Facta Philosophica 4, 2002: 3-9

Can Variables Be Explained Away?

Peter van Inwagen

In his justly famed essay "Variables Explained Away"[1], W.V. Quine has shown how to translate any sentence in the language of first-order logic (that language whose vocabulary consists of quantifiers, variables[2], predicate-letters, sentential connectives, and punctuation marks) into a sentence of a language whose vocabulary consists solely of predicate-letters, punctuation marks, and the following six "predicate operators": Derelativization (Der), Major inversion (Inv), Minor inversion (inv), Reflection (Ref), Negation (Neg), and Cartesian multiplication (... × ...). (These expressions are "predicate operators" in this sense: the result of prefixing any of the first five of them to a predicate-letter is an expression having the syntax of a predicate-letter; the result of surrounding '×' with two predicate-letters is an expression that has the syntax of a predicate-letter.)

The germ of Quine's technique for the elimination of variables can be displayed by showing how to eliminate variables from a very simple language. Consider a fragment of the language of first-order logic that consists only of existential quantifications on a one-place predicate-letter; that is, of sentences like these:

$\exists x\, Fx \qquad \exists x\, Gx \qquad \exists z\, Hz.$

It is easy to eliminate variables from this fragment. The only essential way in which these formulae differ from one another is that each contains a different predicate-letter. We could therefore replace each of them with a formula that contained that one essential ingredient and some symbol that expressed the idea of existential generalization. That is, in fact, just what Quine's operator 'Der' does when it is applied to

1 *Selected Logic Papers* (New York: Random House, 1966), pp. 227–235. (Originally published in 1960 in the *Proceedings* of the American Philosophical Society.)

2 By 'variables', I understand first-order variables: variables that occupy nominal—and not sentential or predicate—positions.

a one-place predicate-letter. Using it to this effect, we replace the above three formulae by:

Der F   Der G   Der H,

If 'Der' were to be applied only to one-place predicate-letters, we could define it as follows: If F is a one-place predicate-letter, the result of writing 'Der' followed by F represents indifferently the existential quantification on F with respect to any variable. (But 'Der' in fact applies to predicate-letters and predicate-letter-like expressions of any number of places. If F is a six-place predicate-letter, the result of prefixing 'Der' to F is an expression that has the syntax of a five-place predicate letter. For the purposes of this note, it will not be necessary to explain how 'Der' works in the general case.) So much for our example and our simple language. Let us now return to the language of first-order logic. In "Variables Explained Away," Quine shows how—by repeated application of his six predicate operators—to represent any closed, truth-functionally simple formula of first-order logic as either an existential quantification on a single variable or the negation of such. Consider, for example, the following formula:

$\forall x$(if Fx, then, if Gxx, $\exists y$(Hyx)).

Quine's technique enables one to translate this formula into a formula of the following form:

$\sim\exists x$ [very complicated expression having the syntax of a predicate-letter and formed by the iterated application of the predicate-operators to 'F', 'G', and 'H'] x.

And this can then be replaced by

Neg Der [very complicated expression ...].

('Neg' is in all essentials the negation sign—but like the other predicate-operators—it applies to expressions having the syntax of a predicate-letter, not to expressions having the syntax of a sentence-letter. The tilde takes a sentence-letter, or something that—like 'p v q' or 'Fxy'—can replace a sentence-letter, and makes a new expression that can replace a sentence-letter. 'Neg' takes a predicate-letter, or something that can replace a predicate-letter, and makes a new expression that can replace a predicate-letter. In the formula '(Neg P)xxy', the expression 'Neg P' has the syntax of a predicate-letter; it is a predicate-letter-like expression such that '(Neg P)xxy' is equivalent to '~ (Pxxy)'.)

One possible translation of '∀*x*(if F*x*, then, if G*xx*, ∃*y*(H*yx*))' into Quine's predicate-operator language is

Neg Der Ref [F × Ref (Ref G × Neg Der Inv H)].

(If this is our translation, the "very complicated expression" is 'Ref [F × Ref (Ref G × Neg Der Inv H)]'.) As Quine puts it, the essence of his technique is to "coax variables ... into positions where we can dispense with them" (229).[3] I shall not give an exposition of the details of this technique. In matters of logical technique, Quine is his own best expositor, and "Variables Explained Away" is short and easily available. Let us assume (what is true) that the technique works, and turn to this question: What philosophical interest is there in the fact that variables can be eliminated using this technique? The following two quotations contain Quine's answer to this question:

> Nor are variables necessarily tied up with generality prefixes or existence prefixes at all. Basically, the variable is best seen as an abstractive pronoun: a device for marking positions in a sentence with a view to abstracting the rest of the sentence as a predicate (228).

> The interest in carrying out the elimination is that the device of the variable thereby receives, in a sense, its full and explicit analysis (229).

My purpose in this note is to raise the question whether Quine's way of eliminating variables does indeed provide the device of the variable with (in any sense) its full and explicit analysis. I doubt whether it does. I will try to explain my doubts.

Consider the formula

∀*x* ∃*y* G*xy*.

3 Quine's technique would, of course, be of no interest if it were not "reversible." But it is designed to be reversible and is reversible. We can, roughly speaking, apply the technique "in reverse" to 'Neg Der Ref [F × Ref (Ref G × Neg Der Inv H)]' and extract thereby from this predicate-operator formula a formula of first-order logic provably equivalent to '∀*x*(if F*x*, then, if G*xx*, ∃*y*(H*yx*))'. Since the predicate-operator formula contains no variables, applying the technique "in reverse" will require us to introduce variables at various points in our application of it, and, subject to obvious constraints, what variables we introduce and the order in which we introduce them will be matters of arbitrary choice. For that reason, the first-order formula we extract from 'Neg Der Ref [F × Ref (Ref G × Neg Der Inv H)]' by applying the "reverse technique"will not necessarily be the original formula; it might well be some "alphabetic variant" of the original, such as '∀*z*(if F*z*, then, if G*zz*, ∃*x*(H*xz*))'. It is because Quine's technique for eliminating variables is in this sense reversible, that the predicate-operator formulas it yields can be said to be equivalent to their first-order originals.

If we apply Quine's technique to this formula we obtain:

Neg Der Neg Der G.

Well and good. The variables are "gone," and the second expression is equivalent to the first. But variables do not occur only in expressions like '$\forall x\ \exists y\ Gxy$', expressions in the language of first-order logic. The formula '$\forall x\ \exists y\ Gxy$', like all expressions containing predicate-letters, is an abstraction, a device for displaying the syntactic features common to an infinite number of sentences. Two of the infinitely many sentences whose common syntactic features are displayed by '$\forall x\ \exists y\ Gxy$' are these:

$\forall x\ \exists y$ *x* is less than *y*

$\forall x\ \exists y$ *y* is the square of *x*.

It is in sentences of *this* sort, sentences that contain not predicate-letters but fragments of natural language like 'is less than' and 'is the square of' that variables live and have their being. Can Quine's technique be used to eliminate variables from sentences like these? The short answer is No, for this technique can be applied only to expressions in which (apart from their occurrences in quantifier-phrases) variables occur only in strings or unbroken clumps—as in, for example, '$Pxxzyx$'—, and this is not how the variables in '$\forall x\ \exists y$ *x* is less than *y*' and '$\forall x\ \exists y$ *y* is the square of *x*' are arranged. But it may be objected, and rightly, that this fact about the expressions to which Quine's technique may be applied constitutes a mere technical difficulty, one easily surmounted. We may surmount it by introducing the idea of a "segregated open sentence."

Let us say that a *predicate* of a given natural language is any expression formed from a declarative sentence of that language as follows: some or all the occurrences of names or terms in that sentence are to be replaced by occurrences of the first *n* numerals, starting with '**1**'. (We write the numerals in bold-face.) Thus, a predicate may contain '**1**' and no other numerals; or it may contain '**1**' and '**2**' and no other numerals; or it may contain '**1**' and '**2**' and '**3**' and no other numerals—and so on. For example, since 'Miami is north of Boston' is a sentence of English, the following expressions are predicates of English:

1 is north of Boston

Miami is north of **1**

1 is north of **1**

1 is north of 2

2 is north of 1.

The highest numeral that occurs in a predicate indicates its *number of places*. Thus, the first three of the five predicates I have displayed are one-place predicates, the other two are two-place predicates, and '3 has married 2 more times than 1 has married 2' is a three-place predicate. (Declarative sentences of English are not, by the terms of our definition, predicates, but, if we wished, we could revise our definition and allow them to count as zero-place predicates.) A *segregated open sentence* is an expression that consists of an *n*-place predicate followed by *n* occurrences of variables. The following expressions are thus segregated open sentences:

1 is north of Boston *z*

1 is north of 2 *xy*

1 is north of 2 *zz*

1 is north of 1 *y*

2 is north of 1 *zy*.

In the first of these expressions, '*z*' is the "first variable"; in the second expression, '*x*' is the "first variable" and '*y*' the "second variable"; in the third expression, '*z*' is both the first and the second variable; in the fourth expression, '*y*' is the first variable; in the final expression, '*z*' is the first variable and '*y*' the second.

Now what do segregated open sentences *mean*? We may give a meaning to segregated open sentences by describing a mechanical procedure for converting any segregated open sentence into an "ordinary" open sentence; we simply declare that a segregated open sentence is equivalent in meaning to the ordinary open sentence the procedure yields. Here is the procedure: replace all occurrences of '1' in the predicate of the segregated open sentence by (occurrences of) its first variable, all occurrences of '2' by its second variable, and so on. Thus, the above segregated open sentences are equivalent, respectively, to

z is north of Boston

x is north of *y*

z is north of *z*

y is north of *y*

y is north of *z*.[4]

Ordinary open sentences may be turned into equivalent segregated open sentences by reversing this procedure (one must of course "invent" appropriate predicates in order to do this)—but there will generally be more than one segregated open sentence that is a correct translation of a given ordinary open sentence. For example, '*x* is north of *y*' is equivalent both to '**1** is north of **2** *xy*' and to '**2** is north of **1** *yx*'.

I said above that segregated open sentences could be used to enable us to apply Quine's technique for the elimination of variables to sentences in which variables occurred in expressions like '*x* is less than *y*' and '*y* is the square of *x*'. Here is the way to do this.

Take any sentence in the quantifier-variable idiom and replace each "inmost" ordinary open sentence (each open sentence that contains no quantifiers) it contains with an equivalent segregated open sentence. In the resulting expression, variables (apart from their occurrence in quantifier-phrases) will occur only in "strings" or "clumps," just as in the language of first-order logic. Quine's technique for the elimination of variables may then be applied to the result, provided we apply his six operators to *predicates* in just the way he applies them to predicate-*letters*. Consider, for example '∀*x* ∃*y* *x* is less than *y*'. We may replace this sentence either with the equivalent '∀*x* ∃*y* **1** is less than **2** *xy*' or the equivalent '∀*x* ∃*y* **2** is less than **1** *yx*'. Application of Quine's technique to the former yields:

Neg Der Neg Der **1** is less than **2**.

4 A moment ago, I said, "The formula '∀*x* ∃*y* *Gxy*', like all expressions containing predicate-letters, is an abstraction, a device for displaying the syntactic features common to an infinite number of sentences. Two of the infinitely many sentences whose common syntactic features are displayed by '∀*x* ∃*y* *Gxy*' are these:

∀*x* ∃*y* *x* is less than *y*

∀*x* ∃*y* *y* is the square of *x*."

This statement may be given a precise sense in terms of the devices I have introduced: Replacing the two-place predicate-letter '*G*' in the formula '∀*x* ∃*y* *Gxy*' with the two-place predicate '**1** is less than **2**', yields '∀*x* ∃*y* **1** is less than **2** *xy*', which is equivalent to '∀*x* ∃*y* *x* is less than *y*' (and similarly for '∀*x* ∃*y* *Gxy*', '**2** is the square of **1**', and '∀*x* ∃*y* *y* is the square of *x*'). '∀*x* ∃*y* *Gxy*' displays the syntactic features common to all and only those sentences that can be obtained from it by substituting a two-place predicate for the two-place predicate-letter '*G*'.

To the latter:

> Neg Der Neg Der Inv 2 is less than 1.

Quine's technique, suitably elaborated, can therefore be used to eliminate variables from expressions that contain open sentences like '*x* is less than *y*', open sentences in which variables occur in the "scattered" nominal positions dictated by the grammatical accidents of natural language. I will assume—and, I concede, the force of my argument does depend on this assumption—that this can be done only by devices essentially equivalent to the ones I have employed.

Now, having solved the technical problem that stood in the way of applying Quine's technique for the elimination of variables to expressions containing open sentences of natural language, let us return to our question: *Has* the device of the variable received—by the application of this technique—"its full and explicit analysis"? I do not think so. The reasons for my doubt can be put in the form of a further question: Can we really understand the structured linguistic items I have called predicates—'1 is north of 2', '2 is less than 1', '1 has married 2 more times than 3 has married 2'—without a prior understanding of variables? I myself, when I consider this question, find that I seem to be able to understand the contrast between the structured items '1 is less than 2' and '2 is less than 1' only by considering the role they play in what I have called segregated open sentences, and I can understand this role only because I understand the rules for translating segregated open sentences into ordinary open sentences (and the rules for translating ordinary open sentences into segregated open sentences). I cannot, therefore, understand predicates unless I have a prior understanding of ordinary open sentences. And that understanding, of course, involves a prior understanding of variables.

Peter van Inwagen
The University of Notre Dame
Department of Philosophy
Notre Dame, Indiana 46556-4619
peter.vaninwagen.1@nd.edu