Carnap's notion of informational containment and monotonicity

An unfinished sketch

Berislav Žarnić

University of Split

Overview

Introduction

The idea that logic is about just one notion of 'logical consequence' is actually one very particular historical stance. It was absent in the work of the great pioneer Bernard Bolzano, who thought that logic should chart the many different consequence relations that we have, depending on the reasoning task at hand.

van Benthem, J.
(2008) Logic and reasoning: Do the facts matter?.
Studia Logica 88:67–84.

Bolzano's consequence relation: Consistency and invariance under substitution

The two conditions:

(a) $A, B, C, D, \ldots, M, N, O, \ldots$ must be compatible.

(b) Every substitution of "variable ideas" which makes all $A, B, C, D, \ldots$ true also makes all $M, N, O, \ldots$ true.

The requirement of compatibility shows that Bolzano's consequence relation is non-monotonic. For example, $p \in Cn(\{p\})$ but $p \notin Cn(\{p, \neg p\})$

true. Obviously, this problem is of some importance. Let us consider, first of all, the case that among the compatible propositions *A*, *B*, *C*, *D*, . . . *M*, *N*, *O*, . . . the following relation obtains: all ideas whose substitution for the variable ideas *i*, *j*, . . . turns a certain part of these propositions, namely *A*, *B*, *C*, *D*, . . . into truths, also have the characteristic of making a certain other part of these propositions, namely *M*. *N*, *O*, . . . true. This special relation which we think between propositions *A*, *B*, *C*, *D*, . . . on the one hand and *M*, *N*, *O*, . . . on the other is of special importance, since it puts us in a position to infer the truth of *M*, *N*, *O*, . . ., once we have recognized the truth of *A*, *B*, *C*, *D*. . . . I wish to give the name of *deducibility* [*Ableitbarkeit*] to this relation between propositions *A*, *B*, *C*, *D*, . . . on one hand and *M*, *N*, *O*, . . . on the other. Hence I say that

Bolzano, B.
([1837] 1972) *The Theory of Science* (Die Wissenschaftslehre oder Versuch einer Neuen Darstellung der Logik), ed. and translated by Rolf George.
University of California Press

Tarski's consequence

- In 1928. in a lecture in Warsaw and in the paper of 1930. Tarski introduced a general perspective on consequence relation.
- It is axiomatized by ten axioms divided into two groups:
  1. Properties of the relation between the sets of sentences considered as wholes. Nowadays called *structural rules*.
  2. Properties of the relation defined in terms of the syntactical properties of elements in the sets. Nowadays called *logical* rules.

Structural properties of consequence relation

Axioms on consequence relation

AXIOM 1. $|S| \leq \aleph_0$.
AXIOM 2. If $X \subseteq S$, then $X \subseteq Cn(X) \subseteq S$.
AXIOM 3. If $X \subseteq S$, then $Cn(Cn(X)) = Cn(X)$.
AXIOM 4. If $X \subseteq S$, then $Cn(X) = \bigcup_{Y \subseteq X \text{ and } |Y| < \aleph_0} Cn(Y)$.
AXIOM 5. There exists a sentence $x \in S$ such that $Cn(\{x\}) = S$.

TARSKI, A.
On some fundamental concepts of metamathematics.
In *Logic, semantics, metamathematics : papers from 1923 to 1938.* Clarendon Press, Oxford, 1956, pp. 30–37. First published in 1930. as
Über einige fundamentale Begriffe der Metamathematik
Comptes Rendus des séances de la Société des Sciences et des Lettres de Varsovie vol. 23.: 22–29

For countable languages S (Axiom 1) it holds that:
– consequences of sentences remain within the same language and premises are their own consequences (reflexivity; Axiom 2),
– consequences of consequences of a set are already consequences of that set (transitivity; Axiom 3),
– consequences of a set X do not exceed the consequences of their finite subsets Y, which are retained in their superset X consequences (compactness, Axiom 4),
– there is at least one sentence in the language such that its consequences include all the sentences of that language (existence of *falsum*, "absurdity," "explosive sentence," "informational breakdown," etc.; Axiom 5).

Axiom 5: compactness and restricted monotonicity

If $X \subseteq S$, then $Cn(X) = \bigcup_{Y \subseteq X \text{ and } |Y| < \aleph_0} Cn(Y)$. The Axiom 5. can be rewritten as the conjunction of the two sentences:

1. L-R $\forall p(p \in Cn(X) \rightarrow \exists Y(|Y| < \aleph_0 \wedge Y \subseteq X \wedge p \in Cn(Y)))$ and
2. R-L $\forall p(\exists Y(|Y| < \aleph_0 \wedge Y \subseteq X \wedge p \in Cn(Y)) \rightarrow p \in Cn(X))$.

The L-R part captures the notion of 'proofs as finite objects (texts)'. It is related to 'compactness theorem'. The R-L part implies monotonicity of the consequence relation only for finite sets.

Monotonicity

Theorem

If $X \subseteq Y \subseteq S$, then $Cn(X) \subseteq Cn(Y)$.

Tarski omits the proof because of its simplicity. The proof requires the use of both directions of Axiom 5.

Proof.

Let $X \subseteq Y$. Assume $p \in Cn(X)$. By Axiom 5. L-R, $\exists Z(|Z| < \aleph_0 \wedge Z \subseteq X \wedge p \in Cn(Z)))$. Since inclusion relation is transitive, $Z \subseteq Y$. Thus the three conditions are satisfied: Z is a finite subset of Y and p belongs to $Cn(Z)$. By Axiom 5. R-L, $p \in Cn(Y)$. □

The historical significance

- Tarski's axioms introduced a bird's-eye view on consequence relation.
- Once *a* structure of consequence relation has been recognized it has become possible to discover structures of other, weaker consequence-like relations in language.
- We proceed to Carnap's semantical notion of consequence as informational containment and propose the hypothesis that it enables introduction of a weak, non-monotonic type of consequence relation.

Possible states of affairs as building blocks for the concept of information

Content

A possible state of affairs of all objects dealt with in a system S with respect to all properties and relations dealt with in S is called an L-state with respect to S. A sentence or sentential class designating an L-state is called a state-description.

Carnap, R.,
(1942) *Introduction to Semantics.*
Harvard University Press

Possible state of affairs	*Pa*	*Pb*	*Qa*	*Qb*	State description
w_1	**t**	**t**	**t**	**t**	$\{Pa, Pb, Qa, Qb\}$
...		...			...
w_5	**t**	**f**	**t**	**t**	$\{Pa, \neg Pb, Qa, Qb\}$
...		...			...
w_{16}	**f**	**f**	**f**	**f**	$\{\neg Pa, \neg Pb, \neg Qa, \neg Qb\}$

The logical range of a sentence

- Carnap explicates the 'L-range of a sentence $\mathfrak{S}$' in semantic terms as: *class of possible states of affairs "admitted by $\mathfrak{S}$".*
- Carnap states the postulates on the parallelism between the inclusion of logical ranges and the consequence relation.

Postulates for L-range

+P18-1. If $\mathfrak{S}_i \underset{L}{\rightarrow} \mathfrak{S}_j$ (in S), then $\mathrm{Lr}\mathfrak{S}_i \subseteq \mathrm{Lr}\mathfrak{S}_i$.

+P18-2. If $\mathrm{Lr}\mathfrak{S}_i \subseteq \mathrm{Lr}\mathfrak{S}_j$ (in S), then $\mathfrak{S}_i \underset{L}{\rightarrow} \mathfrak{S}_j$.

Carnap, R.,
(1942) *Introduction to Semantics.*
Harvard University Press

Consequence and range inclusion

- Tarski's consequence operation and Carnap's logical range inclusion can be related as follows:

Carnap	Tarski
$\mathrm{Lr}\varphi \subseteq \mathrm{Lr}\psi$	$\psi \in Cn(\{\varphi\})$

- The 'L-range of a sentence $\mathfrak{S}$' can be formulated in the modern notation as follows (with W standing for the the class of all possible states of affairs in a given semantical system):

Definition

$$\mathrm{Lr}\varphi = \{w \in W \mid w(\varphi) = \mathrm{t}\} \qquad (1)$$

$$\mathrm{Lr}T = \{w \in W \mid w(\varphi) = \mathrm{t} \text{ for all } \varphi \in T\} \qquad (2)$$

- The aforementioned postulates, which state that L-range inclusion parallels consequence relation, can be proved using the definition.

Is Carnap's consequence monotonic?

- The Tarskian properties of consequence relation can be restated using the notion of logical range and proved using the definition. The symbol for the "semantical system" will be omitted.
- Carnap did not explicitly state the theorem on the monotonicity of consequence relation as based on "logical range" inclusion.
  - Tarski's theorem: *If* $X \subseteq Y$, *then* $Cn(X) \subseteq Cn(Y)$.
  - Tarski's theorem reformulated: If $X \subseteq Y$, then for all φ if $\varphi \in Cn(X)$, then $\varphi \in Cn(Y)$.
  - Corresponding Carnap's theorem would be: *If* $X \subseteq Y$, *then for all* φ *if* $\mathrm{Lr}X \subseteq \mathrm{Lr}\varphi$, *then* $\mathrm{Lr}Y \subseteq \mathrm{Lr}\varphi$.

Monotonicity of logical range inclusion

Theorem

If $X \subseteq Y$*, then for all* φ *if* $\mathrm{Lr}X \subseteq \mathrm{Lr}\varphi$*, then* $\mathrm{Lr}Y \subseteq \mathrm{Lr}\varphi$*.*

Proof.

Let $X \subseteq Y$. Suppose $\mathrm{Lr}X \subseteq \mathrm{Lr}\varphi$. Let w be an arbitrary state of affairs such that $w \in \mathrm{Lr}Y$. Since $Y = X \cup (Y - X)$, w makes true all sentences in X. Therefore, $w \in \mathrm{Lr}X$. Given that $\mathrm{Lr}X \subseteq \mathrm{Lr}\varphi$, it follows that $w \in \mathrm{Lr}\varphi$. □

Logical content as the dual notion of logical range

- For the concept of L-range there is the dual notion of L-content: $\mathrm{Lc}X = W - \mathrm{Lr}X$.
- Cranap prefers the notion of "logical content" since it accords better with what is the common use of expressions for relations of relative informativeness. For example, $Pa \wedge Qa$ is more informative then Pa, but $|\mathrm{Lr}(Pa \wedge Qa| < |\mathrm{Lr}(Pa)|$. On the other hand, the relation between "logical contents" gives what is desired: the more informative sentence has the greater cardinality of the logical content. For example, $|\mathrm{Lc}(Pa \wedge Qa)| > |\mathrm{Lc}(Pa)|$.

Information containment conception of logical consequence

The notion 'logical content of sentence' (1942) reappears in 1952. as the notion of 'semantic information carried by a sentence'

Information

Whenever *i* **L**-implies *j*, *i* asserts all that is asserted by *j*, and possibly more. In other words, the information carried by *i* includes the information carried by *j* as a (perhaps improper) part. Using '*In*(...)' as an abbreviation for the presystematic concept 'the information carried by ...', we can now state the requirement in the following way:
R3-1. $In(i)$ includes $In(j)$ iff *i* **L**-implies *j*.
By this requirement we have committed ourselves to treat information as a set or class of something. This stands in good agreement with common ways of expression,as for example, "The information supplied by this statement is more inclusive than (or is identical with, or overlaps) that supplied by the other statement."

Rudolf Carnap and Yehoshua Bar-Hillel.
An Outline of a Theory of Semantic Information.
Technical Report no. 247. Research Laboratory of Electronics, Massachusetts Institute of Technology, 1952.

Information containment

Carnap's explication of consequence relation in terms of relations between logical contents seems to be the earliest formulation of the "information containment conception" of consequence relation.

An example

The information containment conception: P implies c if and only if then information of c is contained in the information of P. In this sense, if P implies c, then it would be redundant to assert c in a context where the propositions in P have already been asserted; i.e., no information would be added by asserting c.

Jose M. Saguillo.
Logical Consequence Revisited. *The Bulletin of Symbolic Logic* (1997) 3: 216-241

Two readings for the information containment notion of consequence

- The two notions of "information containment" can be derived from the Carnap's explication of consequence relation in terms of "semantic information" (logical content, logical range):

  Strong inf. cont. Conclusion makes no informational change to *any* context that includes all the information contained in premises.

  Weak inf. cont. Conclusion makes no informational change to the context that includes *only* the information contained in premises.

- The weak informational containment is a special case of the strong informational containment.

Is information addition the only sentential operation?

The semantics of sentences can be conceived in terms of an operation performed on the logical content, for example, as "addition of information". The two notions coincide in the case of "information addition" conceived as intersection operation, $\mathrm{Lr}\varphi \cap \mathrm{Lr}\psi$, because of monotonicity (if $\mathrm{Lr}\varphi \subseteq \mathrm{Lr}\varphi'$, then $\mathrm{Lr}\varphi \cap \mathrm{Lr}\psi \subseteq \mathrm{Lr}\varphi'$). On the other hand, if the repertoire of semantic operations includes operations which test some property of the "logical content", like testing whether a certain sentence can be consistently added to it (which can be compared to the first operational step in Bolzano-type consequence), then the difference between the two notions will become visible. If there is a sentence type whose semantic operation is not the "addition of information" but the "test of possibility of informational addition", then the informational content of the sentence type will be context dependent.

1. $\mathrm{Lr}(?\psi) = \mathrm{Lr}\varphi_1 \cap \cdots \cap \mathrm{Lr}\varphi_n$ in the context $\mathrm{Lr}\varphi_1 \cap \cdots \cap \mathrm{Lr}\varphi_n$ if $\mathrm{Lr}\varphi_1 \cap \cdots \cap \mathrm{Lr}\varphi_n \cap \mathrm{Lr}\psi \neq \emptyset$.
2. $\mathrm{Lr}(?\psi) = \emptyset$ in the context $\mathrm{Lr}\varphi_1 \cap \cdots \cap \mathrm{Lr}\varphi_n$ if $\mathrm{Lr}\varphi_1 \cap \cdots \cap \mathrm{Lr}\varphi_n \cap \mathrm{Lr}\psi = \emptyset$.

Practical inference

- The logical theory of practical inference has not as yet reached a clear cut form. There is no generally known theory that serves as a point of reference either for the critique or for the further development. The theoretical reason for the underdevelopment lies in the logical complexity of practical inference.
    1. The number of logical elements (usually treated as modal operators) involved in the practical inference is high and their theories are under dispute.
    2. The consequence relation of practical inference seems not to be Tarskian, but rather a very weak relation (non-transitive, non-reflexive, non-monotonic).
    3. There is a rich variety of candidate means-end relations (sufficient, necessary, INUS, SUIN,... conditions) and this fact produces different forms of the instrumental type of practical inference.
    4. There seems to be an agreement that practical inference is defeasible or non-monotonic.

Geach's description of the properties of consequence relation in practical inference

Quotation

Some years ago I read a letter in a political weekly to some such effect as this. 'I do not dispute Col. Bogey's premises, nor the logic of his inference. But even if a conclusion is validly drawn from acceptable premises, we are not obliged to accept it if those premises are incomplete; and unfortunately there is a vital premise missing from the Colonel's argument-the existence of Communist China.' I do not know what Col. Bogey's original argument had been; whether this criticism of it could be apt depends on whether it was a piece of indicative or of practical reasoning. Indicative reasoning from a set of premises, if valid, could of course not be invalidated because there is a premise "missing" from the set. But a piece of practical reasoning from a set of premises can be invalidated thus: your opponent produces a fiat you have to accept, and the addition of this to the fiats you have already accepted yields a combination with which your conclusion is inconsistent.

Peter Geach.
Dr. Kenny on practical inference.
Analysis (1966) 26: 76–79

Defeasibility of conclusion and completeness of premises

The consequence relation described by Geach has two notable properties:

- ("locality") conclusion holds in virtue of premises but it can be defeated by additional premises;
- (existence of the limit) if the premises are complete the conclusion cannot be defeated (where 'conclusion is defeated' means 'premises are acceptable and conclusion is not acceptable').
- The methodology used by Geach shows that in building the theory of practical inference one can start from the pretheoretical understanding of the specific consequence relation.
- The locality of consequence relation can be explicated using the weak notion of informational containment.

Prima facie consequence

- In a similar manner, Davidson writes that in practical inference one cannot "detach conclusions about what is desirable (or better) or obligatory from the principles that lend those conclusions colour". He uses the term 'prima facie' for this kind of consequence relation. In his representation of practical inference the conclusion is explicitly "localized" and written as follows:

$$\text{prima facie}(Conclusion, \{Premise_1, \ldots, Premise_n\})$$

Davidson, D..
How is weakness of the will possible? (1969).
In *Essays on Actions and Events* (Second ed.). Oxford: Clarendon Press.

- Davidson describes a kind of weak consequence relation that is context dependent.

Concluding remarks

- The selected results within historical development of the understanding of logical consequence show that the notions introduced at one stage need not be preserved at the other although they enable the formulation of alternative notions and the introduction of new notions.
    - Tarski's theses on structural properties of consequence relation have not been preserved within the field of practical logic but Tarski's axioms enabled recognition of other types of consequence relation.
    - Carnap's informational containment notion of logical consequence enabled the subsequent development of the "calculus of the informational content" (for example, in "update semantics") although Carnap's notion has not been preserved in its original form.
    - Carnap's informational containment notion of logical consequence in its weak variant gives a way of understanding the special character of consequence relation in practical logic.

Finiteness

- Axiom 5. reveals that the syntactic way of thinking provides the ground for structural axioms formulation.
- Let us consider an example! Suppose that the number of objects in the domain is enumerably infinite and that the name for each object appears on the list: $a_1, a_2, \ldots, a_n, \ldots$. Assume that for each $i \in \mathbb{N}$ it holds that Pa_i. Obviously one would want to have (a) $\forall x Px \in Cn(\{Pa_1, \ldots, Pa_n, \ldots\})$, and would not like to have (b) $\forall x Px \in Cn(X)$ for any proper subset X of $\{Pa_1, \ldots, Pa_n, \ldots\}$. Axiom 5. L-R forbids (a) because of (b).
- This example shows that that consequence relation characterized by axioms 1–5 does not "coincide with the common concept".

A_0. *0 possesses the given property P*,
A_1. *1 possesses the given property P*,

and, in general, all particular sentences of the form

A_n. *n possesses the given property P*,

where '*n*' represents any symbol which denotes a natural number in a given (e.g. decimal) number system. On the other hand the universal sentence:

A. *Every natural number possesses the given property P*,

cannot be proved on the basis of the theory in question by means of the normal rules of inference. This fact seems to me to speak for itself. It shows that the formalized concept of consequence, as it is generally used by mathematical logicians, by no means coincides with the common concept. Yet intuitively it seems certain that the universal sentence *A* follows in the usual sense from the totality of particular sentences $A_0, A_1, \ldots, A_n, \ldots$. Provided all these sentences are true, the sentence *A* must also be true.

Tarski, A.
(1936) On the concept of logical consequence.
In *Logic, semantics, metamathematics : papers from 1923 to 1938*. Clarendon Press,