Notre Dame Journal of Formal Logic
Volume XIII, Number 2, April 1972
NDJFAM

STRANGE ARGUMENTS

JOHN CORCORAN

In logic one frequently considers a set P of formal sentences together with a single formal sentence c and one asks whether c is a logical consequence of P. In teaching logic it is convenient to follow the philosophers (Mates [2]) and use the term *argument* to indicate such an ordered pair (P, c). If indeed c follows from P then (P, c) is said to be *valid* and otherwise *invalid*.

After learning the formal definition of the logical consequence relation in sentential logic (propositional calculus) students often find it "strange" that there should be valid arguments (P, c) whose *premises* P share no sentential letters with their respective *conclusions* c. Typical examples, of course, are the following facts:

(1) q *follows from* $\{p, \sim p\}$
(2) $(q \supset q)$ *follows from* $\{p\}$

Sometimes students are apt to attribute the "strangeness" to the concept of logical consequence and to feel, on the strength of the attribution, that the formal concept is incorrect, unrealistic, arbitrary, or something of the sort. It is the purpose of this note to indicate a nice way of disabusing thoughtful students of such unjustified feelings while at the same time providing them with some mathematical reasoning involving useful insight into the mathematical implications of the definition.

The Background Let D (the dictionary) be a countably infinite set of sentential letters and let L be the set of formal sentences built-up recursively from D using &, v, $\supset$ and $\sim$ as logical connectives and) and (for punctuation. As usual an *interpretation*, i, of L is function from D into the set $\{t,f\}$ of *truth-values* assigning a truth-value to each sentential letter. Given an interpretation i, truth-values *relative* to i (or *under* i or *on* i) are determined by defining a truth-valuation function V^i from L to $\{t,f\}$ as follows:

(1) $V^i x = ix$, *for each* x *in* D
(2) $V^i{\sim}x = NV^i x$

Received March 9, 1970

$$V^i(x \,\&\, y) = A V^i x V^i y$$
$$V^i(x \vee y) = O V^i x V^i y$$
$$V^i(x \supset y) = C V^i x V^i y$$

where A, O, C and N are functions given by the following tables.

	N
t	f
f	t

	A	O	C
tt	t	t	t
tf	f	t	f
ft	f	t	t
ff	f	f	t

One calls i a *true interpretation of a sentence* x if $V^i x = \mathrm{t}$ and one calls i a true interpretation of a set P if $V^i y = \mathrm{t}$ for all y in P. So, in particular, $\{p, \sim p\}$ has no true interpretations and for that reason is said to be *unsatisfiable* whereas D itself is *satisfied* by the interpretation which maps all sentential letters into $\{\mathrm{t}\}$. Vacuously, every interpretation is a true interpretation of the null set, $\emptyset$.

At this point the formal definition of the logical consequence relation is usually introduced: c is a *logical consequence* of P iff every true interpretation of P is a true interpretation of c. One writes $P \vDash c$ to mean that c is a logical consequence of P. In order to relate this notation to the concepts of argument, validity and invalidity we would define: (P, c) is *valid* if $P \vDash c$ and (P, c) is *invalid* if (P, c) is not valid. $\vDash c$ is an abbreviation of $\emptyset \vDash c$ and some authors say that c is *logically true* (suggesting "true in virtue of the respective meanings, A, O, C and N, of the logical connectives, &, v, $\supset$ and $\sim$") to mean simply $\vDash c$.

In practical cases when one has occasion to point out that (for a particular P and c) $P \vDash c$ one is usually indicating a special, intimate connection between P and c—special at least in this sense: that one could not substitute an arbitrary P' for P and still have $P' \vDash c$, nor could one substitute an arbitrary c' for c and still have $P \vDash c'$. Thus there are two "strange" (i.e., unusual) cases: first, when P is unsatisfiable, $P \vDash c$ for *all* sentences c so in the above sense there is no special relation between P and c; second, when c is logically true $P \vDash c$ for *all* sets P, so again in the above sense there is no special relation between P and c. For pedagogical purposes let us call unsatisfiable sets P *strange* premise sets and let us call logically true sentences c *strange* conclusions.

As noted above students feel that it is "strange" (weird) that there should be cases where $P \vDash c$ but where there are no sentence letters shared by P and c. *Perhaps* the (understandable) source of this feeling is a tacit presupposition that when P and c have no letters in common there could be no special relation (in the above sense) between them. Let us entertain this presupposition as an hypothesis for further investigation. (For purposes of brevity let Dc [the dictionary of c] and DP [the dictionary of P] be respectively the set of sentential letters actually occurring in the sentence c and the set of sentential letters actually occurring in members of P. If $P = \{c_1, c_2, \ldots, c_n\}$ then DP is the union of $\{Dc_1, Dc_2, \ldots, Dc_n\}$; $DL = D$.)

Hypothesis I. *Let* $P \vDash c$. *If* $DP \cap Dc = \emptyset$ *then one of the following conditions hold*: (1) *for all* P', $P' \vDash c$, or (2) *for all* c', $P \vDash c'$.

Let us put aside the hypothesis and return to the situation which suggested it. In order to be more concrete let us imagine that we are considering a particular valid argument (P, c) whose premises, P, share no sentential letters with its conclusion, c. ($P \vDash c$ and $DP \cap Dc = \emptyset$.) Since students feel that such a situation is "strange" let us give recognition to the feeling by defining such arguments to be *strange*. As noted, there is a tendency for the "strangeness" to be attributed to the formal concept of logical consequence. Before submitting to the tendency and regarding the formal concept as "strange" the student should be urged to notice that there are *three* things involved in a *valid* argument: premises, conclusion and the logical consequence relation. Thus, the formal concept could be defended against this attack by proving the following additional hypothesis (whose converse is false).

Hypothesis II. *Every strange argument either has strange premises or a strange conclusion.*

A proof of this hypothesis would show that in every strange argument it is possible to attribute the strangeness to a place other than the logical consequence relation. The balance of this note is to offer a proof of Hypothesis II from which Hypothesis I follows. Our development is intended to be suitable for use in the first or second week of an introductory logic course for advanced undergraduates and/or beginning graduates. Texts suitable for such courses are Mendelson [3] and Robbin [4].

The Mathematics In computing the truth-value, $V^i x$, of a sentence x under an interpretation i, one automatically ignores the interpretations of the letters not occurring in x. One feels that these are irrelevant. Our first step is to note this fact as a lemma, i.e. we want to show that if two interpretations coincide on all sentential letters in a given sentence then that sentence has the same truth-value on each of the two interpretations (regardless of how the interpretations may differ on the sentential letters not occurring in the given sentence).

Lemma I. *Let x be a sentence and let Dx be the set of sentential letters actually occurring in x. Let i and j be two interpretations coinciding on Dx, i.e. such that $iy = jy$ for all y in Dx. Then* $V^i x = V^j x$.

The proof of this lemma is a simple application of mathematical induction on the number of occurrences of logical connectives in x. In my own teaching I call Lemma I "the coincidence lemma" and I suggest the following sloganized paraphrase of it: "If two interpretations coincide on the dictionary of a sentence their valuations coincide on the sentence itself." As an immediate corollary we get the generalization to sets of sentences.

Corollary I. *Let P be a set of sentences and let DP be the set of sentential* *letters actually occurring in P. Let i and j be two interpretations which coincide on DP. Then i and j are both true interpretations of P or neither are true interpretations of P.*

Now we are ready to demonstrate Hypothesis II. By unpacking the definitions we see that Hypothesis II is restated as Theorem I below.

Theorem I. *Let* $P \vDash c$ *with* $DP \cap Dc = \emptyset$. *Then either P is unsatisfiable or c is logically true.*

Proof: Assume the hypothesis, what we want to show is that either P has no true interpretations or else that every interpretation is a true interpretation of c.

Either P has no true interpretations or else it has at least one true interpretation i_0. If the former, we are finished. Suppose the latter. Now we want to show that every interpretation j makes c true. Let j be any interpretation at all and let k be *the* (intermediate) interpretation which coincides with i_0 on DP and which coincides with j on the rest of the sentential letters. Since i_0 is a true interpretation of P and since k coincides with i_0 on DP, k must be a true interpretation of P (by Corollary I). By hypothesis $P \vDash c$, so k is a true interpretation of c. Now since $DP \cap Dc = \emptyset$ and k coincides with j outside of DP, k must coincide with j on Dc. But k is a true interpretation of c. Thus by Lemma I, j must also be a true interpretation of c. Since j was arbitrary every interpretation is a true interpretation of c. QED

Some of the less "mathematically mature" students have found the following discussion employing "visual aids" helpful. Picture the sentential letters ordered as follows: all those in DP, then all those in Dc, then the rest. x's represent the distribution of truth-values given by i_0 (some true interpretation of P). +'s represent the distribution according to j (arbitrarily chosen). The combination of x's and +'s indicates *the* distribution, k, which is intermediate between i_0 and j coinciding with i_0 on DP and with j elsewhere.

	DP	Dc	rest
i_0	x x x x x x x x x x x	x x x x x	x x x x x x x x x x x x x
k	x x x x x x x x x x x	+ + + + +	+ + + + + + + + + + + + +
j	+ + + + + + + + + + +	+ + + + +	+ + + + + + + + + + + + +

Main points:

(1) Since i_0 is a true interpretation of P and since k coincides with i_0 on DP, k is a true interpretation of P by Corollary I.

(2) By hypothesis $P \vDash c$. Thus, since k is a true interpretation of P, k is also a true interpretation of c.

(3) Since $DP \cap Dc = \emptyset$ and k coincides with j outside of DP, k agrees with j on Dc (as pictured above).

(4) Thus, by Lemma I, since k is a true interpretation of c, j must also be a true interpretation of c. QED

The proof of Hypothesis I from Theorem I is an easy exercise which illustrates the interactions of the definitions. The above reasoning for Theorem I should prove to be a helpful pedagogical preliminary to the proof of the Craig interpolation lemma for sentential logic (see [1], p. 5) which trivially implies Theorem I but is itself more involved.

Concerning modification of the above reasoning for use in first and higher order logics the following brief comments should suffice. First, the analogue to Lemma I is: if two interpretations *in the same domain* agree on the non-logical terms of c then their valuations agree on c. (This is an easy generalization of proposition VIII, [3], p. 52.) Second, the analogue to Theorem II is: if $P \vDash c$ where P and c share no non-logical terms then, *for each domain* u, either P has no true interpretations *in* u or else c is true on every interpretation *in* u. As far as applications to other sentential logics are concerned, it should be obvious (1) that the above reasoning did not depend on the *number* of truth-values, as only t was necessary and (2) that the reasoning did not depend on which truth-functions were assigned to the connectives nor on the number of connectives involved.

REFERENCES

[1] Kreisel, G., and J. L. Krivine, *Elements of Mathematical Logic,* North-Holland, Amsterdam (1967).

[2] Mates, Benson, *Elementary Logic,* Oxford, New York (1964).

[3] Mendelson, Elliot, *Introduction to Mathematical Logic,* Van Nostrand, Princeton (1964).

[4] Robbin, Joel W., *Mathematical Logic: A First Course,* Benjamin, New York (1969).

The University of Michigan
Ann Arbor, Michigan