Hacking's Logic

Sundholm, B.G.

HACKING'S LOGIC*

IN his recent study "What Is Logic?"† Ian Hacking attempts to delineate logic by means of a fusion of technical proof theory and standard semantical concepts. The technicalities from proof theory are presented in a deceptively simple way; the initial heuristic idea is obvious enough, but soon Hacking is forced into the notoriously slippery ground of "finitistic" transformations of infinite proof trees. This is due to some constraints on the rules, which are called for in order to exclude unwanted candidates—such as modal logic—from the field of logic proper. Throughout, the paper bears witness to a laudable wish to take into account recent, fairly advanced results from mathematical logic, e.g., Lindström's theorem. But this very comprehensiveness has sometimes led Hacking to a certain unclarity and imprecision, and some remarks of a technical nature may be appropriate. Thus it is the purpose of the present note to spell out in some detail, and with greater precision, the basic concepts of the Hacking setup, such as "semantic framework" and "conservativeness." The role of standard technical results, e.g., "finitistic cut-elimination" (an example which is given heavy emphasis by Hacking), will be reconsidered and sometimes found different from what Hacking envisages. It is in the nature of things that such criticism is of a mainly destructive nature; constructive criticism of Hacking's paper can be found in Christopher Peacocke's "Two Comments".[1]

I. "CONSERVATIVENESS" AND ELIMINATION THEOREMS

Hacking (311) considers a "classical semantic framework" (L, $\vdash$), where L is a language satisfying:

(L1) L is *bivalent;* i.e., "every sentence . . . should be assigned the value true or false but not both" (311).

(L2) *Logical consequence* between finite subsets of L, in symbols $\models_L$, satisfies "Θ is a logical consequence of . . . Γ if no matter what values are assigned to the members of Γ, Θ, some member of Θ is true when every member of Γ is true" (311).

The relation $\vdash$ on the finite subsets of L satisfies the following two conditions:

*I wish to thank Christopher Peacocke and Dana Scott for helpful comments.

†This JOURNAL, LXXVI, 6 (June 1979): 285–318. References in brackets will be to this paper.

[1] This JOURNAL, this issue, pp. 168–175.

0022-362X/81/7803/0160$00.90

($\vdash$ 1) $\vdash$ is a *deducibility relation*, in symbols Ded ($\vdash$); i.e., the three characteristic properties

Reflexivity	$A \vdash A$
Dilution	If $\Gamma \vdash \Theta$, then $\Gamma, A \vdash \Theta$ and $\Gamma \vdash A, \Theta$.
Transitivity (Cut)	If $\Gamma \vdash A, \Theta$ and $\Gamma, A \vdash \Theta$, then $\Gamma \vdash \Theta$.

hold for all $A \epsilon$L and all finite subsets Γ, Θ of L (see 293).

($\vdash$ 2) $\vdash$ is sound for logical consequence in L, in symbols $\vdash \subseteq \models_L$ (292).

The language L can be extended to L′ by the addition of logical constants, say, for the sake of simplicity, &, with the obvious formation rule for conjunction. A relation $\vdash'$ between the finite subsets of L′ is then defined inductively:

(0) $\vdash \subseteq \vdash'$
(1) If $\Gamma \vdash' \Delta, A$ and $\Gamma \vdash' \Delta, B$, then $\Gamma \vdash' \Delta, A \mathbin{\&} B$.
(2) If $A, \Gamma \vdash' \Delta$, then $A \mathbin{\&} B, \Gamma \vdash' \Delta$.
(3) If $B, \Gamma \vdash' \Delta$, then $A \mathbin{\&} B, \Gamma \vdash' \Delta$.

This is, in effect, what goes on in sections VI and VII, although Hacking fails to bring out the crucial fact that *two* relations, $\vdash$ and $\vdash'$, in *different* languages are involved. Evidence for this interpretation of Hacking can be found in, among other passages, page 296, lines 20–24, and page 314, line 4 from the bottom.

Hacking then considers a notion of "conservativeness" (296), which, with our $\vdash / \vdash'$ distinction, can be spelled out as:

(C1) For any Γ, Δ in L, if $\Gamma \vdash' \Delta$, then $\Gamma \vdash \Delta$.
(C2) For any Γ, Δ in L, if $\Gamma \vdash \Delta$, then $\Gamma \vdash' \Delta$.

For (C1) compare "No more . . ." (296, lines 9-11), and for (C2), "Nothing less . . ." (line 11). It is emphatically claimed that "*Proving cut-elimination is one ingredient in showing that the operational rules are conservative definitions*" (296) and that "Cut-elimination, dilution-elimination, and identicals-elimination (all for complex formulas) are necessary conditions for [conservativeness of the definitions of the logical constants]" (298). Unfortunately, this is not correct, because conservativeness of $\vdash'$ over $\vdash$ is readily established without having to prove the elimination results. We can, for example, carry over to the present context the trivial model-theoretic observation that, for first-order theories T,T′, if every model of T is expandable to a model of T′, then T′ is conservative over T, and hence there are other ways to prove conservativeness which do not use cut-elimination. This proof of conservativeness is prima facie not finitistic—it uses completeness—and perhaps Hacking can modify his claim that cut-elimination is the only way to obtain conservativeness to claim rather that it is the only *finitistic* way. At the present state of knowledge in proof theory, however, we do not possess even a framework in which questions like "Is cut-elimination the only way in which one can prove conservativeness finitistically?" can be formulated and answered. This would be a part of the subject, call it *abstract proof theory,* which would stand to proof theory as so-called "soft" model theory stands to model theory. (Note that without the subformula property it is not even clear that the elimination theorems are sufficient for conservativeness; however, the subformula property and the elimination theorems together entail conservativeness.)

No, the proper role for the elimination theorems in the Hacking setup is, not to establish conservativeness, but to guarantee that $\vdash'$ is a deducibility relation. Certainly Ded($\vdash'$) has to be established, and for this task the available tools are the definition of $\vdash'$ and the fact that Ded($\vdash$) holds. So one must show that, say, Dilution is a derivable rule for $\vdash'$; but this is nothing but proving the elimination result for Dilution. And so on.

Hacking sums up this part of his analysis in the opinion that "operational rules introducing a constant should (i) have the subformula property, and (ii) be conservative with respect to the basic facts of deducibility. The second clause means . . . provability of the elimination theorems" (304). The rider concerning (ii) appears to give a different *definition* of conservativeness from the earlier one; and, though the claims about the elimination theorems, by definition, are trivially correct for this latter reading, it seems a most awkward use of the term. Much of the unclarity in the paper seems to stem from Hacking's conflation of these two notions of conservativeness.

Thus he claims that he "can give no defense of [the subformula property] as a requirement that definitions should be conservative" (299). With the second reading of conservativeness this claim seems justified; conservativeness then simply *means* provability of elimination theorems, and nothing must necessarily hold about the subformula property. With the original reading, however, the subformula property and the elimination theorems are "defended" from conservativeness to exactly the same degree: neither alone will suffice for conservativeness, but together they are sufficient.

A semblance of a case for (i) and (ii)—with the original reading of conservativeness—may perhaps be put together along the following lines. *The subformula property* is defended from the general requirement on definitions that they be *non-impredicative;* if the operational rules for the constant K fail to have the subformula property, the inductive clauses for the extended relation may fail to yield a well-determined concept. The situation can be illuminated by an analogy with the concept "autological" from Grelling's paradox: an adjective 'abc' is autological if and only if 'abc' is abc. This definition works very well except for 'autological' where we get the result that 'autological' is autological if and only if 'autological' is autological. So the meaning of 'autological' is left underdetermined by the introduction-rule for 'autological'. *The elimination theorems* are, as we have seen, needed to ensure that the extended relation is a deducibility relation.

At this point we may deal with a possible objection. It might be argued that, if the elimination theorems were not provable, we could simply *add* Reflexivity, Dilution, and Transitivity as postulates for the extended relation. This would be a *creative* definition, but only of the deducibility relation. The meaning of the logical constant thus introduced is uniquely given by its truth and falsity conditions as laid down by the left and right subformula-property introduction rules. Plausible as it may seem, such an objection would be question-begging, because Hacking's do-it-yourself semantics for reading off the truth and falsity conditions presupposes that cut-free derivations are available. So until the semantics has been employed to read off the relevant information, the above objection will not take off. But the semantics can be employed only after the elimination theorems have been proved.

II. "FINITISTIC" PROVABILITY OF THE ELIMINATION THEOREMS

It is not enough, Hacking claims, that the elimination theorems simply be true: they must be finitistically provable (295, 309). His reason for this demand is that "only when logic is construed in terms of finitistically provable elimination theorems do we have a theory whose consistency can be established by logic itself" (309). Be that as it may, in order to exclude modal logic, Hacking is led to consider only "local" rules, which pose no restriction on the side formulas. He even claims that "dilution-elimination means that all rules are local" (312). This claim would appear ill founded for at least two reasons: the nonlocal standard first-order quantifier rules admit dilution-elimination as well as the "equivalent" (313) infinitary rules (so perhaps we ought to use 'local' for rules that have "equivalent" local variants as well); and, secondly, even with an emended use of locality, one would have to *prove* that, say, for modal logic there was *no* "equivalent" set of local rules that generated the same theorems.

The standard $\forall$-rules are, as remarked already, nonlocal (*eigen*-parameter condition!), so Hacking is led to consider the "equivalent ω-rule" (313):

$$\text{If, for all } t_i,\ \Gamma \vdash \Delta, At_i, \text{ then } \Gamma \vdash \Delta,\ \forall x Ax.$$

It is a long-lived misconception, to which Hacking unfortunately subscribes (295,309,316), that cut-elimination for the ω-rule in arithmetic presupposes transfinite induction up to the ordinal ϵ_0. This is not so; on the contrary, in primitive recursive arithmetic, PRA, which certainly is a finitistic system, one can prove the following characterization theorem for first-order arithmetic, PA:

THEOREM: PA $\vdash A$ iff PRA $\vdash$ "there is a Kalmar-elementary function that is cut-free ω-proof of A with ordinal height less than ϵ_0."[2]

No transfinite induction is needed for the proof, and the techniques used are entirely finitist. But if one wants *to apply* this cut-elimination theorem, say to prove the consistency of PA, then applications of transfinite ϵ_0-induction are often called for. Indeed, for an arbitrary PA derivation d, one considers its cut-free ω-version d' and proves by induction on the ordinal height of d' that the sequent $\vdash 0 = 1$ does not occur at any node in d'.

The restriction Hacking wants, in order to exclude arithmetic from logic proper, is not to be found by considering the principles used in the proof of the cut-elimination theorem; what is called for is that there be a *uniform* finite bound on the lengths of all the branches of the ω-proofs involved. The above proof techniques will then transform a finitistically given finitist ω-proof tree with a finitistically given uniform finite bound on the height of the tree into a cut-free finitist ω-proof tree, also finitistically of finite height.

Perhaps Hacking can find a way to justify his very strong demand of the self-provability of consistency along these lines. Another—and in my opinion more promising—way of arguing for the finitist provability of the elimination theorems would be based on the traditional topic-neutrality of logic. One of Peacocke's "Comments" tries to fault the Hacking criterion for logical constanthood by considering the same sort of *extensional* counterexamples that have been used to sink early versions of the Davidsonian program.[3] In our present setup for the Hacking strategy such a criticism would appear unfounded; "A logical constant is a constant that can be added to any language of a certain sort" (314), viz., one that is a classical semantic framework. One then has to verify, with the given framework acting as a *parameter,* that the extension of this framework has certain properties, and this verification should be carried out using simple reasoning about the transformation of syntactic signs. It is indeed difficult to see how

[2] Technical support for this and other claims about the ω-rule in arithmetic can be found in my B. Phil. thesis, *The Omega-rule: A Survey,* Oxford University, 1978.

[3] See, e.g., J. A. Foster, "Meaning and Truth Theory," in Gareth Evans and John McDowell, eds., *Truth and Meaning: Essays in Semantics* (New York: Oxford, 1976).

the Peacocke objection can be carried through the heavy intensionality of finitist provability.

III. BIVALENCE AND SEMANTIC FRAMEWORKS

We claimed above that Ded($\vdash'$) had to be established. But why? No argument was given by me, nor can we find one in Hacking's paper. The reason is simple, however. We may wish to add yet another logical constant to the language L′. To do this we must be able to use the same framework properties for (L′, $\vdash'$) that were used for (L, $\vdash$) in the step from $\vdash$ to $\vdash'$. This is a place where Hacking's conflation of the two notions of conservativeness has been at play, one suspects; because if it is realized that in the step from L to L′ certain properties have to be *preserved* (rather than conserved), then it is clear that quite a few more than just dilution, reflexivity, and transitivity have to be preserved. In particular, (L′, $\vdash'$) has to be a classical semantic framework; so (L′1), (L′2), and ($\vdash'$2) have to be established [as well as ($\vdash'$1), which is the provability of the elimination theorems].

With the DIY-semantics from section XVI, (L′2) is taken to be true by definition, but the remaining two properties require proof. We have here a very important point, and it is exactly the place where Hacking's delineation fails to be reductive. Even if the basic prelogical language is constructivist, the bivalence of L′ resulting from the addition of the universal quantifier can be established only nonconstructively. The DIY-semantics for $\forall$ gives the following truth and falsity conditions (where At_i may be assumed bivalent and $At_i \in$ T and $At_i \in$ F are both decidable, for every term t_i):

$$\forall xAx \in \mathrm{T} \quad \text{iff} \quad \forall t_i(At_i \in \mathrm{T})$$
$$\forall xAx \in \mathrm{F} \quad \text{iff} \quad \exists t_i(At_i \in \mathrm{F})$$

To prove bivalence, assume that $\forall xAx \notin \mathrm{T}$. We have to show that $\forall xAx \in \mathrm{F}$.

From the assumption it follows by definition that $\neg(\forall t_i(At_i \in \mathrm{T}))$, which, by the decidability and bivalence of A, is equivalent to $\neg\forall t_i(At_i \notin \mathrm{F})$. If and only if we are permitted the step from sentences of the form $\neg\forall x\neg Cx$ to $\exists xCx$, for decidable C, can we reach the wanted conclusion. It is well known, however, that such a step is not constructively valid in general.[4]

Similarly, under the same assumptions on T, F and A, the *soundness* of the ω-rule with respect to logical consequence cannot be established without essential use of the constructively invalid

[4] See, e.g., Michael Dummett, *Elements of Intuitionism* (New York: Oxford, 1977), pp. 21/2.

distribution principle $\forall x(Bx \vee C) \supset (\forall xBx \vee C)$. Consider the following application of the ω-rule:

$$\frac{\text{For all } t_i,\ M \vdash N, At_i}{M \vdash N, \forall xAx}$$

Under the assumption of bivalence for M,N and all the At_i, the soundness proof goes over into the inference of

$$M \,\varepsilon\, F \vee N \,\varepsilon\, T \vee \forall xAx \,\varepsilon\, T$$

from

$$\forall t_i(M \,\varepsilon\, F \vee N \,\varepsilon\, T \vee At_i \,\varepsilon\, T)$$

It is a perhaps little-known fact that the addition of the $\forall/\vee$ distribution law to an intuitionistic system with decidable prime formulas will in fact produce the corresponding classical system. [One proves $\forall x(Ax \vee \neg Ax)$ by induction on the build-up of A, where the quantifier step is taken care of by the distribution law.]

So in fact both of the further assumptions on the semantic frameworks demand full classical logic for the proof of their preservation. This seems to me to pose a grave methodological obstacle to Hacking's approach, because *no reason is given for restricting the methods of proof in some of the preservation results, e.g., the elimination theorems, but not in others, e.g., soundness.* It seems not unfair to say that Hacking has not even noticed these differences, and without a detailed philosophical analysis and subsequent motivation this discrepancy appears entirely ad hoc. Of course, there is also the possibility that the whole idea of defining classical logical constanthood in terms of cut-free rules is drastically mistaken; from an intuitionistic viewpoint, on the other hand, such an approach appears much more plausible, as has been shown in recent writings by Dag Prawitz and Jeffrey Zucker.[5]

IV. LINDSTRÖM'S THEOREM AND RAMIFIED SECOND-ORDER LOGIC

Lindström's remarkable result[6] says, roughly, that first-order logic is singled out from a wide class of possible logics by three of its properties: (a) containing first-order logic, (b) satisfying a downward Löwenheim-Skolem theorem for sentences, and (c) being compact.

[5] Prawitz, "Meaning and Proofs," *Theoria,* XLIII, 1 (1977): 2-40; "Proofs and the Meaning and Completeness of the Logical Constants," in Jaakko Hintikka *et al.*, eds., *Essays on Mathematical and Philosophical Logic* (Boston: Reidel, 1978), pp. 25-39. J. I. Zucker and Robert Tragesser, "The Adequacy Problem for Inferential Logic, *Journal of Philosophical Logic,* VII, 4 (November 1978): 502-516.

[6] Per Lindström, "On Extensions of Elementary Logic," *Theoria,* XXXV, 1 (1969): 1-11.

Hacking goes on to argue that "a Löwenheim-Skolem theorem holds for anything which, on my delineation, is logic. It follows from Per Lindström's result that logic is not compact" (307). By "logic" Hacking here means ramified second-order logic. If we inspect the rules for this logic we find that they are, of course, perfectly finitary, just like the standard rules for first-order logic. The "equivalent ω-rule" will in this case look like:

If, for all terms T_i^k of level k, $\Gamma \vdash \Delta, A(T_i^k)$, then $\Gamma \vdash \Delta, \forall X^k A(X^k)$.

which rule is "equivalent," exactly as in the first-order case, to its *eigen*-parameter version:

$$\text{If } \Gamma \vdash \Delta, A(P^k), \text{ then } \Gamma \vdash \Delta, \forall X^k (X^k).$$

where P^k does not occur in Γ, Δ. But then, obviously, this *is* a compact logic, contrary to what was claimed by Hacking, and all the premises of Lindström's theorem are satisfied. The conclusion, however, is false.

Two possibilities suggest themselves: (1) Lindström's proof is radically wrong, or (2) Hacking's ramified logic falls outside the range of the possible logics in Lindström's theorem. The latter is, of course, the case: the "extensional semantics" of Hacking[7] has the property that all the relevant information about a structure for ramified second-order logic is not determined simply by specifying the universe of individuals and the predicates and functions thereon. But Lindström presupposes this; that is why *full* second-order logic—not the compact, Henkin-complete variety—is a possible Lindström logic: as soon as the domain of individuals has been given, the set-universe is uniquely determined as the class of all subsets of the domain.[8] Both the Henkin-complete and Hacking's ramified logic may also be regarded as *many-sorted first-order logics,* and then they do obey both the premise and the conclusion of the Lindström theorem; when regarded as second-order logics they are not possible Lindström logics.

V. HACKING'S ADEQUACY CRITERIA

Hacking's use of the notion of "adequacy criteria" is very loaded (304). In order that his demarcation should fit in with the "analytical program" he introduces three "adequacy criteria" (A)-(C) and goes on to argue that, because his particular delineation happens to satisfy

[7] In the paper referred to in his footnote 18, Hacking claims that his semantics for ramified logic "points in the same direction" as that of Saul Kripke, "Is There a Problem about Substitutional Quantification?", in Evans and McDowell, *op. cit.*, p. 368.

[8] Monk, *Mathematical Logic* (New York: Springer, 1976), ch. 30, is a good reference for the technicalities involved. Monk also has a nice exposition of Lindström's theorem.

(A)-(C), "rules for pure logic [should] have the subformula property and be conservative".

At best one would call these conditions "*desiderata*"; their role is to rule out unacceptable delineation, but they do not guarantee that any delineation that happens to satisfy them is *the* true delineation. Tarski's Convention T, on the other hand, is a proper adequacy criterion in the sense that *any* definition of truth that satisfies Convention T will do.

There are also reasons for doubting that (A)-(C) are even desiderata; we saw above that second-order ramified logic was nothing but a many-sorted first-order theory if we wanted it to be a Lindström logic, or, if we preferred to view it as a second-order theory, it was not a Lindström logic. There is also the voluminous literature on various alternative logics; so why should any delineation be held to be automatically adequate just because it happens to give ramified logic as its result?

In conclusion, Hacking's paper may be summed up by a working proof-theorist as: Consider a logic for which cut- and other elimination results are finitistically provable. As the rules are supposed to have the subformula property, the well-known Beth-Hintikka-Kanger-Schütte technique[9] for proving the completeness of cut-free rules may be applied backwards to read off a semantics.

It is hard to see how such a technical groundwork can be made to hold the grand superstructure Hacking wishes to erect on its basis.

GÖRAN SUNDHOLM

Magdalen College, Oxford University

HACKING ON LOGIC: TWO COMMENTS*

IAN HACKING's "What Is Logic?"[1] is manifestly an important contribution to one kind of theory about the classical conception of logical consequence and logical constants. But his theory, to be adequate, must be regarded as making two tacit assumptions. The point of this note is to make explicit these assumptions. Once these assumptions are recognized, Hacking's theory can be seen to have important elements in common with other less proof-theoretical accounts of classical logic: Hacking's theory may be regarded as an

[9] Hacking's paper has 33 footnotes, but not a single reference to this method. A lucid presentation is in Raymond Smullyan, *First-order Logic* (New York: Springer, 1968).

* I am grateful to Göran Sundholm and Charles Parsons for helpful comments.

[1] This JOURNAL, LXXVI, 6 (June 1979): 285-319.

0022-362X/81/7803/0168$00.70