Binding bound variables*

Brian Rabern

Consider this formula of first-order logic:

(1) $\forall x \exists x Fx$

This is a paradigm example of what is called "vacuous quantification". The occurrence of the variable 'x' in 'Fx' is already bound by '$\exists x$' in the subformula '$\exists x Fx$', thus prefixing '$\forall x$' is idle—the universal quantifier is *vacuous*. For example, the following is a theorem concerning vacuous quantification from introductory logic texts (see, e.g., Kalish and Montague 1964: 164-65).

(2) $\forall x \exists x Fx \leftrightarrow \exists x Fx$

In general, one might insist on the following principle concerning binding and vacuity.

> THE PRINCIPLE OF VACUOUS QUANTIFICATION. If all the variables in a formula ϕ are bound, then for any quantifier Σ, $\Sigma\phi \leftrightarrow \phi$.

In slogan: *You can't bind a bound variable!* Or can you? There might be more going on with so-called "vacuous quantification" than is commonly recognised. In fact, I will argue that the principle is false.

To do so I will introduce a simple language that sticks close to the syntax and semantics of first-order logic—this helps to demonstrate that there is nothing tricky going on in my counterexample.

First we define the syntax. In addition to parenthesis, the basic symbols consist of the following:

Variables: $x, y, z, \ldots$
Predicates: $F, G, H, \ldots$
Connectives: $\neg$, $\wedge$
Quantifiers: $\exists$, ᴚ

For any sequence of variables $\alpha_1, \ldots, \alpha_n$, any n-place predicate π, and any variable α, the sentences of the language are provided by the following grammar:

$\phi ::= \pi\alpha_1 \ldots \alpha_n \mid \neg\phi \mid (\phi \wedge \phi) \mid \exists\alpha\phi \mid$ ᴚ$\alpha\phi$

This language looks essentially like predicate logic, and we can define the other usual operators (e.g. '$\vee$', '$\rightarrow$', '$\forall$', etc.) as abbreviations in terms of our basic symbols. The only novel thing about the language, thus far, is the addition of a new quantifier symbol 'ᴚ'. There is nothing interesting about it syntactically, and although it will be given an interpretation that is different from '$\exists$' it is essentially a kind of existential quantifier.

*For helpful comments and feedback many thanks to Wolfgang Schwarz, Landon Rabern, Dilip Ninan, Martin Smith, Bryan Pickel, and audiences at the St. Andrews Philosophy Club and the Edinburgh Language Workshop.

Turning to the semantics, let a model $\mathfrak{A} = \{D, I\}$, where D is a (non-empty) set of individuals and I is an interpretation function, which assigns values to the predicates. Since our language has variable-binding operators we relativise to an *assignment*, which assigns values to the variables. An assignment g is a function from the set of variables to set of individuals D. We provide the following recursive semantics, in the style of Tarski, by recursively defining 1 (truth or satisfaction) relative to an assignment g:[1]

- $[\![\alpha]\!]^g = g(\alpha)$
- $[\![\pi\alpha_1 \ldots \alpha_n]\!]^g = 1 \quad \text{iff} \quad \langle [\![\alpha_1]\!]^g, \ldots, [\![\alpha_n]\!]^g \rangle \in I(\pi)$
- $[\![\neg\phi]\!]^g = 1 \quad \text{iff} \quad [\![\phi]\!]^g = 0$
- $[\![\phi \wedge \psi]\!]^g = 1 \quad \text{iff} \quad [\![\phi]\!]^g = 1 \text{ and } [\![\psi]\!]^g = 1$
- $[\![\exists x\phi]\!]^g = 1$ iff for some g' that differs from g at most in that $g'(x) \neq g(x)$, $[\![\phi]\!]^{g'} = 1$
- $[\![\text{ꓤ} x\phi]\!]^g = 1$ iff for some g' that differs from g at most in that $g'(x) > g(x)$, $[\![\phi]\!]^{g'} = 1$

The last clause deserves comment, since it appeals to the *greater than* relation. For this to make sense, of course, the individuals in D have to be ordered: we could impose an ordering on any domain, but let's instead just assume that D is the set of natural numbers with their natural ordering. It is not essential to my argument that we use this particular relation, nor that we order the domian. But a nice relation like this helps to keep the initial set up simple, and then one can generalise after seeing the key point. Notice that 'ꓤ' is very similar to '∃' in that it is an existential quantifier, but it only "looks" at a subset of the assignments that '∃' looks at. '∃' is the "for some" quantifier, while 'ꓤ' is the "for some greater" quantifier.

Consider the following sentence of our language:

(3) $\text{ꓤ} xFx$

Certainly, in (3) the variable in the embedded formula 'Fx' is bound by the quantifier '$\text{ꓤ} x$'.

CLAIM 1. All the variables in '$\text{ꓤ} xFx$' are bound.

This claim seems innocent enough, but some may suspect some kind of trickery: some sleight of hand or some variable up my sleeve. One might worry that I pulling a fast one with respect to the operative notion of "binding"—are all the variables in '$\text{ꓤ} xFx$' really bound? Standardly, an occurrence of a variable α is said to be *bound* in a formula just in case it is immediately attached to the quantifier or within the scope of a quantifier that is indexed with α (and free otherwise). So in a formula such as '$(\exists xFx \wedge Gy)$' both occurrences of 'x' are bound, while the occurrence of 'y' is free. Clearly, given this standard definition, all the occurrences of variables in '$\text{ꓤ} xFx$' are bound.

Thus, if I can establish that prefixing a quantifier, such as '$\exists x$', to (3) is not idle, then we will have a counterexample to the Principle of Vacuous Quantification. More precisely, I will show the following (for some model $\mathfrak{A}$):

CLAIM 2. $[\![\exists x \text{ꓤ} xFx]\!] \neq [\![\text{ꓤ} xFx]\!]$

[1]Here we will not worry about the distinction between "satisfaction by a sequence" and "truth"—of course, Tarski reserves *truth* for formulae that are satisfied by all sequences.

We are assuming that D is the set of natural numbers, and let's also assume a particular interpretation for 'F', namely $I(F) = \{8\}$. So 'F' is only true of 8. Provided this model, consider the truth conditions of (3) relative to some assignment g, which are calculated as follows:

$$\begin{aligned}
[\![\text{ꓤ}xFx]\!]^g = 1 \ \ &\text{iff for some } g' \text{ that differs from } g \text{ at most in that } g'(x) > g(x), [\![Fx]\!]^{g'} = 1 \\
&\text{iff for some } g' \text{ that differs from } g \text{ at most in that } g'(x) > g(x), g'(x) \in I(F) \\
&\text{iff for some } n \geq g(x), n = 8.
\end{aligned}$$

Assume $g(x) = 10$, then since, of course, there isn't an $n \geq 10$ such that $n = 8$, it follows that $[\![\text{ꓤ}xFx]\!]^g = 0$. Now let's drop the hammer: embed '$\text{ꓤ}xFx$' under '$\exists x$':

(4) $\exists x \text{ꓤ} xFx$

By calculating the truth conditions (relative to the same model) we see that the outer quantifier is not vacuous:

$$\begin{aligned}
[\![\exists x\text{ꓤ}xFx]\!]^g = 1 \ \ &\text{iff for some } g' \text{ that differs from } g \text{ at most in that } g'(x) \neq g(x), [\![\text{ꓤ}xFx]\!]^{g'} = 1 \\
&\text{iff for some } g' \text{ that differs from } g \text{ at most in that } g'(x) \neq g(x), \\
&\quad \text{for some } g'' \text{ that differs from } g' \text{ at most in that } g''(x) > g'(x), g''(x) \in I(F) \\
&\text{iff for some } m, \text{ for some } n \geq m, n = 8.
\end{aligned}$$

Since there is an m and an n such that $n \geq m$ and $n = 8$, it follows that $[\![\exists x\text{ꓤ}xFx]\!]^g = 1$. Thus, relative to this model and assignment g, $[\![\text{ꓤ}xFx]\!]^g = 0$, while $[\![\exists x\text{ꓤ}xFx]\!]^g = 1$. This completes the proof that $[\![\exists x\text{ꓤ}xFx]\!] \neq [\![\text{ꓤ}xFx]\!]$, and thus that following biconditional is not valid (it is false at g):

(5) $\exists x\text{ꓤ}xFx \leftrightarrow \text{ꓤ}xFx$

This provides a simple counterexample to the Principle of Vacuous Quantification. Thus, even when all the variables in a formula ϕ are bound, prefixing a quantifier Σ to ϕ can be non-vacuous.[2]

Let me conclude with two comments.

Comment A. If we restrict our focus to standard first-order logic, then the Principle of Vacuous Binding clearly holds, so what is the essential difference introduced by the 'ꓤ' quantifier? First-order quantifiers (e.g. '$\exists$' and '$\forall$') are standardly appointed a special kind of stability—one that

[2] But what about the claim that bound variables can be re-bound? Does prefixing '$\exists x$' to '$\text{ꓤ}xFx$' result in 'x' being rebound? I think so, but I'm not so interested in explicitly defending it here, since this turns on some subtle terminological issues. We'd need a definition of when a quantifier *binds* a particular occurrence of a variable in a formula. There is a standard definition of this in terms of syntax which stipulates that a bound variable can't be re-bound (see e.g. Heim and Kratzer 1998: 120). So, even though a natural way to describe the counterexample above would be to say that '$\exists x$' rebinds the last occurrence of 'x', this is ruled out by a standard definition. What's going on? Under the stress of the counterexample the syntactic definition of binding is pulling apart from the background semantic understanding of binding. The class of "variables" and "quantifiers" (or variables-binders in general) are grouped together due to their interesting *semantic* properties, not their syntactic properties. Thus, in this more fundamental semantic sense we should say that a quantifier binds an occurrence of a variable in a formula when the sensitivity of the variable is affected by the shifting induced by the quantifier. It is in the semantic sense that you can bind a bound variable.

is not essential to their status as variable-binding operators. The set of α-variants that '∃' looks to when assessing its embedded formula are not at all constrained by what the initial variable assignment assigns to α, so shifting what the input assignment assigns to α will be idle. Whereas, the set of α-variants that 'Ɐ' looks to when assessing its embedded formula *are* constrained by what the initial variable assignment assigns to α—they have to assign something greater than (or equal to) what the initial variable assignment assigns to α—thus shifting the input assignment can make a difference. In this way, '∃' standardly has a certain indifference to the input assignment, whereas for 'Ɐ' the input assignment genuinely matters—'Ɐ' is "context" sensitive.[3]

This suggests that the essential difference between '∃' and 'Ɐ' concerns the *accessibility relations* involved. Thus, it can be illuminating at this point to view first-order logic as a modal logic in the way associated with Amsterdam (see Van Benthem 1977 and especially Blackburn et al. 2002, §7.5 on reverse correspondence theory). On this way of viewing things, the assignments are the "worlds", and the model includes a stock of binary accessibility relations R^α that hold between assignments (relative to a variable α). Standard first-order logic is only concerned with a special subset of all the possible such models for first-order languages—that is, it constrains itself to a particular accessibility relation:

$$g \equiv^\alpha g' \text{ iff } g' \text{ differs from } g \text{ at most in that } g'(\alpha) \neq g(\alpha)$$

Notice that $\equiv^\alpha$ is reflexive, transitive, and symmetric. So given standard assumptions the accessibility relation between assignments is an equivalence relation, and thus will validate the relevant formulae corresponding to S5, which will include the following theorems concerning vacuous quantification:

$$\forall x \forall x \phi \leftrightarrow \forall x \phi$$
$$\exists x \forall x \phi \leftrightarrow \forall x \phi$$
$$\exists x \exists x \phi \leftrightarrow \exists x \phi$$
$$\forall x \exists x \phi \leftrightarrow \exists x \phi$$

Prefixing further quantifiers to a closed sentence in standard first-order logic is "vacuous"—just as adding further boxes and diamonds is vacuous in an S5 modal logic. But—just as in modal logic—the relevant equivalencies only hold given particular restrictions on the accessibility relation. For example, $\Box\phi \to \Box\Box\phi$ is invalid unless the frame is transitive. Likewise, if R^x is not transitive, then $\forall x \phi \to \forall x \forall x \phi$ will be invalid. To round off this point, consider again the the quantifier 'Ɐ' and the accessibility relation that it appeals to:

$$g \geq^\alpha g' \text{ iff } g' \text{ differs from } g \text{ at most in that } g'(\alpha) > g(\alpha)$$

This relation is reflexive and transitive, but it is not symmetric. And our counterexample to the vacuous quantification principle implicitly exploited the fact that the relation was not symmetric. Notice that, in general, if the accessibility relation between assignments is not assumed to be symmetric then the paradigm example of vacuous quantification mentioned at the outset becomes invalid.

[3] An alternative way to put the difference here is to say that standard first-order quantifiers take an "external" perspective, whereas a quantifier like 'Ɐ' must take an "internal" perspective on the relevant relational structures (see Recanati 2007, 65-71 and Blackburn et al. 2002, xi-x).

(2) $\forall x \exists x Fx \leftrightarrow \exists x Fx$

Consider the right-to-left direction: If someone taller than me is happy, it doesn't follow that everyone taller than me is such that there is someone taller than them that is happy.[4]

Comment B. But is there anything of interest or use concerning these non-standard quantifiers and accessibility relations? One can imagine that by generalising and playing around with the first-order accessibility relations there are things of intrinsic interest to metalogic, e.g. concerning decidability (see discussion in Blackburn et al. 2002: 466-469), but is there anything more applicable that the non-standard quantifiers can help to model? I think there are a few good answers to this. Let me mention one. There are enlightening and perhaps surprising connections to the extant literature on counterpart semantics (Lewis 1968). Here the falsity of the Principle of Vacuous Binding is already, in a sense, presupposed. Lewis provides a first-order translation for formulae of quantified modal logic such as

$$\Box Fx \quad \approx \forall y \forall z((Wy \wedge Izy \wedge Czx) \rightarrow Fz)$$

This says (roughly) that every counterpart of x, in any world, is an F. And such a translation generalises for any modalised open sentence '$\Box \pi \alpha_1 \ldots \alpha_n$'. We can follow the translation procedure, and then provide the resulting first-order formulae with their standard first-order interpretation, thereby endowing the modal formulae with truth-conditions. But can we instead directly provide a model-theoretic semantics for the language of quantified modal logic that corresponds to Lewis' translation rules? Yes.[5] But in order to interpret an open modalised formulae such as '$\Box Fx$' we need to evaluate the embedded open formula 'Fx' relative to every counterpart of x at every world. Thus, if '$\Box Fx$' is evaluated at a variable assignment g, then 'Fx' must be evaluated at assignments g' where $g'(x)$ is a counterpart of $g(x)$ at the the relevant world. So the model should include a counterpart relation C. One natural way to implement this—see Schwarz (2012) for a slight alternative—is to let C be a binary relation that holds between assignment-world pairs, that is a relation on $D^{\mathbb{N}} \times W$.[6] In this way we have the modal operator evaluate its embedded formula at assignment-world pairs that are "accessible" via the relation on worlds and the counterpart relation.

- $[\![\Box\phi]\!]^{g,w} = 1$ iff for all w' and g' such that wRw' and $\langle w, g\rangle C \langle w', g'\rangle$, $[\![\phi]\!]^{g',w'} = 1$

Notice that the modal operator effectively binds all the variables in its scope. But since the assignments that '$\Box$' looks to when assessing its embedded formula ϕ *are* constrained by what the

[4]There is a tight connection here between '$\forall\!\!\!\!\exists$' and the kinds of quantifiers used for knowledge representation by description logics, e.g. '$\exists R$' (see Blackburn 2006 and Baader et al. 2003): they both restrict to the set of individuals that bear a relation to the input individual. The sentences of description logic function much like the sentences of Priori's Egocentric logic such as "Someone-more-perfect standing", which is true at an individual a iff there is an x such that x is more perfect than a and x is standing (see Prior 1968).

[5]See Schwarz (2012), and early discussion in Hazen (1979).

[6]Let a model $\mathfrak{A} = \{W, R, D, C, I\}$, where W is a (non-empty) set of worlds, R is a binary accessibility relation on W, D is a (non-empty) set of individuals, C is the counterpart relation, a binary relation on $D^{\mathbb{N}} \times W$. My presentation differs from the presentation in Schwarz (2012). Besides the mere notational differences, Schwarz has the counterpart relation hold between individual-world pairs, and then uses this to construct the required alternative sequences (see p. 13*ff*), whereas I have a counterpart relation on sequence of individuals and world pairs. There is an issue here concerning multiple counterparts at one world (see Hazen 1979, 328-330, and Lewis 1983, 44-45), but we will gloss over it here.

initial variable assignment assigns to the variables in ϕ, prefixing a quantifier to '$\Box\phi$' needn't be vacuous. So although the modal operator binds the variables, they can be re-bound.

In the postscript Lewis, in fact, mentions that his modal operators bind the variables in their scope when addressing a certain objection (Lewis 1983). The objection is that his counterpart theory seems to invalidate Leibniz's Law, since the following fails:

(6) $\forall x \forall y(x = y \rightarrow (\Diamond x \neq y \leftrightarrow \Diamond y \neq y))$

Lewis insists that on his view (6) is not an instance of Leibniz's Law for much the same reason that (7) isn't an instance.

(7) $\forall x \forall y(x = y \rightarrow (\exists y\, x \neq y \leftrightarrow \exists y\, y \neq y))$

Clearly, to think that (7) is an instance of Leibniz's Law is, as Lewis says, "to commit a fallacy of confusing bound variables". Lewis insists that the same holds for (6), since the modals bind the variables in their scope.[7]

> The abbreviated notation of quantified modal logic conceals the true pattern of binding . . . The diamonds conceal quantifiers that bind the occurences of 'x' and 'y' that follow. (Lewis 1983: 46)

Recently theorists have been appealing to "assignment-shifting operators" in treatments of certain natural language constructions such as attitude verbs and epistemic modals, see, e.g., Cumming (2008), Santorio (2012), Ninan (2012), and Pickel (2015). These views all fall within the family of counterpart semantics, broadly construed, and share the feature that the modals or attitude verbs bind all the variables in their scope. If quantifying into the relevant constructions is vacuous, then such views are hopelessly misguided and empirically inadequate. One attempt at stating such views schematically (where '$\Box$' stands for the a belief operator or an epistemic necessity modal) would be as follows:

$$[\![\Box\phi]\!]^{g,w} = 1 \text{ iff for some } \langle g', w'\rangle Rw,\ [\![\phi]\!]^{g',w'} = 1$$

According to such a semantics the truth of '$\Box Fx$' does not depend on what the input assignment assigns to 'x', thus we get the undesirable result that "quantifying in" is blocked (see Pickel 2015: 339-341 for a criticism of Cumming 2008 that essentially exploits this feature of his view). Among these authors there is disagreement over whether '$\Box Fx$' should be sensitive to the input assignment, and if so how best to achieve this sensitivity.[8] But in light of our discussion here a very promising approach is to follow the lead of counterpart semantics (as implemented above) and provide a semantics that appeals to assignment-world pairs that are accessible from the input assignment and world. Schematically, this would go as follows:

$$[\![\Box\phi]\!]^{g,w} = 1 \text{ iff for some } \langle g', w'\rangle R\langle g, w\rangle,\ [\![\phi]\!]^{g',w'} = 1$$

[7] See Schwarz (2012: 17-18) for discussion of the appropriate restrictions on substitution in a counterpart semantics.

[8] For example, Cumming (2008) appeals to machinery inspired by Kaplan's (1968) treatment of de re attitude ascriptions, while Pickel (2015) argues against this view, and instead provides a two-factor view where the truth of '$\Box Fx$' depends on both the truth (at all relevant worlds) of 'Fx' relative to shifted assignments and relative to the *input* assignment. The latter conjunct gives rise to the arguably bad result that "Olivia believes that Hesperus is distinction from Phosphorus" cannot be true—one diagnosis might be that attitude reports have become *overly* sensitive to the input assignment. Santorio (2012) doesn't mention quantifying in, but it seems that prefixing a quantifier needn't be vacuous since his counterpart relations appeal to the initial variable assignment. Ninan (2012) perhaps mostly closely resembles the counterpart semantics described here, though it has additional bells and whistles.

References

Baader, F., Calvanese, D., McGuinness, D., Nardi, D. and Patel-Schneider, P.: 2003, *The Description Logic Handbook: Theory, Implementation and Applications*, Cambridge University Press.

Blackburn, P.: 2006, Arthur Prior and hybrid logic, *Synthese* **150**(3), 329–372.

Blackburn, P., De Rijke, M. and Venema, Y.: 2002, *Modal Logic*, Vol. 53: Cambridge Tracts in Theoretical Computer Science, Cambridge University Press.

Cumming, S.: 2008, Variablism, *Philosophical Review* **117**(4), 605–631.

Hazen, A.: 1979, Counterpart-theoretic semantics for modal logic, *The Journal of Philosophy* **76**(6), 319–338.

Heim, I. and Kratzer, A.: 1998, *Semantics in Generative Grammar*, Blackwell Publishers.

Kalish, D. and Montague, R.: 1964, *Logic: techniques of formal reasoning*, Oxford University Press.

Kaplan, D.: 1968, Quantifying in, *Synthese* **19**(1-2), 178–214.

Lewis, D.: 1968, Counterpart theory and quantified modal logic, *The Journal of Philosophy* pp. 113–126.

Lewis, D.: 1983, Postscripts to "Counterpart theory and quantified modal logic", *Philosophical Papers I*, Oxford: Oxford University Press, pp. 39–46.

Ninan, D.: 2012, Counterfactual attitudes and multi-centered worlds, *Semantics and Pragmatics* **5**(5), 1–57.

Pickel, B.: 2015, Variables and attitudes, *Noûs* **49**(2), 333–356.

Prior, A.: 1968, Egocentric logic, *Noûs* **2**(3), 191–207.

Recanati, F.: 2007, *Perspectival Thought: A Plea for (Moderate) Relativism*, Oxford University Press.

Santorio, P.: 2012, Reference and monstrosity, *Philosophical Review* **121**(3), 359–406.

Schwarz, W.: 2012, How things are elsewhere: Adventures in counterpart semantics, *in* G. Restall and G. Russell (eds), *New Waves in Philosophical Logic*, Palgrave Macmillian, pp. 8–29.

Van Benthem, J.: 1977, *Correspondence theory*, PhD thesis, University of Amsterdam.