Notre Dame Journal of Formal Logic
Volume XVII, Number 2, April 1976
NDJFAM

CONSISTENCY OF n-ORDER LOGICS

WILLIAM J. THOMAS

1 *Introduction** A very natural second-order generalization of first-order predicate logic might be expected to result from an application of the intuition that predicate letters, and perhaps function letters, can be treated in much the same fashion as individual variables. The content of this intuition would seem to be that predicate letters are considered a kind of variable (rather than a kind of constant, as in the usual treatment of first-order logic). Both axioms and rules of inference which affect quantification of individual variables may be extended to license the same operations on predicate letters. Finally, predicate letters are permitted to appear in other than the initial positions of atomic well-formed formulae ("wfs" hereinafter). This last provision may, in some accounts, be accompanied by the introduction of predicate letters of higher type (predicates of predicates) which occur only in the initial position, and over which quantification is not permitted.

Motivating arguments for higher order extensions of predicate logic usually proceed by producing an example of a clearly valid argument which seems to be most naturally rendered schematically by means of the higher order apparatus. "Richard has all of George's good qualities. Candor is a good quality. George is candid. Therefore, Richard is candid," would seem a fairly typical example.

Formal treatments of higher order logics tend to be concerned with disguised pieces of set theory, with comprehension axioms and the other trappings of set theory, or they tend to assume the restrictions of a type theory, or both. Informal treatments, like that in Copi [1], are too vague in their specification of the generalization to permit meaningful discussion of the consistency question. Quine, in [3], has charged that any "natural" (in our sense) generalization must be inconsistent.

*This paper was presented to the 1972-73 Annual Meeting of the Association for Symbolic Logic, Dallas, Texas, January 25-26, 1973. Funding, in the form of a UNCC Summer Research Grant, is gratefully acknowledged.

Received May 7, 1973

In the next section we make formally precise the intuitive higher order generalizations of first-order logic. It is necessary to adopt the vocabulary of type theory in order to distinguish the n-order from the $(n + 1)$-order theory, but we do not assume the restrictive onus of a type-theoretic doctrine about meaningfulness (well-formedness). For purposes of proving consistency in the third section, our liberal policy in regard to formation rules is an advantage in any event: the consistency of systems which differ from ours, essentially only in respect of having more restrictive formation rules, will follow from our result. The consistency proof proceeds by providing an effective provability-preserving translation from n-order logic to first-order logic. We apply the translation to Quine's derivation (in [3]) of an inconsistency in the "natural" second-order logic.

In the fourth and final section, we give a second theorem: the inverse of the translation function used in the consistency argument is provability-preserving too. It then follows that an n-order wf is provable in the n-order logic just in case its first-order translation is provable in first-order logic. We draw the philosophical conclusion that n-order "natural" generalizations are dispensable. We apply the philosophical conclusion to the motivating example given above. Finally, we offer a few remarks about completeness.

2 *n-Order Generalizations*

2.1 *Formation rules for n-order logics* $\mathbf{K}_n$ Well-formed formulae of order n (n-wfs) are defined in terms of the elements of $\mathrm{T} = \{{}_ix_j^k | i, j \in N, k = \alpha \vee k \in N\}$, where N is the set of numerals, α is an arbitrary symbol distinct from the numerals. We call T the set of "terms." The intended interpretation of the super and subscripts on the terms is as follows: ${}_ix_j^k$ is to be understood as an i-adic variable of order j; the superscript k serves to distinguish infinitely many different variables of each order and for each number of argument places. 'α' is used to distinguish certain variables for reasons of convenience in the consistency argument.

1. ${}_ix_2^k$, ${}_jx_1^{m_1}$, . . ., ${}_nx_1^{m_i}$, where $k \in N$ or $k = \alpha$, is 1-wf. Note that we allow "individual variables" to be polyadic (or to have left-hand subscripts which so indicate, if you prefer).
2. ${}_ix_{p_1}^{m_1}$, . . ., ${}_jx_{p_{i-1}}^{m_{i-1}}$, is $(n + 1)$-wf (for $n \geqslant 1$) if $p_1 \geqslant \max\{p_i\}$ and $(n + 1) \geqslant p_1$.
3. If F is n-wf, then so is $\sim F$.
4. If F and G are n-wf, so is $(F \supset G)$.
5. If F is n-wf, then so is $({}_ix_j^k)\ F$, provided that $j \leqslant n$.
6. Nothing else is n-wf, for any n.

Obviously, the other truth-functional connectives could have been included, but to have done so would have been a needless complication. These formation rules capture and generalize in an exact way the intuitive idea that higher order logic results from "treating the predicate letters like the variables." It may also be noted that whatever is an n-wf is also an $(n + 1)$-wf.

2.2 *Axiomatization of the* K_n Our procedure will be to give the axioms for K_n in terms of axiom schemata, inductively from a specification of the axioms for K_1. Our axiomatization for K_1 is a modification of the system K, described by Mendelson (in [2]). We use only one rule of inference, *viz. modus ponens.*

K_1 Axiom Schemata: (we use '$A(x)$', '$B(x)$', . . . as meta-linguistic variables ranging over n-wfs containing the free variable 'x').

K_1 (1.) (*Every tautologous* 1-*wf is an axiom.*)
K_1 (2.) $({}_ix_1^k)\ A({}_ix_1^k) \supset A({}_ix_1^j)$, *where* $A({}_ix_1^k)$ *is* 1-*wf, and* ${}_ix_1^j$ *is a term free for* ${}_ix_1^k$ *in* $A({}_ix_1^k)$. *Again, k, j may be numerals, or* 'α'.
K_1 (3.) $({}_ix_1^k)(A \supset B) \supset (A \supset ({}_ix_1^k)\,B)$, *where A and B are* 1-*wf, A containing no free occurrences of* ${}_ix_1^k$.
K_1 (4.) $A \supset ({}_ix_1^k)\ A$, *where A is* 1-*wf.*

Example: '$({}_0x_1^6)\ {}_2x_2^{14}({}_4x_1^8,\ {}_0x_1^6) \supset {}_2x_2^{14}({}_4x_1^8,\ {}_0x_1^7)$ is an axiom, by K_1 (2.).

Now, $\mathsf{K}_{(n+1)}$ arises from K_n by adding axiom schemata to K_n which license treating predicate letters of K_n quantificationally as variables are treated in K_n. This admittedly vague specification may be made precise, as follows:

$\mathsf{K}_{(n+1)}$ Axiom Schemata:

$\mathsf{K}_{(n+1)}$ (1.) (*Every axiom of* K_n *is an axiom of* $\mathsf{K}_{(n+1)}$, *as is every tautologous* $(n+1)$-*wf.*)
$\mathsf{K}_{(n+1)}$ (2.) $({}_ix_{(n+1)}^k)\ A({}_ix_{(n+1)}^k) \supset A({}_ix_{(n+1)}^j)$, *where* $A({}_ix_{(n+1)}^k)$ *is* $(n+1)$-*wf, and* ${}_ix_{(n+1)}^j$ *is a term free for* ${}_ix_{(n+1)}^k$ *in* $A({}_ix_{(n+1)}^k)$.
$\mathsf{K}_{(n+1)}$ (3.) $({}_ix_{(n+1)}^k)(A \supset B) \supset (A \supset ({}_ix_{(n+1)}^k)B)$, *where A and B are* $(n+1)$-*wf, and A contains no free occurrences of* ${}_ix_{(n+1)}^k$.
$\mathsf{K}_{(n+1)}$ (4.) $A \supset ({}_ix_{(n+1)}^k)A$, *where A is* $(n+1)$-*wf.*

At this point it is appropriate to remark that the $\{\mathsf{K}_n\}$ is not the theory of types (simple or ramified), and it is not included within the theory of types. For in the K_n, $(n > 1)$, all instances of ${}_1x_{(n-1)}^k,\ {}_1x_n^j \supset {}_1x_{(n-1)}^k,\ {}_1x_n^j$ are *theorems*, whereas the corresponding expressions in type theory (something like: ${}_1x_n^j \in {}_1x_{(n-1)}^k \supset {}_1x_n^j \in {}_1x_{(n-1)}^k$) are not even well-formed.

3 *Consistency of* K_n

3.1 *The consistency argument* We require an effective map $\mathbf{t}$ which takes $\{{}_ix_j^k\} \xrightarrow{1-1} \{{}_ix_1^m\}$. This is easily secured in terms of an effective map ϕ: $\{k, j\} \xrightarrow{1-1} N$. The map ϕ can be, e.g., a simple Gödel numbering code, such as: $\phi(\langle k, j\rangle) = 2^k \cdot 3^j$. Since the exact choice of ϕ is a matter of no consequence for our purposes, we leave it unspecified.

The translation function $\mathbf{m}$ is defined in terms of $\mathbf{t}$. The precise definition of $\mathbf{t}$ is:

$$\mathbf{t}({}_ix_j^k) = {}_ix_1^{\phi(\langle k, j\rangle)}$$

$\mathbf{m}$ is defined as follows:

where $A = {}_{i_1}x_{j_1}^{k_1}, \ldots, {}_{i_n}x_{j_n}^{k_n}$, $\mathbf{m}(A) = {}_n x_2^{\alpha}, \mathbf{t}({}_i x_j^k), \ldots, \mathbf{t}({}_{i_n}x_{j_n}^{k_n})$,

where $A = \sim B$, B n-wf, $\mathbf{m}(A) = \sim\mathbf{m}(B)$,

where $A = (B \supset C)$, B, C n-wf, $\mathbf{m}(A) = (\mathbf{m}(B) \supset \mathbf{m}(C))$,

where $A = ({}_i x_j^k)\, B$, B n-wf, $\mathbf{m}(A) = \mathbf{t}({}_i x_j^k)\mathbf{m}(B)$.

We are now in a position to state Theorem 1.

Theorem 1 *For all n, $\mathbf{K}_n$ is consistent.*

Proof: The reader may be spared a rigorous inductive argument to show that $\mathbf{m}$ is provability preserving, and yet be convinced by the following considerations: It is easily verified that $\mathbf{m}$ takes axiom schemata to axiom schemata. Truth-functional structure is obviously preserved. Instances of axioms in which multiple quantification occurs will be mapped in such a way as to keep distinct variables distinct, since $\mathbf{t}$, the function from n-order variables to first-order variables is one-one. Hence, if an n-wf A contains variables ${}_i x_n^k$ and ${}_m x_n^j$, ${}_i x_n^k$ will be free for ${}_m x_n^j$ in A just in case $\mathbf{t}({}_i x_n^k)$ is free for $\mathbf{t}({}_m x_n^j)$ in $\mathbf{m}(A)$. Thus, quantificational structure will be preserved, and so axiomhood is preserved. The rule of inference, *modus ponens*, is the same for all the $\mathbf{K}_n$. Hence $\mathbf{m}$ is provability preserving. With a view to *reductio*, suppose that there is a $\mathbf{K}_n$ which is inconsistent. Then there exists an n-wf, A, such that A and $\sim A$ are both provable in $\mathbf{K}_n$. But then $\mathbf{m}(A)$ and $\mathbf{m}(\sim A)$ will both be provable in $\mathbf{K}_1$. Since $\mathbf{m}(\sim A) = \sim\mathbf{m}(A)$, it follows that $\mathbf{K}_1$ is inconsistent, which is absurd.

3.2 *Quine's derivation* It is instructive to apply the translation $\mathbf{m}$ to Quine's derivation (in [3]) of an inconsistent second-order formula. Quine's axiomatization of first-order logic differs from the one treated here, and the notation is different as well. We may ignore these differences, and focus on one inference in his derivation which is erroneous. For simplicity, let us suppose that we have extended all our definitions in the obvious way in order to cope with existential quantification, expressed '$(\mathsf{E}\, {}_i x_j^k)$', and the biconditional, expressed '$\equiv$'.

Expressed in our notation, the interesting inference is:

$$(\mathsf{E}\, {}_1x_2^3)({}_1x_2^2)({}_1x_2^3,\ {}_1x_2^2 \equiv {}_1x_2^1,\ {}_1x_2^2) / \therefore (\mathsf{E}\, {}_1x_2^3)({}_1x_2^2)({}_1x_2^3,\ {}_1x_2^2 \equiv \sim {}_1x_2^2,\ {}_1x_2^2).$$

(This corresponds to the inference from the fifth to the sixth line of Quine's derivation in [3].) In order to effect the translation, let:

$$\begin{aligned} \mathbf{t}({}_1x_2^1) &= {}_1x_1^{\phi(\langle 1,2\rangle)} = {}_1x_1^i \\ \mathbf{t}({}_1x_2^2) &= {}_1x_1^{\phi(\langle 2,2\rangle)} = {}_1x_1^j \\ \mathbf{t}({}_1x_2^3) &= {}_1x_1^{\phi(\langle 3,2\rangle)} = {}_1x_1^k \end{aligned}$$

Notice that i, j, and k are mutually distinct, by the one-oneness of ϕ, and so the three first-order variables which are the values, under $\mathbf{t}$, of the three second-order variables, are mutually distinct as well. The translation, under $\mathbf{m}$, of the inference being examined becomes:

$$(\mathrm{E}\,{}_1x_1^k)({}_1x_1^j)({}_2x_2^\alpha,\ {}_1x_1^k,\ {}_1x_1^j \equiv {}_2x_2^\alpha,\ {}_1x_1^i,\ {}_1x_1^j)/\therefore(\mathrm{E}\,{}_1x_1^k)({}_1x_1^j)({}_2x_2^\alpha,\ {}_1x_1^k,\ {}_1x_1^j \equiv \sim{}_2x_2^\alpha,\ {}_1x_1^j,\ {}_1x_1^j).$$

This is an invalid first-order inference, as one can see by considering the same inference expressed in more familiar notation:

$$(\mathrm{E}\,x)(y)(\phi xy \equiv \phi zy)/\therefore(\mathrm{E}\,x)(y)(\phi xy \equiv \phi yy).$$

In $\mathbf{K}_1$, as in any consistent first-order logic, derivation of such an inference is blocked by restrictions on newly substituted variables being bound by already present quantification.

4 *Dispensability* Having established that the "natural generalizations" $\mathbf{K}_n$, are "safe," in this section we argue that they are also unnecessary. As an initial step in the argument, we have:

Theorem 2 *The inverse,* $\mathbf{m}^{-1}$, *of the translation function* $\mathbf{m}$, *is provability-preserving.*

Proof: Again the reader may be spared the details. The convincing considerations are essentially the same as those presented in the proof of Theorem 1, with the one-oneness of $\mathbf{t}$ assuring preservation of quantificational structure by $\mathbf{m}^{-1}$. It is to be noted that $\mathbf{m}$ is *strictly into* $\mathbf{K}_1$, and so not all 1-wfs will have n-wf translations under $\mathbf{m}^{-1}$. Truth-functional structure, axiomhood, and derivability under *modus ponens* are all preserved. Hence $\mathbf{m}^{-1}$ is provability-preserving.

We recall that $\mathbf{m}$ was defined so as to be effective. It follows from the effectiveness and one-oneness of $\mathbf{m}$ that $\mathbf{m}^{-1}$ is effective as well. Hence, in $\mathbf{m}$ we have an effective translation from $\mathbf{K}_n$ to $\mathbf{K}_1$ with the property that an n-wf will be provable in $\mathbf{K}_n$ if and only if its translation is provable in $\mathbf{K}_1$. It follows that any inference problem which is solvable in a "natural generalization" $\mathbf{K}_n$ can also be solved in $\mathbf{K}_1$; what is more, the move from $\mathbf{K}_n$ to $\mathbf{K}_1$ is effective. The moral may be drawn that the pure quantificational structure of the first-order predicate calculus is formally no weaker than that of higher order "natural" generalizations. Those higher order theories would seem to be dispensable in favor of first-order logic.

As an illustrative example, consider the motivating example given above (p. 257). Again allowing ourselves the luxury of some additional truth-functional connectives and more familiar notation, we find that the example becomes, under translation, an argument of the form:

$$\begin{aligned}&(x)((\phi xy \cdot \phi zx) \supset \phi xw)\\ &\phi zu\\ &\phi uy/\therefore\ \phi uw.\end{aligned}$$

This form is certainly one which is valid, and so provable in any complete first-order logic, such as $\mathbf{K}_1$.

Theorem 2 provides a kind of completeness proof for the $\mathbf{K}_n$. For, by Theorem 2, any n-wff A which is not provable in $\mathbf{K}_n$ has a first-order translation $\mathbf{m}(A)$ which is not provable in $\mathbf{K}_1$. By the completeness of $\mathbf{K}_1$, $\{\sim\mathbf{m}(A)\}$ has a model. This model will also be a model of $\sim A$, though perhaps not a "standard" model. (I am supposing that a "standard" model is one in which, if ${}_i x_{j_1}^{k_1}, \ldots, {}_m x_{j_i}^{k_i}$ is true in the model, then the membership relation obtains between $\left\langle {}_n x_{j_2}^{k_2}, \ldots, {}_m x_{j_i}^{k_i} \right\rangle$ and ${}_i x_{j_1}^{k_1}$.) The completeness property assured by Theorem 2 is perhaps too weak, but surely we ought not to expect any higher order formal logic to capture all the truths of its "standard" set-theoretic model.

REFERENCES

[1] Copi, I. M., *Symbolic Logic*, third edition, Macmillan, New York (1967).

[2] Mendelson, E., *Introduction to Mathematical Logic*, Van Nostrand, Princeton, New Jersey (1964).

[3] Quine, W. V., "On universals," *The Journal of Symbolic Logic*, vol. 12 (1947), pp. 74-84.

The University of North Carolina at Charlotte
Charlotte, North Carolina