Review of Kit Fine, 'Failures of the Interpolation Lemma in Quantified Modal Logic' (Journal of Symbolic Logic, 44 (1979), pp. 201-206)

Saul A. Kripke

This is the published version of Kripke, S. A. (1983), 'Kit Fine. Failures of the Interpolation Lemma in Quantified Modal Logic. *The Journal of Symbolic Logic*, vol. 44, (1979), pp. 201-206.' *Journal of Symbolic Logic*, 48(2): 486-488, which can be obtained from the publisher at https://doi.org/10.2307/2273567. It is reproduced here by permission of the Association for Symbolic Logic which holds the copyright.

KIT FINE. *Failures of the interpolation lemma in quantified modal logic.* ***The journal of symbolic logic,*** vol. 44, (1979), pp. 201–206.

The author proves that the Craig interpolation lemma fails for quantified S5 with variable domain, and for all quantified modal logics with constant domain between K and S5. The theorems refute some erroneous published "proofs" of the lemma for certain modal logics; a vague presupposition that the lemma ordinarily holds in modal logic may have been widespread even earlier. Although the author does not say so, the results appear to be by-products of his interest in showing that various notions seeming to require a "possibilist" ontology of worlds and possible individuals for their definition can in fact be defined in an "actualist" language whose means of expression are restricted, roughly speaking, to modal operators, truth functions, and quantification over actual individuals and propositions. (See the author's work in XLIV 654 (8, 9).)

Consider an arbitrary universal quantificational frame (model structure) for quantified S5, where **K** is the set of all worlds, and $\psi(\mathbf{H})$ is the domain of each $\mathbf{H} \in \mathbf{K}$. (A frame is universal if and only if the accessibility relation $\boldsymbol{R}$ is $\mathbf{K} \times \mathbf{K}$; that is, all worlds are related to each other. It is well known that quantified S5 is sound and complete for such frames.) In a model based on the frame, given the assignment to all free variables of a formula A except x, the *inner* (actualist) *quantifier* $(x)A$ is characterized semantically by stipulating that $(x)A$ is to be true in a world **H** if and only if A is true in **H** for every assignment to x of an element a of $\psi(\mathbf{H})$. The *outer* (possibilist) *quantifier* $(\forall x)A$ is correspondingly characterized by the stipulation that $(\forall x)A$ is true in **H** if and only if A is true in **H** for every assignment to x of an element of $\mathbf{U}(= \bigcup_{\mathbf{H} \in \mathbf{K}} \psi(\mathbf{H}))$. The author shows, in effect (he does not explicity state this—see the next paragraph for the author's version), that, for S5, the outer quantifier is not definable in the usual first-order modal language with the inner quantifier. (Probably others have obtained this result independently of the author. Allen Hazen states the same result, but without giving a complete proof, in *Expressive completeness in modal language,* ***Journal of philosophical logic,*** vol. 5(1976), pp. 25–46; see p. 35. The reviewer once heard David Kaplan announce the result in a talk at a conference in 1975, and there may be others.) Nevertheless the outer quantifier $(\forall x)A$ is definable in S5 if the first-order modal language with the inner quantifier is augmented by propositional quantifiers (interpreted as ranging over arbitrary subsets of **K**). $(\exists p)(p \wedge \Box(x)\Box(p \supset A))$, or alternatively $(p)(p \supset \Box(x)\Diamond(p \wedge A))$, where p does not occur in A, are both appropriate definitions. (The author does not explicitly state this here, but see his discussion of such matters in XLIV 654 (8, 9), and see the next paragraph. In the aforementioned talk, Kaplan announced the related result that the outer quantifier is definable in S5 with inner individual quantifiers and an actuality operator.) This immediately refutes the "Souslin–Kleene" theorem for quantified S5 with variable domain; that is, $(\forall x)Fx$ is definable in the second-order language both as an existential and as a universal formula, but it is not definable in the first-order language. Since, as is well known, the Craig interpolation lemma implies the Souslin–Kleene theorem, the Craig interpolation lemma is refuted as well. In fact, it is immediate that $p \wedge \Box(x)\Box(p \supset Fx) . \supset . q \supset \Box(x)\Diamond(q \wedge Fx)$ is a valid formula with no interpolant.

Actually, the author does not state that the Souslin-Kleene theorem is disproved; but by a similar argument he refutes another consequence of the Craig interpolation lemma, the Beth definability theorem. (The resulting argument is somewhat less perspicuously connected to the question of the definability of the outer quantifier, and the refutation of the interpolation lemma is perhaps slightly less direct.) Consider the theory T whose axioms are $\Box(x)\Box(p \supset Fx)$ and $\Diamond(\exists x)\Box(Fx \supset p)$. The author shows that if both axioms of T are true in a given world of a universal S5 model, then p is true in that world if and only if $(\forall x)Fx$ is true there, which establishes that p is implicitly definable in T (with quantified S5 as the underlying logic). That is, if one adds $\Box(x)\Box(q \supset Fx)$ and $\Diamond(\exists x)\Box(Fx \supset q)$ to T, then $p \equiv q$ follows from T in quantified S5. On the other hand, the author uses essentially the same argument that would show that $(\forall x)$ is not definable in the first-order modal language with inner quantifiers to show that p is not explicitly definable in T. That is, for no formula B of the first-order modal language does $p \equiv B$ follow from T in quantified S5. (In fact, it is easy to show that the two claims are equivalent.) (Actually, the author presents his argument in terms of the outer *existential* quantifier; the presentation above, which uses the outer *universal* quantifier, modifies the author's argument accordingly.) The author also gives the same example as in the previous paragraph of a conditional valid in quantified S5 that has no interpolant, and he shows directly that in quantified S5 with identity it has no interpolant even if Fx is replaced by $(\exists y)(y = x)$.

It is worth noting, since the author does not state the fact, that his argument can easily be adapted to refute the interpolation lemma for quantified B (the "Brouwersche" system, characterized by symmetry and reflexivity of the accessibility relation). Redefine $(\forall x)A$ as a "local" outer quantifier; $(\forall x)A$ is true in **H** if and only if A is true in **H** no matter what element of $\mathbf{U}'(\mathbf{H})$ is assigned to x, where $\mathbf{U}'(\mathbf{H}) = \bigcup_{\mathbf{H}R\mathbf{H}'} \psi(\mathbf{H}')$, and R is the accessibility relation between worlds. The arguments just stated then go through for B. It is easy to modify the arguments further so that they will refute the interpolation lemma for quantified KB. Symmetry of the accessibility relation thus appears to be the key to these arguments.

It is natural to ask next what the situation is for modal logics with constant domain, since this corresponds to the case where the outer quantifier is taken as primitive. In this case the author refutes the interpolation lemma for all systems between quantified K (plus the schema $(x)\Box A \equiv \Box(x)A$) and quantified S5 (with the same schema). Call these systems K* and S5*, respectively. Consider a frame (**K**, R), where **K** is the set of worlds and R is the accessibility relation, with a constant domain **D**. Introduce a quantifier $(\mathrm{Q}x)$ by stipulating that, given a fixed assignment to the free variables of A other than x, $(\mathrm{Q}x)A$ is to be true in **H** if and only if there is an **H**$'$ such that $\mathbf{H}R\mathbf{H}'$ and for no $a \in \mathbf{D}$ is A true in both **H** and **H**$'$ when x is assigned a. (That is, $(\mathrm{Q}x)Fx$ says that it is possible that the extension of F should be disjoint from its true extension.) The author shows (in effect) that $(\mathrm{Q}x)A$ is not definable in the first-order modal language even if we restrict ourselves to universal frames with constant domain. (On page 39 of his paper, Hazen conjectures the same result but is unable to provide a complete proof.) In the second-order language, however, $(\mathrm{Q}x)A$ is definable both as $(\exists F)((x)((A \supset \Box Fx) \wedge (\Diamond Fx \supset A)) \wedge \Diamond(x)\sim(A \wedge Fx))$ and as $(F)((x)((A \supset \Box Fx) \wedge (\Diamond Fx \supset A)) \supset \Diamond(x)\sim(A \wedge Fx))$, where F does not occur in A, and the definitions are valid in all frames with constant domain. Thus, the Souslin–Kleene theorem, and hence, the interpolation lemma, fail for all systems between K* and S5*. More precisely, the positive definability result for $(\mathrm{Q}x)$ implies, given the completeness of K* for validity in arbitrary frames with constant domain, that

$$(x)((Hx \supset \Box Fx) \wedge (\Diamond Fx \supset Hx)) \supset \Box(x)\sim(Hx \wedge Fx) \,.\supset.\; (x)((Hx \supset \Box Gx) \wedge (\Diamond Gx \supset Hx)) \supset \Diamond(x)\sim(Hx \wedge Gx)$$

is provable in K*, while the negative result implies that no interpolant can be found even in S5*. Hence the interpolation lemma fails for all systems between K* and S5*.

As before, the author actually refutes the Beth definability theorem, not the Souslin–Kleene theorem, for all systems between K* and S5*, and does not explicitly put the argument in terms of the definability of $(\mathrm{Q}x)$. Here the theory T' has the two axioms $p \supset \Diamond(x)(Fx \supset \Box(p \supset \sim Fx))$ and $\sim p \supset \Box(\exists x)(Fx \wedge \Box(\sim p \supset Fx))$. The author shows that, in any universal S5 model with constant domain, if the axioms of T' are true in a world, then p is true in that world if and only if $(\mathrm{Q}x)Fx$ is true there, so that p is implicitly definable in T' with S5* as the underlying logic. However, p is not explicitly definable in T' with S5* as the underlying logic. This refutes the Beth definability theorem for S5*. To refute the Beth definability theorem for all systems between K* and S5*, the author uses a supplementary argument to reduce the result to the case for S5*. (The argument of the previous paragraph does not require such a supplementary argument.)

The results on quantified S5 with constant domain refute published claims by J. Czermak (abstract in this JOURNAL, vol. 39 (1974), p. 416), and Kenneth Bowen (*Normal modal model theory*, ***Journal of philosophical logic,*** vol. 4 (1975), pp. 97–131, Theorems 10.2 and 10.3). The author states that Czermak (op. cit.) and Gabbay (*Craig's interpolation theorem for modal logics*, XL 510(16)) have correctly proved the interpolation lemma for many standard systems with the schema $\Box(x)A \equiv (x)\Box A$, i.e., where models are required to have cumulative (but not necessarily constant) domains. The author seems to say that he himself has also proved such results. Melvin Fitting's proofs of the interpolation lemma for certain quantified modal logics with cumulative domains (*Model existence theorems for modal and intuitionistic logics*, this JOURNAL, vol. 38 (for 1973, pub. 1974), pp. 613–627)—which the author does not mention—also accord with the author's assertions and are presumably correct. (The reviewer has not checked them.) The author further asserts that he can prove the interpolation lemma for many systems that lack the schema

$\Box(x)A \equiv (x)\Box A$ and hence permit models without cumulative domains. He does not characterize or list, even partially, the systems for which he has proved the interpolation lemma. Note that the positive results imply, when they hold, that if the outer quantifier is definable by using inner quantifiers and propositional quantifiers, it cannot be definable in both existential and universal forms.

The author asserts that "it would appear that the failures persist when the language is tricked out with possibilist quantifiers, actuality constants, and other such devices. But there may be some natural extension of the modal language which falls short of a full classical language and for which the classical results still hold." (p. 206) The arguments from the Souslin–Kleene theorem suggest that one could start by attempting to characterize all those (generalized) quantifiers that are expressible in both existential and universal forms in the second-order modal language and adding them to the first-order language, since their expressibility is a necessary condition for success.

The reviewer would like to emphasize the philosophical interest of the (very simple) positive definability results. Hazen (in the above-mentioned paper) and others have argued that since the outer quantifier and such quantifiers as (Qx) are not definable in the usual first-order modal logics with the inner quantifier, and since various ordinary locutions seem to invoke such notions, the language with the modal operators and the inner quantifier alone is inadequate. Some have suggested further that ontological conclusions may follow, e.g. that possible individuals must be taken as "real." The second-order definability results cast considerable doubt on any such conclusions. Indeed, provided quantifiers over arbitrary propositions are accepted, it would seem that, at least for languages based on modal operators of the S5 type, the actualist and possibilist languages differ very little (if at all) in expressive power. The differences come out only in the interpretations assigned to various locutions. (For example, the S5 actualist, even though he accepts ($\forall x$) as meaningful in terms of one of the definitions above, does not interpret it the same way as the possibilist, who takes it as primitive.) The author has treated these matters from his own philosophical point of view in XLIV 654(8, 9) and elsewhere.

Are there quantified modal logics that validate the Souslin–Kleene theorem, or the Beth definability theorem, but not the Craig interpolation lemma? Are there quantified modal logics that validate the Souslin–Kleene theorem, but not the Beth definability theorem, or vice-versa?

SAUL A. KRIPKE

S. K. THOMASON. *Noncompactness in propositional modal logic.* ***The journal of symbolic logic,*** vol. 37 no. 4 (for 1972, pub. 1973), pp. 716–720.

KIT FINE. *An incomplete logic containing S4.* ***Theoria,*** vol. 40 (1974), pp. 23–29.

S. K. THOMASON. *An incompleteness theorem in modal logic.* Ibid., pp. 30–34.

MARTIN GERSON. *The inadequacy of the neighbourhood semantics for modal logic.* ***The journal of symbolic logic,*** vol. 40 (1975), pp. 141–148.

MARTIN SEBASTIAN GERSON. *An extension of S4 complete for the neighbourhood semantics but incomplete for the relational semantics.* ***Studia logica,*** vol. 34 (1975), pp. 333–342.

MARTIN GERSON. *A neighbourhood frame for T with no equivalent relational frame.* ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik,*** vol. 22 (1976), pp. 29–34.

V. B. ŠEHTMAN. *On incomplete propositional logics.* ***Soviet mathematics,*** vol. 18 (1977), pp. 985–989. (English translation by B. F. Wells of *O népolnyh logikah vyskazyvanij,* ***Doklady Akadémii Nauk SSSR,*** vol. 235 (1977), pp. 542–545.)

J. F. A. K. VAN BENTHEM. *Two simple incomplete modal logics.* ***Theoria,*** vol. 44 (1978), pp. 25–37.

J. F. A. K. VAN BENTHEM and W. J. BLOK. *Transitivity follows from Dummett's axiom.* Ibid., pp. 117–118.

J. F. A. K. VAN BENTHEM. *Syntactic aspects of modal incompleteness theorems.* Ibid., vol. 45 (1979), pp. 63–77.

Various semantics are available for modal logics, including algebraic "semantics," the Kripke relational semantics, and neighbourhood semantics. There is also the so-called general relational semantics, a variation on the Kripke relational semantics in which only a specified class of valuations is permitted. It is easy to show that the general relational semantics are equivalent to the