184
Notre Dame Journal of Formal Logic
Volume XVII, Number 2, April 1976
NDJFAM

NEGATION DISARMED

ROBERT K. MEYER

The purpose of the present note is to extend the methods of [1] to show, for several interesting systems of quantificational logic, that their negation-free fragments are determined by their negation-free axioms, even in the presence of strong classical negation axioms. Among these systems, as in [1], are the relevant logics of Anderson and Belnap, presented here in their first-order versions **RQ**, **EQ**, etc. We generalize the results of [1] to the extent that they apply here not merely to positive logics $\mathbf{L}^+$ but to positive $\mathbf{L}^+$-theories; i.e., it turns out for the relevant logics (and some others) that the set of negation-free theorems of a first-order theory all of whose non-logical axioms are negation-free is completely determined on applying negation-free logical axioms and rules to these non-logical axioms.

Aside from its intrinsic interest, the point of this result lies in the fact that the negation-free part of the relevant logics is intuitionistically acceptable, though its negation axioms are not. This acceptability extends to possession of certain interesting structural properties, e.g., if $A \vee B$ is a negation-free theorem of one of the relevant logics, so is at least one of A and B, as was noted at the sentential level in [2]; similarly, as is to be shown in a paper in preparation, if $\exists x A(x)$ is a negation-free theorem, so is an instance $A(t)$ for some term t; both properties, of course, are intuitionistic. What we want to show, accordingly, is that there are no theorems in the constructively acceptable negation-free parts of the relevant logics that are only provable by constructively unacceptable methods, i.e., by detours through the properties of classical negation. (The point is unlikely to be missed, but the claim is that relevant logics have certain *formal* properties that are intuitionistically acceptable; as is usual in these matters, no such claim is entered for the *informal* arguments employed to establish this result.)

The reader is presumed to have access to [1], and so its methods and terminology are used freely. Arguments, by and large, are old, being adapted here only as is necessary for the richer context. References to axioms (e.g., A1), etc., are to [1].

Received October 13, 1971

1 We deal here only with first-order languages $\mathcal{J}$ of a particularly simple kind; they may be complicated to taste in well-known ways without disturbing the arguments. We assume that $\mathcal{J}$ is built up from denumerably many individual variables, for which we use 'x', etc., and from an unspecified number of individual and predicate constants, the latter of any degree from 0 on. *Terms* and *formulas* are built up without restriction, using the quantifiers '( )' and '$\exists$', the connectives &, $\vee$, $\rightarrow$, and perhaps other sentential connectives and constants. 'A', 'B', etc., are our syntactical variables for formulas; 't', 'u', etc., for terms; $A(t/u)$ is the result of *proper* substitution of t for free u in A, bound variables being rewritten if necessary according to some definite plan to avoid confusion. A *sentence* is a formula in which no individual variable occurs free. We shall in general identify a language with the set of its formulas.

Let $\mathbf{LQ}^+$ be a *positive quantificational logic*, i.e., one whose language $\mathcal{J}^+$ does not contain the negation sign -. The *negation completion* $\mathbf{LQ}$ of $\mathbf{LQ}^+$ is defined following [1] as the result of adding negation to the formation apparatus to get an enriched language $\mathcal{J}$, and of taking as new axioms (i) all instances of old axiom schemes in the richer language, and (ii) all instances of the strong classical axiom schemes A1-A3 of [1], in conclusion closing this doubly enriched set of axioms under the rules of $\mathbf{LQ}^+$.

As in [1], certain quantificational logics are *rigorous*. The *basic positive* one, $\mathbf{BRQ}^+$, builds on the corresponding sentential logic $\mathbf{BR}^+$ of [1], having a language $\mathcal{J}^+$ with primitive connectives &, $\vee$, $\rightarrow$, with rules of *modus ponens* for $\rightarrow$, adjunction for &, and universal generalization for ( ). Its axioms are all instances of the schemes B1-B9 of [1], together with all instances of the following quantificational schemata.

B10. $(x)A \rightarrow A(t/x)$.
B11. $A(t/x) \rightarrow \exists xA$.
B12. $(x)(A \rightarrow B) \rightarrow ((x)A \rightarrow (x)B)$.
B13. $(x)(A \rightarrow B) \rightarrow (\exists xA \rightarrow \exists xB)$.
B14. $A \rightarrow (x)A$, *if x is not free in A*.
B15. $\exists xA \rightarrow A$, *if x is not free in A*.

A quantificational logic $\mathbf{LQ}^+$ is a *positive rigorous logic* if it can be formulated with the same connectives, axiom schemes, and rules of inference as $\mathbf{BRQ}^+$, with perhaps as additional axiom schemes one or more of C1-C7 of [1] or of the following (C8 and C8′ go together, both to be adopted or neither):

C8. $(x)(A \vee B) \rightarrow ((x)A \vee B)$, *if x is not free in B*.
C8′. $(\exists x A \;\&\; B) \rightarrow \exists x(A \;\&\; B)$, *if x is not free in B*.
C9. $(x)\mathsf{N}A \rightarrow \mathsf{N}(x)A$.[1]

Finally, $\mathbf{LQ}$ is a *rigorous logic* if it is the negation-completion of some positive rigorous logic $\mathbf{LQ}^+$. Examples of positive rigorous logics are the systems $\mathbf{EQ}^+$, formulated with sentential axioms as in [1] and *all* the above quantificational schemes, and $\mathbf{PQ}^+$, sentential axioms as in [1] and all the above schemes but C9 as quantificational axioms. The Anderson-Belnap systems **EQ** of entailment and **PQ** of ticket entailment, with quantification, are the negation-completions of $\mathbf{EQ}^+$ and $\mathbf{PQ}^+$ respectively; for remarks about other rigorous logics, consult [1].

2 With each first-order logic $\mathbf{LQ}^+$ or **LQ** we associate the corresponding sentential logic $\mathbf{L}^+$ or **L** that results from deleting quantificational machinery; e.g., we associate **E** with **EQ**. Then the notions of *possible* $\mathfrak{L}$-*matrix* and $\mathfrak{L}$-*matrix* as defined in [1] make sense. Let **LQ** be a first-order logic (with or without negation, for the moment). Let S be any set of *sentences* of the language of **LQ**. Then $[S]$ shall be the corresponding *regular* **LQ**-*theory*, i.e., $[S]$ shall be the smallest set of *formulas* of **LQ** that contains all members of S, all theorems of **LQ**, and which is closed under *modus ponens* and adjunction.[2] A set T of formulas is an **LQ**-theory if it is $[S]$ for some set S of sentences of the language of **LQ**, and nothing else is an **LQ**-theory. There is a smallest **LQ**-theory (in the sense explained in the footnote), namely the set (which we simply call **LQ**) of theorems of **LQ**. Certain technical purposes are served by the definitions chosen, worth calling to the reader's attention (and dropping if other purposes were to be served). As characterized, members of S serve as non-logical axioms for the theory $[S]$; restricting S, though not $[S]$, to sentences assures us that 'x', etc., will function as ambiguous names. Note, too, that each member of $[S]$ has a *derivation* from members of S and theorems of **LQ**, and that moreover, though we have not specifically required theories to be closed under generalization, in all logics thus far (and to be) considered, generalization holds anyway, by a simple inductive argument on length of derivation.

Let **LQ** be a first-order logic. Sometimes we look at **LQ**-theories T as matrices $\langle F, O, T\rangle$, where F is the set of formulas of the language of **LQ** and O is its set of connectives; T is then the set of designated elements of the matrix. The ambiguity being harmless, we let context determine whether 'T' means the matrix $\langle F, O, T\rangle$ or its third (and critical) member. For the logics **LQ** considered here, considered as a matrix each **LQ**-theory T is trivially an $\mathfrak{L}$-matrix in the sense of [1], i.e., it validates the axioms and rules, and hence the theorems, of the sentential base **L** of **LQ**. As a matrix, we call **LQ** itself the *canonical* $\mathfrak{LQ}$-*matrix*.

Let $\mathfrak{LQ}^+$ be a positive quantificational logic, with language $\mathcal{F}^+$, and let $[S]$ be an $\mathbf{LQ}^+$-theory. Let **LQ** be the negation-completion of $\mathbf{LQ}^+$, and let $[[S]]$ be the corresponding **LQ**-theory determined by S. We call $[[S]]$ then the *negation-completion* of $[S]$. (In context, if we are just thinking of the theories and not of axioms for them, we shall write T^+ for $[S]$ and correspondingly T for $[[S]]$.) Clearly $S \subseteq [S] \subseteq [[S]]$; if $[S] = F^+ \cap [[S]]$, we call $[[S]]$ a *conservative extension* of $[S]$. (Intuitively, $[[S]]$ is what one gets out of $[S]$ when as explained in motivating remarks one is allowed to use non-constructive arguments (as respects negation) to derive additional theorems. The result for which we are aiming, where **LQ** is a rigorous logic, is that one gets essentially nothing new, leaving classically acceptable arguments at one's service in the derivation of constructively acceptable results.) In fact, for rigorous logics, the result is at hand.

Theorem 1 *Let* $\mathbf{LQ}^+$ *be a positive rigorous logic, and let* T^+ *be a regular* $\mathbf{LQ}^+$*-theory. Then the negation-completion* T *of* T^+ *is a conservative extension of* T^+*. As a special case, the rigorous logic* $\mathbf{LQ}$ *is a conservative extension of* $\mathbf{LQ}^+$*.*

Proof: Trivially, T^+ is an $\mathfrak{L}^+$-matrix. Form the rigorous enlargement $\langle M, O, D\rangle$ of T^+ as in section **4** of [1]. One gets $\langle M, O, D\rangle$, it will be recalled by adding (intuitively) the least and greatest elements O and I to the set F^+ of formulas of $\mathbf{LQ}^+$, and then, to account for negation, adding a copy of each element in the enlarged F^+, getting M; the designated subset D of M consists of T^+, the element I, and all the copies. Moreover $\langle M, O, D\rangle$, by the proof of Theorem 1 of [1], is an $\mathfrak{L}$-matrix.

It suffices now to finish the proof of Theorem 1 to find an interpretation I of the set F of formulas of $\mathbf{LQ}$ in $\langle M, O, D\rangle$ such that (1) $I(A) \in D$ for every member of T, and (2) $I(A) \notin D$ for every member of $F^+ - T^+$. Operations corresponding to the connectives of $\mathbf{LQ}$, including -, being defined on M, the trick is to interpret the quantifiers. We do so, in effect, by recursively defining I on F as follows:

(a) $I(A) = A$, if A is atomic.
(b) $I(A \,\&\, B) = I(A) \,\&\, I(B)$.
(c) $I(A \vee B) = I(A) \vee I(B)$.
(d) $I(\bar{A}) = -I(A)$.
(e) $I(A \rightarrow B) = I(A) \rightarrow I(B)$.
(f) If $I(A)$ is one of O, I, -O, -I, $I((x)A) = I(A) = I(\exists xA)$.
(g) If $I(A) \in F^+$, $I((x)A) = (x)I(A)$ and $I(\exists xA) = \exists xI(A)$.
(h) If $-I(A) \in F^+$, $I((x)A) = -\exists x - I(A) \,\&\, I(\exists xA) = -(x) - I(A)$.

Since members of F^+ are formulas, so that prefacing quantifiers makes sense, and since all cases have been covered it is clear that I is well-defined. Moreover, since $I(A) = A$ is obvious for all formulas A of F^+, that (2) above holds is trivial. Accordingly, we may finish the proof of the theorem by showing that (1) holds.

If $A \in T$, it has a derivation from theorems of $\mathbf{LQ}$ and members of the subset S of F^+ such that $[[S]] = T$ using *modus ponens* and adjunction. It is already noted that if $A \in S$ then $I(A) \in D$, while on the inductive hypothesis that premisses of rules belong under I to D, the $\mathfrak{L}$-matrixhood of $\langle M, O, D\rangle$ assures that their conclusions will. It remains only to show that if A is a theorem of $\mathbf{LQ}$, $I(A) \in D$. In this case, A has a derivation also, from logical axioms, by *modus ponens*, adjunction, and generalization. Sentential axioms and rules are once more taken care of, in the latter case again on inductive hypothesis, by the $\mathfrak{L}$-matrixhood of $\langle M, O, D\rangle$. Suppose A comes by generalization and, on inductive hypothesis, $I(A) \in D$. If $I(A)$ is I, -I, or -O, it is trivial by (f) that $I(A) \in D$; if $I(A) \in T^+$, since generalization holds as noted in T^+, by (g) $I((x)A) \in T^+ \subseteq D$. If $-I(A) \in F^+$, then by (h) $-I((x)A) = \exists x - I(A) \in F^+$, whence $I((x)A) \in D$.

I turn finally to verification of logical, quantificational axioms. As in [1], this is an argument by cases, of which I do some typical ones and leave the rest to the reader.

Ad B11. We neglect the clause in the definition of $A(t/x)$ which rested on rewriting bound variables, forthcoming from the other axioms in any event. We note next that $I(A)$ and $I(A(t/x))$, for any term t properly substitutable for x in A (without rewriting) always end up in the same place, i.e., under I both together are members of F^+, are negations of members of F^+, or are identical to the same constant among O, I, -O, -I. (This is proved by induction on the length of A.) This leaves as the most non-trivial case in the verification of B11 the case $-I(A) \in F^+$, in which case by (h) and the fact just noted $I(A(t/x) \rightarrow \exists xA) = I(A(t/x)) \rightarrow -(x) - I(A)$, which by definition of rigorous enlargement is $(x) - I(A) \rightarrow -I(A(t/x))$; another tedious inductive argument, omitted here, shows this to be an instance of B10 and so a member of T^+, ending the verification.

Ad C8. We must show $I((x)(A \vee B)) \rightarrow (I((x)A) \vee I(B)) \in D$. The hard case is when neither $I(A)$ nor $I(B)$ is a constant. Suppose first that $I(A)$ is the negation of a member of F^+ and $I(B) \in F^+$. Then $-I(A \vee B) = -I(A) \in F^+$ by definition of rigorous enlargement, whence left and right sides of C8 both reduce to $-\exists x - I(A)$; the case $I(A), -I(B) \in F^+$ is similar. If both $I(A)$, $I(B) \in F^+$, the left side of C8 becomes $(x)(I(A) \vee I(B))$ and the right becomes $(x)I(A) \vee I(B)$; noting that x is free in $I(B)$ only if free in B, C8 belongs to T^+ as an instance of the logical axiom C8 of **LQ**$^+$. The case $-I(A), -I(B) \in F^+$ is similar, requiring some contraposition by appeal to definition in [1] of rigorous enlargement and verified in the end as an instance of C8′. (This explains, in case the reader wondered, which is probably doubtful, why we required C8′ to hold if C8 does.)

Ad C9. The point, consulting definitions and [1], is that $I(\mathsf{N}A)$ must be O, a member of F^+, or I, as $I(A)$ is O, a member of F^+, or anything else. In the final case, the consequent goes to I, which suffices; in the first case, the antecedent goes to O, which suffices. The middle case reduces to an instance of C9, which completes the verification of C9 and the proof of Theorem 1.

3 Theorem 1 shows that the conservative extension properties of rigorous logics extend to their quantificational variants, with respect to negation. This includes all the relevant logics but the system **RQ**; we recall that the system **R** of relevant implication, the strongest and most interesting of the relevant logics, was not rigorous and accordingly required special technique in [1] to establish the conservative extension property. It also required the introduction of an intuitionistically acceptable connective of intensional conjunction, though this turns out to be eliminable by the somewhat different methods of [4] and [5]; we retain here, for simplicity, the formulation of **R**$^+$ and its negation completion **R** of [1], getting **RQ**$^+$ and **RQ** by adding B10-B15, C8-C8′ (C9 proving redundant and, since **R** is non-modal, uninteresting.) The hard work having been done, either above or in [1], we have almost immediately

Theorem 2 *Let T^+ be a regular* **RQ**$^+$*-theory. Then the negation-completion T of T^+ is a conservative extension of T and, as a special case,* **RQ** *is a conservative extension of* **RQ**$^+$.

Proof: Like Theorem 1, *mutatis mutandis*.

We leave the reader with a thought, and a corollary. The thought is that, although in this and related papers we have been proving Brouwerian facts about relevant logics, in the sense that relevant logics are in many ways closely related to intuitionist logic and share many of its most striking properties, these Brouwerian facts have Hilbertian import. For the fact that, unlike intuitionist calculi, relevant logics do not break down in the presence of classical negation suggests that there may have been something wrong with intuitionist formulation of intuitionist insight in the first place, and that a correction which preserves what is right about the formulation is most welcome to, of all people, the formalists.

What is wrong with intuitionist logic, of course, is its indifference to relevance as a motivating condition in the analysis of inference. Heyting et al. are hardly to be blamed for this; even the beginnings of a clear formal approach to the problem of relevance required twenty years more, and probably would not have been undertaken at all but for the light thrown on the deduction theorem by the intuitionist sentential calculus. But, with the advantage of hindsight, it may very plausibly be argued that it is precisely disregard for relevance, not the success of intuitionist logic in catching intuitionist views about, say, excluded middle or double negation, that makes it impossible for intuitionist *logic* to accommodate classical negation.[3] Take, e.g., the classical theorem $p \vee (p \rightarrow q)$, which is certainly intuitionistically unacceptable but which would become provable, even in Johansson's minimal logic, in the presence of double negation and excluded middle; indeed, double negation alone will do.

An examination of an intuitive proof of the formula in question, which pretty well reconstructs the formal proof in Johansson's system, shows what is going wrong. Suppose p is either true or false. In the latter case, clearly there is something false that it implies—itself, for example. Letting f be the intuitive disjunction of all falsehoods, accordingly p is either true or p implies f. (If we had analyzed negation inferentially to begin with, which seems more in the spirit of Johansson and indeed of Heyting, we would have started here.) But f implies that q's implying f implies f, whence, by double negation, f implies q. (Intuitionist negation being absurd to begin with, in Curry's phrase, this step is not necessary for it.) Accordingly, by elementary properties of disjunction and implication, either p is true or p implies q.

This argument fails, relevantly, on the point $f \rightarrow q$. Once relevance is considered at all, it certainly is not to be accepted as an evident principle, as in intuitionist logic proper. Nor does it follow from double negation, since the necessary theorem, $f \rightarrow ((q \rightarrow f) \rightarrow f)$, is an instance of the relevantly fallacious $A \rightarrow (B \rightarrow A)$. Note, however, that what one might take to be the central feature of a constructive analysis of negation, i.e., that it be analyzed inferentially, is quite compatible with relevant insights. Nor do any basic relevant insights *require* double negation or excluded middle; the point is that they do not *forbid* them, and insofar as there is an argument that a constructive implication *ought* to forbid such principles on pain of collapse, it is fallacious.

It is, I have suggested, *formalism* which gets a boost from results like the present one. For formalists, no less than intuitionists, are attached to constructive techniques (though differently). Unlike the intuitionist, however, the formalist has no objection to non-constructive arguments, so long as he is assured that he gets into no trouble. That is what we have assured him. To balance accounts, we conclude with a Brouwerian fact.

Corollary 2.1 *Suppose* $A \vee B$ *is a theorem of* **RQ**, *where* A *and* B *are negation-free; then one of* A *and* B *is a theorem of* **RQ**. *Moreover, if* A *is negation-free in* **RQ**, *and if* $\vdash_{\mathbf{RQ}} \exists x A$, *then for some term* t, $\vdash_{\mathbf{RQ}} A(t/x)$.

Proof: By Theorem 2 above, [2], and Theorem 5 of [3].[4]

NOTES

1. 'N' is presumed defined, as in [1] (following Anderson-Belnap) by $\mathsf{N}A =_{Df} (A \to A) \to A$. Anderson and Belnap take C9 as an axiom of **EQ**, but, as is well-known, there is philosophical debate about whether it, or its converse, is wanted; I do not join this debate here, noting merely that the converse of C9 holds for **EQ** and that, were we to follow Ackermann in principle by adding a sentential constant t, the intuitive conjunction of all logical truths, and to formalize $\mathsf{N}A$ as $\mathsf{t} \to A$, both C9 and its converse would hold without special quantificational assumptions. On other technical maneuvers, e.g., taking 'N' as primitive—C9 would of course fail.

2. 'Regular **LQ**-theory' means what on the usage of [3] and preceding papers would have been called just '**LQ**-theory'. But Routley, as is noted in [4], has discovered a use for *irregular* theories—theories that do not contain all the logical truths; no such use is relevant here, so that here '**LQ**-theory' is always tacitly understood as 'regular **LQ**-theory'.

3. For a beautiful analysis of intuitionist and related negations, *cf*. H. B. Curry's chapter on negation in his *Foundations of Mathematical Logic*, McGraw-Hill, New York (1963).

4. My thanks are due to the National Science Foundation, for partial support of this research through grant GS-2648.

REFERENCES

[1] Meyer, R. K., "On conserving positive logics," *Notre Dame Journal of Formal Logic*, vol. XIV (1973), pp. 224-236.

[2] Meyer, R. K. "On relevantly derivable disjunctions," *Notre Dame Journal of Formal Logic*, vol. XIII (1972), pp. 476-480.

[3] Meyer, R. K., J. M. Dunn, and H. Leblanc, "Completeness of relevant quantification theories," *Notre Dame Journal of Formal Logic*, vol. XV (1974), pp. 97-121.

[4] Routley, R., and R. K. Meyer, "The semantics of entailment, I," in H. Leblanc, ed., *Truth, Syntax, Modality*, North-Holland, Amsterdam (1973).

[5] Meyer, R. K., "Conservative extension in relevant implication," *Studia Logica*, vol. 31 (1973), pp. 39-48.

The Australian National University
Canberra, Australia