I

Modality

1

Ruth Barcan Marcus and the Barcan Formula

TERENCE PARSONS

This smart, well-spoken and beautiful woman enjoyed among her students and also in wider circles wondering admiration, . . . ; she was also sought out by holders of political powers and enjoyed great influence. Morally without blemish, she moved with secure freedom even among men.

said of Hypatia, 4th–5th century logician[1]

Ruth Barcan Marcus is and will remain famous for the Barcan formula. The questions I want to address are these: What is it? Did she really invent it? and Is it worth being famous for? (There is no need for drama; the answers to the last two questions are yes.) I will not discuss the question of whether she deserves to be known for other things as well; the answer to that is also yes, but that is beyond the scope of this essay.

What Is the Barcan Formula?

Part of Ruth Barcan's dissertation project was to give a systematic axiomatic treatment of modal logic as developed by Lewis and Langford in their 1932 textbook, including especially the predicate calculus part containing quantifiers. She did this, and published the results in a relatively new journal, the *Journal of Symbolic Logic* (volume 11), in 1946. This is the first systematic published work that addressed the question of combining unrestricted quantification with unrestricted modality, and the Barcan formula is an essential

I first met Ruth Marcus in September 1965, when I walked into her office to introduce myself. This was not a social call; she was my new boss, having hired me fresh out of Stanford to a brand new department in the newly opened Chicago campus of the University of Illinois. I'll never forget that meeting. She was pleasant, charming, gracious, hospitable, and she scared me to death. What scared me was her asking me whether my dissertation advisor was urging me to publish my thesis. He wasn't. It wasn't good enough. The message was clear: I was in the big leagues. And I wasn't sure I could measure up. Fortunately, it wasn't long before Ruth was letting me know that in her eyes I was measuring up, and that was a great comfort to me. That combination, of setting high standards while also being supportive is just what one wants in a mentor, and I can admit today that I needed a mentor then. I thank Ruth for filling that role for me.

part of her account. The formula is

$$\Diamond(\exists x)S \equiv (\exists x)\Diamond S,$$

or, equivalently, in terms of the universal quantifier and necessity:

$$\Box(x)S \equiv (x)\Box S.$$

Aficionados know that in one direction these biconditionals are called the Barcan formula, and, in the other direction, the converse Barcan formula, but most people (myself included) can rarely remember which is which. Marcus endorsed both directions, and the formula is important and controversial in both directions. (In the first example, the left-to-right implication is the Barcan formula and the converse implication is the converse Barcan formula. Conversely, in the latter formula the right-to-left implication is the Barcan formula and the converse implication is the converse Barcan formula.)

These formulas are important for two reasons. One is the use of notation that clearly distinguishes two different readings, and the other is a proposal for how they are related.

Background

To appreciate the significance of the Barcan formula, we need to know something about the historical background of how modalities related to quantification before 1946. The interaction between modality and quantifiers was discussed by Aristotle, who considered the difference between

An animal is necessarily moral.

and

Necessarily, an animal is mortal.

Aristotle's terminology for the distinction was *composition* and *division*, which is essentially one of grammatical scope. This terminology was used for centuries, and persists today. It is a relatively theory-neutral way to mark the distinction. A new step was taken near the end of the twelfth century when an anonymous writer suggested that the first reading could be considered to attribute a necessity to a *thing*, and the latter to attribute necessity to *what is said*; the distinction came down to us as the contrast between *de re* and *de dicto* modality.[2] The distinction was carefully respected by the great logicians of the thirteenth and fourteenth centuries, though it became less central as logic languished in the centuries thereafter, and you will find what we would call *de re/de dicto* confusions throughout the modern era.

None of this earlier work that respected the *de re/de dicto* distinction was in a position to address the issue Marcus addressed, for two reasons. First, the authors usually had in mind not logical necessity, but something else, something that we would see as more restricted. Second, before Frege, unrestricted quantification was unknown; all logicians used restricted quantification. Let me expand on this latter point. The most sophisticated treatment of quantification prior to Frege was found in supposition theory, as developed by people like Ockham, Buridan, and Paul of Venice. Here is an example of the sort of theoretical apparatus they were dealing with. The term 'donkey', they would say, has *determinate supposition* in the sentence:

Some donkey is brown.

This is because from that sentence one may infer:

This donkey is brown, or that donkey is brown, or that one, and so on for all the donkeys.

This is like saying that 'donkey' is existentially quantified, with wide scope. Similar tests were used for distributive supposition of a term (something like being universally quantified, though not necessarily with wide scope) and merely confused supposition.

It is still not clear today just what was supposed to be accomplished by analyses of the suppositional status of terms, but it is apparent from the sort of tests that were used that an essential precondition for the application of supposition theory is that its goal is a study of restricted quantification. The theory does not classify quantifiers regarding quantity and scope; it classifies terms (predicates). In 'Some donkey is an animal' the occurrence of the quantifier word 'some' is what *gives* the word 'donkey' determinate supposition, but the thing that *has* determinate supposition is the word 'donkey', not the word 'some'. This is a theory that classifies the words that restrict quantifiers, not the quantifiers themselves, and so it is essentially a theory of restricted quantification. It was only when Frege tore apart the quantifiers from their restrictors that unrestricted quantification came to exist, when he symbolized 'Some donkey is an animal' as '$(\exists x)(Dx \,\&\, Ax)$', and then gave a semantics for the quantifier apart from the conjunction. Before Frege, the Barcan formula could not have been written.

Frege, of course, was unknown by most English-speaking philosophers, and it was only after Russell and Whitehead's *Principia Mathematica* that attention to something like modern quantification theory got momentum, impelled in part by Russell's provocative applications of it to philosophical puzzles. These often employed something like his theory of definite and indefinite descriptions, which explicitly uses unrestricted quantifiers to analyze restricted ones. A couple of decades later, modalities were being looked at seriously in modern terms by Lewis and Langford in 1932,[3] and fourteen years after that Ruth Barcan published the first systematic treatise mixing logical modality and unrestricted quantification.

How do I know Marcus was first? My evidence is of three kinds. First, I've asked several knowledgeable people about antecedents, and I've searched some of the likely literature myself. That's the least conclusive part of the evidence. Second, I asked Marcus herself. Surely, given her fame in connection with the Barcan formulas, someone would have pointed out to her antecedents, had there by any. Nobody has. Third – and here is proof beyond Cartesian doubt: Alonzo Church says so. In his first essay on sense and denotation he says, ''The first systematic treatments [of modality in connection with quantifiers] are those of Barcan and Carnap (independent, and nearly simultaneous).'' This narrows it down to Barcan and Carnap, and Barcan published first.[4]

Contemporary Uses of the Barcan Formula

The term 'Barcan formula' has come to have at least three different uses:

1. Any mixing of modalitylike things with quantifiers is often referred to as an instance of the Barcan formula. And this is fair. It wasn't just that modalities hadn't been carefully and systematically mixed with quantifiers before Marcus did it; other intentional operators such as 'knows that' and 'ought' hadn't either, and once a paradigm was in the air, people's interest was directed to important issues. Often in such applications the analogs of the Barcan formulas do not hold, but that is not the important point. By providing an examplar, Marcus focused attention on addressing logically and semantically important issues.

2. In a more restricted context, ''Barcan formula'' sometimes refers to the mixing of any alethic modality with any quantification. Here there are still plenty of counterexamples to the formulas, using restricted kinds of necessity and quantification. For example, consider the inference:

No existing American can run a mile in under 3:45.
∴ It's impossible for an American to run the mile in under 3:45.

If you reject this inference, then you do not accept what is sometimes called the Barcan formula for this case. This is the Barcan formula extended so as to apply to old-fashioned restricted quantification and restricted modality; the quantification is restricted to Americans, and the modality is restricted to what is possible consistent with laws of physiology, gravity, and existing constitutions of human bodies. Construed in such ways, the Barcan formula usually does not hold. But, as I have noted, this is not what Marcus proposed in her original papers. She had in mind unrestricted modality and quantification.

The restricted versions can be expressed in logical notation in ways well known from Frege and Russell, something like this:

(x) (x is an American → □ (physical laws are such and such →
  ¬ (x runs a mile in under 3:45))
∴ □ (physical laws are such and such →
  (x)(x is an American → ¬(x runs a mile in under 3:45))

Nowhere is the pure Barcan formula to be found here. The viability of the Barcan formula as a law of pure logic for unrestricted quantification and modality is untouched by such examples.

3. The third use of the term 'Barcan formula' is its strict use in its original application to unrestricted modality and quantification. And here, it may be completely defensible. This is controversial, because many contemporary philosophers think the Barcan formulas are incorrect as originally proposed. I think that the jury is still out on the issue. To be specific, the arguments that are currently popularly accepted as disproving the Barcan formulas beg questions by explicitly or implicitly using restricted quantification or restricted modality, or by utilizing logical devices that transcend the framework within which the formulas were originally proposed. For the remainder of this paper I will address these points.

But first let me clarify my defense. The Barcan biconditional is sufficiently controversial that I am not comfortable defending it as an a priori principle of logic. It settles too many issues that are debated by reasonable people to be a logical principle. But that leaves open the possibility that every instance of the formula is true. This is what I will concentrate on.

Putative Counterexample to the Converse Barcan Formula

I begin with a putative counterexample to the converse Barcan formula, usually attributed to Kripke:

Necessarily, everything exists. TRUE
∴ Everything necessarily exists. FALSE

In symbols:

□ (x)(x exists) TRUE
∴ (x) □ (x exists) FALSE

The converse Barcan formula endorses this inference, and it allegedly leads from truth to falsehood.

Is this a good counterexample? I think not. It plays on an equivocation in 'exists'. To see this, begin by asking why we think the premise is true. Virtually all nonphilosophers as well as most philosophers do not think that everything exists; numbers don't exist, properties don't exist, and so on. So it is doubtful that many people would agree that necessarily, everything exists. The premise is apparently not plausible at all. The rejoinder to this is that people who object to the premise are taking 'exist' in a restricted sense, one that does not exhaust the range of the quantifiers. To get the counterexample, you can't restrict 'exists' at all. This can be made precise by replacing 'x exists' by '$(\exists y)(y = x)$'. The counterexample then becomes

$\Box (x)(\exists y)(x = y)$ TRUE
$\therefore (x) \Box (\exists y)(x = y)$ FALSE

So construed, the premise is plausible. But now why reject the conclusion? The conclusion no longer says 'Everything has necessary existence', it says 'Everything is necessarily something' and that sounds plausible, not implausible. After all, everything is necessarily itself, so why isn't everything necessarily something? Remember, being something does not *mean* being something that exists; we are being careful not to sneak in the assumption that if I am necessarily something then I necessarily exist. The Barcan formula only gets us to the claim that I am necessarily something, it does not get us the further inference from that claim to the claim that I necessarily exist. This further inference is accepted by many (though not all) current possible worlds theorists, but it is typically there by assumption, not by argument.

I have argued this way before, so I am aware of the usual rejoinder. This is to note that quantifiers have domains, and that my inference from 'Everything is necessarily itself' to 'Everything is necessarily something (namely, itself)' illegitimately imports a domain along with the 'something'. Quantifiers have domains, these change when you go from world to world, and their presence needs to be taken account of.

Suppose this is right; where does it get us? This requires a slight excursion into domains. Suppose I say, "Everything's calm today," and you reply with "Oh yeah, what about that tenement fire in Minsk?" That is probably not a pertinent rejoinder. You have countered my claim:

(x) (Calm x)

with a true claim:

$\neg$ (x) (Calm x)

but your claim and mine are not inconsistent in spite of appearances; my domain is local events, and your is everything on earth. How can logic handle such phenomena? One way is to go contextual, and to claim that the relativity of domains to quantifiers depends on context. On this approach there is no straightforward application of the logical inferences we teach about in logic courses. I am not at all opposed to this approach, but we really do not have a well-developed theory of how it works. Another equally good approach is to handle such examples by being explicit about the domains. When we say 'everything' we symbolize it as 'every *thing*', and then if you use the same word with a different meaning from my word, we use different words to symbolize our remarks. My claim is symbolized:

$(x)\,(\text{Thing}_1\ x \rightarrow \text{Calm}\ x)$

and yours as

$\neg\ (x)\,(\text{Thing}_2\ x \rightarrow \text{Calm}\ x)$

Then your response is not pertinent to mine because you mean something different by 'Thing_2' than I mean by 'Thing_1'.

We can do modal logic in this way as well, being explicit about the domains of our quantifiers. The counterexample proposed before is then symbolized as

$\Box\,(x)(x \text{ is a thing} \rightarrow (\exists y)(y \text{ is a thing} \mathbin{\&} x = y))$ TRUE
$\therefore (x)(x \text{ is a thing} \rightarrow \Box\,(\exists y)y \text{ is a thing} \mathbin{\&} x = y))$ FALSE

The domains then change from possible world to possible world in the same way that the extensions of predicates do.

I now agree that the premise is true and the conclusion false; this argument is not a good one. But it is also no longer a counterexample to the Barcan formula, because it is no longer an instance of the Barcan formula. The introduction of explicit domains has turned the example into one containing restricted quantification. I am not aware of any way to improve on this line of argumentation without begging the question. (For example, moving the modal sign in the conclusion to just inside the initial quantifier produces an example about which partisans will naturally disagree.)

Putative Counterexample to the Barcan Formula

Let me turn now to the argument that most people see as the clincher against the Barcan formula proper. Couldn't there be things different from any actual thing? Not different in kind, but different numerically? If so, don't we have an easy refutation of the Barcan formula? Suppose we use 'actual' to pick out things in the domain of the actual world, and suppose that we alter our modal logic to treat this predicate in a logically special way: It is to be unaffected by modal operators. The counterexample is then:

Possibly, there is something different from any actual thing.
∴ There is something that is possibly different from any actual thing.

Allegedly, the premise is true, because the actual world is not necessarily as large as any world can be; some possible world is bigger, and so it has objects in it that are different from any actual ones. And the conclusion is false because each actual thing is necessarily itself, and so any actual thing is necessarily a thing that in the actual world is actual.

$\Diamond\ (\exists x)(x \text{ is a thing } \&\ (y)(y \text{ is actual} \rightarrow x \neq y))$ TRUE
$\therefore\ (\exists x)\ \Diamond\ (x \text{ is a thing } \&\ (y)(y \text{ is actual} \rightarrow x \neq y))$ FALSE

This bypasses any maneuvers with domains. However we read '*x* is a thing', the conclusion follows from the premise by the Barcan formula, and, it is alleged, when '*y*' is restricted in possible worlds to things that are actual in the actual world, the premise is true and the conclusion false.

There are two natural responses to this alleged counterexample. First, it is important to point out the inconclusiveness of the premise. It assumes that this world does not necessarily contain as many things as it could. But is this so? Much recent work in set theory has concentrated on large cardinal assumptions, and there is a natural inclination to demand largeness without limit. Maybe the actual world does contain as many things as a world could possibly contain. And if the existence of pure sets is a matter of necessity, then every world contains as many things as it can. This, together with a large dose of antiessentialism, undercuts the claim that there might have been things in addition to the things that there are.[5]

Suppose you do not find the preceding speculation persuasive. There is still room to question the significance of the counterexample to the Barcan formula. This proposed counterexample is reminiscent of arguments based on Cambridge change. Cambridge change is like this: Suppose that yesterday you were not equal in size to me, but today you are. I didn't change height, but you grew. Then, in a sense, I didn't change, though something is true of me now that was not true of me yesterday. I underwent a Cambridge change: I changed without there being a change in me.

Some people are skeptical of essentialism. Perhaps necessarily cyclists are two-legged (*de dicto*), but is each cyclist two-legged (*de re*) of necessity? Certainly not. Nor are donkeys necessarily animals, if necessity is construed *de re*.[6] This denial of essentialism has wide (though not universal) support. But then people like Plantinga come along and argue that you *have* to endorse essentialism.[7] Here, for example, is an essence of Ruth Marcus: It is true of her in every possible world that she is a philosopher *in the actual world*. It's true in each world that she's a philosopher in this one. Is that essentialism? That's Cambridge essentialism! It's an artificial construction that fits the formula for essentialism: a property that Marcus has in every world. That is the sort of thing that is going on in the counterexample to the Barcan formula. We have a predicate, 'is actual', that is true of a thing in a possible world if something is true of it in this world. We express this Cambridge essence in our language, and as a result, the Barcan formula turns out to be false. Being actual in this sense is a nonlogical essence, but it's a Cambridge essence.

If the only way to conclusively refute the Barcan formula is to expand the notation of modal logic so as to create logically special predicates so as to be able to express Cambridge essences in the object language, then maybe there is something deeply true about what is under attack, something that we have not yet completely fathomed.

NOTES

1. Pauly's *Real-Encyclopädie der Classichen Altertumswissenshcaft*, entry for ''Hypatia.''
2. Personal communication from Calvin Normore, based on a manuscript translated by Norman Kretzman.
3. C. I. Lewis and C. H. Langford, *Symbolic Logic* (New York: Dover, 1932).
4. The Church reference is to ''A Formulation of the Logic of Sense and Denotation,'' P. Henle, H. M. Kallen, and S. K. Langer (eds.), *Structure, Method and Meaning* (New York: The Liberal Arts Press, 1951); p. 22 , n. 23. Marcus's article is ''A Functional Calculus of First Order Based on Strict Implication,'' *Journal of Symbolic Logic* 11 (March 1946; submitted in September 1945); Carnap's is ''Modalities and Quantification,'' *Journal of Symbolic Logic* 11 (June 1946; submitted in November 1945).
5. Antiessentialism is required as follows. Certainly there might have been more porcupines than there are. So there must be another possible world in which there are things that are porcupines that are not porcupines in this world. The cardinality assumption guarantees that the alternate world has enough things in it to play the roles of the additional porcupines, but the additional things will not be porcupines in the actual world. We thus need the possibility that a thing that is not a porcupine in this world is one in the other world; that thing thus cannot essentially be or not be a porcupine. The point becomes more poignant if we assume that there could be more concrete objects in another world than in this one, thus requiring (by a similar argument) that being concrete or abstract cannot be essential to at least some things.
6. William Ockham even attributes to Aristotle the view that it is false that men are necessarily animals, when this is read *de re* (when it is read as equipollent to a modal proposition in the sense of division). See A. Freddoso and H. Schuurman (trans.), *Ockham's Theory of Propositions* (Notre Dame, Ind.: University of Notre Dame Press, 1980), pp. 111–12.
7. *The Nature of Necessity* (Oxford: Clarendon Press, 1974).