formulae, but cannot be logical consequence from the probabilistic or veritative-functional points of view.

(FC4) A formula is only informational logical consequence of a given set of formulae if that set of premises is probabilistically equivalent to the conclusion, or if the conclusion is informationally null (Theorem 4.8). This result shows that a probabilistically contradictory formula can be informational consequence of a given set of informative formulae. As an example: $\varphi \nvDash \neg\varphi \wedge \varphi$.

If, on one hand, informational logical consequence is distinct from traditional classical deductive logical consequence, on the other hand, it should not be considered to be an inductive inference. This is because a characteristic of induction is that it permits the conclusion to possess more information than its set of premises.

(FC5) The logic underlying informational logical consequence is, at the least, paraconsistent *sensu lato* (Theorem 4.8b). One of the motives for arriving at this conclusion is that in paraconsistent logics, the principle of explosion is not valid: from a contradiction does not follow any formula. A proposal for future work is to analyze the elementary characteristics of a paraconsistent system, and show that they are satisfied by informational logical consequence.

The logic underlying informational logical consequence is not classical logic, since in classical formal logical systems a contradiction generates any formula. Consequently, the complementary logics are also unable to provide a basis for informational logical consequence, given that they satisfy the principles of classical logic. Heterodox logics, such as the intuitionistic and temporal systems, are also unable to fulfill this role, as shown in the commentaries on Theorems 4.8b and 4.5d.

From Theorem 4.8b, we find that, in contrast to the veritative-functional and probabilistic perspectives, in formal logical systems that adopt informational logical consequence, the inconsistency of a given theory does not imply its triviality, as shown by Alves (2012b). From the point of view of formal logical systems such as the classical ones, this signifies that any formula would be considered a theorem of the given theory. Consequently, from the theorem of completeness, for these systems we have $\vdash_{\mathcal{V}} \psi$ and $\vdash_{\mathbb{P}} \psi$, for any ψ. Meanwhile, since in informational logical consequence, in general, $\bot_{\mathbb{P}} \nvDash \psi$, it is not possible to conclude that the inconsistency of a theory implies its triviality.

(FC6) The corresponding semantics of the Theorem of Deduction is not valid in the informational perspective of logical consequence (Theorem 4.9c). The last three items of Theorem 4.9 illustrate other significant results specific to the informational perspective. In the veritative-functional and probabilistic perspectives, if two formulae are logical consequence of an empty set of inferences, then they are equivalent, as they are both valid. In the informational version, as expressed in Theorem 4.9d, this does not possess general validity. A probabilistically valid formula and another that is probabilistically inconsistent provide an example to demonstrate the invalidity of this result.

Theorem 4.9d illustrates another specificity of informational logical consequence. This result can be easily shown for the probabilistic version, as reported by Alves (2012b). In the case of the veritative-functional version, it is necessary to adapt the right hand side of the result: the true value of the set of premises is smaller than the true value of the conclusion, in all evaluations. This signifies that it is never possible that all the premises can be true while the conclusion is false in the same single evaluation. Finally, the last item of the theorem states that informational and veritative-functional logical consequence cannot be satisfied by the same sets of sentences.

(FC7) Informational logical consequence is not a Tarskian logical consequence (Theorem 4.10). A large part of the characteristics specific to the informational perspective of logical consequence, such as that which refers to non-monotonicity, derives from cases

in which the quantity of information is null. The specific difference of this perspective is found in cases that concern extreme probability values, whether in individual situations or in all situations.

While this may be the basic difference, it should not be considered a small difference. It produces results that might be considered discrepant, when compared to the traditional perspectives of logical consequence. Amongst these, we recall that some rules of inference do not constitute valid arguments, and that not every formula is informational logical consequence of a contradictory set of formulae. Furthermore, it shows that logical consequence, when analyzed from the informational viewpoint in question, ceases to be Tarskian, and that the logic underlying this perspective is not classical, but at the very least paraconsistent.

(FC8) The set of premises of an informational inference is always finite, and the sample space is constituted of a finite set of events. These two characteristics represent serious restrictions in our proposal. The first restriction indicates the impossibility of dealing with arguments using a potentially infinite set of premises, such as the ω-arguments, described by Tarski (1956b). This author claims to have constructed a theory in which sentences of the type "*n* possesses the property P", for natural *n*, are theorems of the theory, and the sentence "every natural number possesses the property P" cannot be proved in the theory. Hence, this sentence is not their logical consequence, which in this case seems absurd.

The second restriction limits the possible models for a language of a formal system. A proposal for future work would be to consider informational logical consequence based on a definition that involves an infinite probability space. It would then become viable to analyze informational logical consequence for languages of first-order theories, which has not been addressed in the present work.

References

ALVES, M. A. Informação e conteúdo informacional: notas para um estudo da ação. In: GONZALEZ, M. E. Q; BROENS, M. C. (Orgs). *Informação, Conhecimento e Ação Ética*. São Paulo: Cultura Acadêmica, p. 98-112, 2012a.

_____. *Lógica e Informação: Uma Análise da Consequência Lógica a Partir de uma Perspectiva Quantitativa da Informação*. Campinas/SP: Universidade Estadual de Campinas, 2012b (Doctorate thesis).

BRESCIANI FILHO, E; D'OTTAVIANO, I. M. L. Conceitos básicos de sistêmica. In: D'OTTAVIANO, I. M. L; GONZALES, M. E. Q. (Orgs.). *Auto-organização: Estudos Interdisciplinares*. Campinas: Universidade Estadual de Campinas/CLE, p. 283-306, 1990 (Coleção CLE, v. 30).

DA COSTA, N. C. A. *Lógica Indutiva e Probabilidade*. 2nd ed. São Paulo: Ed. HUCITEC, 1993.

_____. *Logiques Classiques et Non Classiques*: Essai sur les Fundements de la Logique. Paris: Masson, 1997.

D'OTTAVIANO, I. M. L. A lógica clássica e o surgimento das lógicas não-clássicas. In: ÉVORA, F. R. R. (Org.). *Século XIX: O Nascimento da Ciência Contemporânea*. Campinas: Universidade Estadual de Campinas/CLE, p.65-93, 1992 (Coleção CLE, v. 11).

DRETSKE, F. *Knowledge and the Flow of Information*. Oxford: Basil Blackwell, 1981.

ENDERTON, H. B. *Elements of set theory*. San Diego: Academic Press, 1977.

GONZALEZ, M. E. Q. Informação e cognição: uma proposta de (dis)solução do problema mente-corpo. In: Encontro Brasileiro/Internacional de Ciências Cognitivas, 2, 1996, Campos dos Goytacazes. *Anais*... Campos de Goytacazes: Universidade Estadual do Norte Fluminense, p. 53-60, 1996.

FEITOSA, H. A; D'OTTAVIANO, I. M. L. Um olhar algébrico sobre as traduções intuicionistas. In: SAUTTER, F. T.; FEITOSA, H. A. (Orgs.) *Lógica: Teoria, Aplicações e Reflexões*. Campinas: Universidade Estadual de Campinas/CLE, p. 59-90, 2004 (Coleção CLE, v. 39).

HAACK, S. *Philosophy of logics*. Cambridge: Cambridge University Press, 1978.

HARTLEY, R. Transmission of information. *Bell System Technical Journal*, v. 7, p. 535-563, 1927.

HERSHBERGER, W. *Principles of Communication Systems.* New York: Prentice-Hall, 1955.

MATES, B. *Elementary logic*. New York: Oxford University Press, Inc., 1972.

MENDELSON, E. *Introduction to Mathematical Logic*. Princeton, NJ: D. Van Nostrand, 1964.

SHANNON, C. A mathematical theory of information. *The Bell System Technical Journal*, vol. 27, p. 379-423, 1948.

_____; WEAVER, W. *The Mathematical Theory of Information.* Urbana: University of Illinois Press, 1949.

SHOENFIELD, J. *Mathematical Logic.* Reading, Mass.: Addison Wesley Publishing Company, 1967.

TARSKI, A. On some fundamental concepts of metamathematics. In: Tarski. A. (ed.) *Logic, semantics, metamathematics*: papers from 1923 to 1938. Translated by J. H. Woodger, pp. 30-37. Oxford: Clarendon Press, 1956a

_____. Fundamental concepts of methodology of deductive sciences. In: Tarski. A. (ed.) *Logic, semantics, metamathematics*: papers from 1923 to 1938. Translated by J. H. Woodger, pp. 60-109. Oxford: Clarendon Press, 1956b.

_____. Concept of truth in formalized languages. In: Tarski. A. (ed.) *Logic, semantics, metamathematics*: papers from 1923 to 1938. Translated by J. H. Woodger, pp. 152-278. Oxford: Clarendon Press, 1956c.

_____. On the concept of logical consequence. In: Tarski. A. (ed.) *Logic, semantics, metamathematics*: papers from 1923 to 1938. Translated by J. H. Woodger, pp. 409-420. Oxford: Clarendon Press, 1956d.