Notre Dame Journal of Formal Logic
Volume 44, Number 3, 2003

A Remark on Henkin Sentences and Their Contraries

John P. Burgess

Abstract That the result of flipping quantifiers and negating what comes after, applied to branching-quantifier sentences, is not equivalent to the negation of the original has been known for as long as such sentences have been studied. It is here pointed out that this syntactic operation fails in the strongest possible sense to correspond to any operation on classes of models.

Any first-order sentence is equivalent first to some *prenex* sentence, consisting of a *prefix* or string of quantifiers followed by a *matrix* or quantifier-free formula, and any prenex sentence is equivalent to an existential second-order sentence. For example, (1) below is equivalent to (2) below:

$$\forall u \exists x \forall v \exists y R(u, v, x, y) \quad (1) \qquad \exists f \exists g \forall u \forall v R(u, f(u), v, g(u, v)) \quad (2)$$

Logicians have for almost a half century also considered *Henkin* sentences, consisting of one *or more* prefixes (customarily written in a vertical column, though they could be written on one line, say using a slash to represent line breaks, as is done when quoting verse) followed by a single matrix. The intended sense of such a sentence is explained as that of an associated existential second-order sentence. Thus (3) below is defined to be equivalent to (4) below:

$$\begin{matrix} \forall u \exists x \\ \forall v \exists y \end{matrix} R(u, v, x, y) \quad (3) \qquad \exists f \exists g \forall u \forall v R(u, f(u), v, g(v)) \quad (4)$$

The difference between (2) and (4) is that g is a two-place function in the former and a one-place function in the latter. Where (1) says that for every u there exists an x and for every v there exists a y such that $R(u, v, x, y)$, (3) says the same with the addition that y *depends only on* v.

Received July 16, 2003; accepted September 8, 2003; printed July 7, 2004
2000 Mathematics Subject Classification: Primary, 03C80
Keywords: branching quantifiers, Henkin sentences

According to a theorem of Enderton [2] and Walkoe [4] (to which the reader is referred for the details of the general definition of branching quantifiers) *every* existential second-order sentence is equivalent to a Henkin sentence. Thus all the theorems, long familiar to logicians, about models of existential second-order sentences, immediately become applicable to Henkin sentences. So, for instance,

1. The compactness theorem holds (a set of Henkin sentences has a model if every finite subset does);
2. There is a Henkin sentence whose models are precisely the interpretations with an infinite domain; and
3. It is not the case that the negation of any Henkin sentence is equivalent to a Henkin sentence.

(It is worth mentioning that since (1) is incompatible with the existence of a sentence whose models are precisely the interpretations with a *finite* domain, (1) and (2) together *imply* (3).)

By contrast to this last, the conjunction (or disjunction) of any two Henkin sentences Φ and Ψ *is* equivalent to a Henkin sentence $\wedge\Phi\Psi$ (respectively, $\vee\Phi\Psi$). Relettering if necessary, we may assume that Φ and Ψ have no variables in common, and then we may take as $\wedge\Phi\Psi$ (respectively, $\vee\Phi\Psi$) the Henkin sentence whose prefixes are those of Φ together with those of Ψ, and whose matrix is the conjunction (respectively, disjunction) of that of Φ with that of Ψ.

To avoid trivialities, in the logic of first-order sentences it is conventional to exclude models with an empty domain, while in the logic of Henkin sentences it will be convenient to exclude models with a one-element domain as well. We write $[\Phi]$ for the class of models of Φ (tacitly excluding those with one-element domains). The syntactic operation $\wedge$ (respectively, $\vee$) on Henkin sentences corresponds to the semantic operation of intersection (respectively, union) on classes of models. As a result, all the usual Boolean laws hold for $\wedge$ and $\vee$. For instance, $\wedge\Phi\Psi$ is equivalent to $\wedge\Psi\Phi$, since $[\wedge\Phi\Psi] = [\Phi] \cap [\Psi] = [\Psi] \cap [\Phi] = [\wedge\Psi\Phi]$.

If Φ has but a single prefix, or in other words, is a first-order sentence in prenex form, then a prenex sentence $\neg\Phi$ equivalent to the negation $\sim\Phi$ of Φ can be obtained by *flipping* the prefix of Φ, replacing each $\forall$ by $\exists$ and vice versa, and negating the matrix of Φ. We may also form for any Henkin sentence Φ a Henkin sentence $\neg\Phi$ whose prefixes are the results of flipping those of Φ, and whose matrix is the negation of that of Φ. This $\neg\Phi$ is sometimes called the *contrary* of Φ, while the ordinary negation $\sim\Phi$, which is not in general equivalent to a Henkin sentence, would be called the *contradictory* of Φ. For example, the contrary of (3) above, equivalent to (4) above, would be (5) below, equivalent to (6) below.

$$\begin{matrix}\exists u\forall x\\ \exists v\forall y\end{matrix} \sim R(u, v, x, y) \quad (5) \qquad\qquad \exists i\exists j\forall x\forall y \sim R(i, x, j, y) \quad (6)$$

$$\begin{matrix}\forall x\\ \exists y\end{matrix}\ (x = y) \quad (7) \qquad\qquad \begin{matrix}\exists x\\ \forall y\end{matrix}\ (x \neq y) \quad (8)$$

Similarly if we call the sentence in (7) above Θ_0, then the sentence in (8) above is $\neg\Theta_0$. This example (a replacement, called to my attention by Väänänen, for an unnecessarily complicated one used in an earlier draft of this note) is the simplest case of the failure of equivalence of contrary and contradictory. For $\neg\Theta_0$ is true in no interpretations at all, and Θ_0 only in interpretations with single-element domains, which we are excluding; so $[\Theta_0] = [\neg\Theta_0] = \varnothing$.

Some but not all the usual Boolean laws hold for $\neg$. Among those that do hold are the De Morgan laws: $\neg \wedge \Phi\Psi$ is equivalent to $\vee\neg\Phi\neg\Psi$ (they have the same prefixes, and matrices that are equivalent by the De Morgan laws of sentential logic), and $\neg \vee \Phi\Psi$ to $\wedge\neg\Phi\neg\Psi$. Another that holds is the law of noncontradiction: Φ and $\neg\Phi$ are *incompatible*, in the sense of having no common model. For instance, (3) and (5), or equivalently (4) and (6), can have no common model, since otherwise we would have both $R(i, f(i), j, g(j))$ and $\sim R(i, f(i), j, g(j))$. A law that does *not* hold is excluded middle. An interpretation may fail to be a model of either Φ or $\neg\Phi$.

The syntactic operation $\neg$ on sentences not only does not correspond to the semantic operation of complementation on classes of models, but further it does not correspond to any semantic operation § at all. Setting $\S[\Phi] = [\neg\Phi]$ does not give a well-defined operation § on classes of models, since one may have Henkin sentences Φ and Φ' with $[\Phi] = [\Phi']$ but $[\neg\Phi] \neq [\neg\Phi']$. Simply knowing only the class of models of a Henkin sentence, and not the sentence itself, leaves the class of models of its contrary completely undetermined, apart from the fact that the latter class is also the class of models of some Henkin sentence, and the fact that the two classes are disjoint. Such is the content of the following corollary of the Craig interpolation theorem.

Corollary 1 *Let Φ_0 and Φ_1 be incompatible Henkin sentences. Then there is a Henkin sentence Θ such that Θ is equivalent to Φ_0 and its contrary $\neg\Theta$ is equivalent to Φ_1.*

Proof Recall that there is a Henkin sentence Θ_0 with $[\Theta_0] = \varnothing$ and $[\neg\Theta_0] = \varnothing$. Given Φ_0 and Φ_1, let Ψ_0 and Ψ_1 be $\vee\Phi_0\Theta_0$ and $\vee\Phi_1\Theta_0$, respectively. Then we have

$$[\Psi_0] = [\vee\Phi_0\Theta_0] = [\Phi_0] \cup [\Theta_0] = [\Phi_0] \cup \varnothing = [\Phi_0]$$

$$[\neg\Psi_0] = [\neg \vee \Phi_0\Theta_0] = [\wedge\neg\Phi_0\neg\Theta_0] = [\neg\Phi_0] \cap [\neg\Theta_0] = [\neg\Phi_0] \cap \varnothing = \varnothing$$

and similarly $[\Psi_1] = [\Phi_1]$ while $[\neg\Psi_1] = \varnothing$. Since each of $[\Psi_0] = [\Phi_0]$ and $[\Psi_1] = [\Phi_1]$ is the class of models of a existential second-order sentence, and since the two classes are disjoint, by Craig's theorem there is a first-order sentence Ψ such that $[\Psi_0] \subseteq [\Psi]$ and $[\Psi_1] \subseteq [\sim\Psi]$. This Ψ may be taken to be in prenex form, so that $\neg\Psi$ is defined and $[\neg\Psi] = [\sim\Psi]$. Then we may take

$$\Theta = \wedge\Psi_0(\vee\neg\Psi_1\Psi) = \wedge(\vee\Phi_0\Theta_0)(\vee\neg(\vee\Phi_1\Theta_0)\Psi)$$

We will have the following, to complete the proof:

$$[\Theta] = [\wedge\Psi_0(\vee\neg\Psi_1\Psi)] = [\Psi_0] \cap ([\neg\Psi_1] \cup [\Psi]) = [\Psi_0] \cap [\Psi] = [\Psi_0] = [\Phi_0]$$

$$[\neg\Theta] = [\vee\neg\Psi_0(\wedge\Psi_1\neg\Psi)] = [\neg\Psi_0] \cup ([\Psi_1] \cap [\neg\Psi]) = [\Psi_1] - [\Psi] = [\Psi_1] = [\Phi_1]$$

□

In recent years Hintikka [3] and co-workers have revived a variant version of the logic of Henkin sentences under the label "independence-friendly" logic, have restated many theorems about existential second-order sentences for this "new" logic, and have made very large claims about the philosophical importance of the theorems thus restated. In discussion, pro and con, of such philosophical claims it has not been sufficiently emphasized that contrariety, the only kind of "negation" available, fails to correspond to any operation on classes of models. For this reason it seemed worthwhile to set down, in the form of the corollary above, a clear statement of just how total the failure is. But even apart from recent controversies it is hoped that the corollary is of some interest as an exercise in applying the interpolation theorem. One immediate implication of the corollary just proved is worth mentioning. An *elementary class* or EC is a class of the form $[\Phi]$ with Φ a first-order sentence, and a *pseudo-elementary* class or PC is one of the form $[\Phi]$ with Φ an existential second-order sentence. The Enderton-Walkoe theorem says that for any PC, call it K, there is a Henkin sentence Θ such that $K = [\Theta]$. The corollary just proved allows this theorem to be strengthened to say that for any two disjoint PCs, call them K_0 and K_1, there is a Henkin sentence Θ such that $K_0 = [\Theta]$ and $K_1 = [\neg\Theta]$. This observation answers (modulo the exclusion of interpretations with a one-element domain) a question of Caicedo and Krynicki [1] about "independence-friendly" logic.

References

[1] Caicedo, X., and M. Krynicki, "Quantifiers for reasoning with imperfect information and Σ^1_1-logic," pp. 17–31 in *Advances in Contemporary Logic and Computer Science (Salvador, 1996)*, edited by W. A. Carnielli, vol. 235 of *Contemporary Mathematics*, American Mathematical Society, Providence, 1999. Zbl 0935.03048. MR 2001b:03037. 188

[2] Enderton, H. B., "Finite partially-ordered quantifiers," *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, vol. 16 (1970), pp. 393–97. Zbl 0193.29405. MR 44:1546. 186

[3] Hintikka, J., *The Principles of Mathematics Revisited*, Cambridge University Press, Cambridge, 1996. Zbl 0869.03003. MR 97j:03005. 187

[4] Walkoe, W. J., Jr., "Finite partially-ordered quantification," *The Journal of Symbolic Logic*, vol. 35 (1970), pp. 535–55. Zbl 0219.02008. MR 43:4646. 186

Department of Philosophy
Princeton University
Princeton NJ 08544-1006
jburgess@princeton.edu