Repairing the Interpolation Theorem in Quantified Modal Logic

Carlos Areces[a] Patrick Blackburn[b]
Maarten Marx[a]
[a] ILLC, University of Amsterdam,
Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands
[b] INRIA Lorraine,
615 Rue du Jardin Botanique, 54602 Villers lès Nancy Cedex, France

October 30, 2001

Abstract

Quantified hybrid logic is quantified modal logic extended with apparatus for naming states and asserting that a formula is true at a named state While interpolation and Beth's definability theorem fail in a number of well known quantified modal logics (for example in quantified modal **K**, **T**, **D**, **S4**, **S4.3** and **S5** with constant domains), their counterparts in quantified hybrid logic have these properties. These are special cases of the main result of the paper: the quantified hybrid logic of any class of frames definable in the bounded fragment of first-order logic has the interpolation property, irrespective of whether varying, constant, expanding, or contracting domains are assumed.

keyword quantified modal logic, quantified hybrid logic, interpolation, Beth definability, bounded fragment.

Hybrid logics are extensions of orthodox modal logics in which it is possible to name states (or worlds, or times, or locations…) and to assert that a formula is true at a named state (see for example [10, 9, 3, 8, 1], Chapter 7 of [2], and the webpage `www.hylo.net`). Arguably such extensions are interesting from an applied perspective: we often want to reason about what happens at particular states, worlds, times, or locations, and this isn't possible in orthodox modal logics. For the purposes of the present paper, however, the important point about such extensions is that they often yield logics better behaved than the original modal logic. For example, many modally undefinable properties (such as irreflexivity) are definable in hybrid logic, and general completeness and interpolation results are easier to come by.

Most previous work on hybrid logic has examined the effects of hybridizing *propositional* modal logics. What happens when a *quantified* (first-order) modal logic is hybridized? Orthodox quantified modal logics are often badly behaved. In particular, they typically lack the interpolation property: Kit Fine [5] showed that both Beth's definability theorem and Craig's interpolation theorem fail for first-order **S5** without any assumption on the domains, and that both results fail for any logic between **K** and **S5** when the constant domain axiom schema is added.

Does hybridization improve the situation? As we shall show, yes. Hybrid logics offer precisely the features needed to prove interpolation theorems. Firstly, the basic hybrid machinery of *nominals* (propositional symbols that name states) and satisfaction operators (which enable us to assert that a formula holds at a named state) enable us to express Robinson diagrams of models for quantified modal logic. That is, with the help of nominals and satisfaction operators we can describe the set of states, the accessibility relation, the universe of every state, and the interpretation of the relations symbols in each state. Secondly, the hybrid binder $\downarrow$ lets us replace nominals ('constants') by bound variables. This enables us to express the properties required by interpolants without leaving the original signature.

The paper is organised as follows. The first section reviews orthodox quantified modal logic. The second section adds the apparatus of hybrid logic and uses it to express Robinson diagrams. The third section reviews the counterexamples from Fine [5] and provides interpolants and explicit definitions in the richer hybrid language. The fourth section links hybrid logic with the bounded fragment of first-order logic. The last section contains the main result: *the quantified hybrid logic of any class of frames definable in the bounded fragment of first-order logic has the interpolation property, irrespective of whether varying, constant, expanding, or contracting domains are assumed.* As an immediate corollary, Beth's definability theorem also holds for all such logics.

1 Quantified modal logic

We quickly review the relevant model theory for quantified modal logic (QML). For a thorough treatment the reader is referred to [7]. The language of QML is obtained from the language of classical first-order predicate logic with identity by adding a unary operator $\Diamond$. We only consider signatures without function symbols, but allow constants denoting individuals and nullary predicate symbols (that is, propositional symbols). We use the *existence predicate* $\mathsf{E}x$ as an abbreviation for $\exists y(y = x)$ and may be read as "x exists". Note that $\mathsf{E}x$ is a formula in the empty signature.

A *frame* is an ordered tuple (W, R) with W a non-empty set of states and R a binary relation on R. A *skeleton* is an ordered triple (W, R, D), with (W, R) a frame and D a function with domain W assigning to each state $s \in W$, a non-empty set D_s ("s's domain of individuals"). Let D denote $\bigcup_{s \in W} D_s$. A *model* is an ordered quadruple (W, R, D, I) with (W, R, D) a skeleton and I a (valuation) function defined over the non–logical constants of the language such that $I(c) \in \mathsf{D}$ for each individual constant c and $I(P) \subseteq W \times \mathsf{D} \times \ldots \times \mathsf{D}$ (n

D's) for each n-place predicate symbol P ($n \geq 0$). Note that the individual constants are treated as *rigid designators*. We say that a skeleton or model has *constant domains* if $D_s = D_t$ for all $s, t \in W$.

Now for the satisfaction definition. When considering models without constant domains a number of choices have to be made concerning the definition of the quantifiers and the modalities. We adopt the choices made by Fine [5]. That is, we shall relativise quantification to objects existing in the current state of evaluation. On the other hand, again following Fine, we don't relativise the meaning of the modalities to objects existing in successor states. Some authors (for example, [11]) insist on this second relativisation, but the strategy we adopt is more general, for we can mimic the relativized meaning in the object language with the help of the existence predicate. For example, *She is always angry* would be translated by $\Box(\mathsf{E}(x) \rightarrow \mathsf{Angry}(x))$ if this sentence is taken to mean that she is angry at all points of time during her existence.

An assignment is a function from the variables to D. For A an assignment and t a term, we let $[A, I](t)$ denote $A(t)$ if t is a variable, and $I(t)$ if it is a constant.

Let $\mathfrak{M} = (W, R, D, I)$ be a model and ϕ a formula. $\mathfrak{M} \models \phi\ [s, A]$ (*ϕ holds in $\mathfrak{M}$ at s under A*) is defined inductively, for all states $s \in W$, and assignments A, according to the clauses:

1. $\mathfrak{M} \models Pt_1 \ldots t_n\ [s, A]$ iff $(s, [A, I](t_1), \ldots, [A, I](t_n)) \in I(P)$,
2. $\mathfrak{M} \models t_1 = t_2\ [s, A]$ iff $[A, I](t_1) = [A, I](t_2)$,
3. $\mathfrak{M} \models \neg\phi\ [s, A]$ iff $\mathfrak{M} \not\models \phi\ [s, A]$,
4. $\mathfrak{M} \models \phi \wedge \psi\ [s, A]$ iff $\mathfrak{M} \models \phi\ [s, A]$ and $\mathfrak{M} \models \psi\ [s, A]$,
5. $\mathfrak{M} \models \Diamond\phi\ [s, A]$ iff there exists a $t \in W$ such that Rst and $\mathfrak{M} \models \phi\ [t, A]$,
6. $\mathfrak{M} \models \exists x\phi\ [s, A]$ iff there exists an assignment A' different from A at most in that $A'(x) \neq A(x)$, with $A'(x) \in D_s$ and $\mathfrak{M} \models \phi\ [s, A']$.

Let F be a class of skeletons, and ϕ a formula. We say that ϕ is F-valid (notation $\models_{\mathsf{F}} \phi$) if for every model $\mathfrak{M}$ on every skeleton from F, $\mathfrak{M} \models \phi\ [s, A]$ holds for every s and A. Validity of ϕ on the class of all skeletons is denoted by $\models \phi$, without subscript.

The language of quantified modal logic can be viewed as a subset of a two-sorted first-order logic. In [11], the standard translation of propositional modal logic to first-order logic is extended to languages of quantified modal logic. by defining a two-sorted first-order correspondence language with one sort for individuals and one for states. To be precise, just as in the propositional case, the correspondence language contains a distinguished binary predicate symbol R denoting a relation between elements of the state sort. Two extra devices are needed to cope with quantified modal logic. First, a binary relation E between the two sorts, with $E(s, d)$ meaning that $d \in D_s$, or "d exists in state s", is required. Second, the language has, for every n-place predicate symbol P

of the original quantified modal language, an $n+1$-place predicate symbol $\bar{P}$ denoting the predicate P relativized to states. The standard translation from propositional modal logic to first-order logic can then be extended to the case of quantified modal logic as follows. Let s be an arbitrary state variable, then

$$\begin{array}{lcl}
ST(Pt_1 \ldots t_n) & := & \bar{P}st_1 \ldots t_n \\
ST(t_i = t_j) & := & t_i = t_j \\
ST(\neg\phi) & := & \neg ST(\phi) \\
ST(\phi \wedge \psi) & := & ST(\phi) \wedge ST(\psi) \\
ST(\exists x\phi) & := & \exists x(Esx \wedge ST(\phi)) \\
ST(\Diamond\phi) & := & \exists t(Rst \wedge \exists s(s = t \wedge ST(\phi))).
\end{array}$$

This is essentially the translation given by van Benthem in [11], but adjusted to our non-relativized definition of $\Diamond$ and presented in the variable free format of Vardi [12]. The translation is truth preserving when we interpret the $n+1$-place relations $\bar{P}$ and the existence predicate in the obvious way on a model (W, R, D, I).

As is usual with first-order languages, the correspondence language can express the Robinson diagram of a model. Consider a model $\mathfrak{M} = (W, R, D, I)$ for a modal predicate language in signature L and let $(\mathfrak{M}, \bar{n}, \bar{c})$ be an expansion in which every state $s \in W$ is named by a constant in $\bar{n}$ and every individual in D by a constant in $\bar{c}$. Then $(\mathfrak{M}, \bar{n}, \bar{c})$ can be described up to isomorphism in the two-sorted correspondence language, as follows. The frame part (W, R) is described by the literals of the form $(\neg)Rnm$ and $(\neg)(n = m)$ true in $\mathfrak{M}$. The function D assigning to every $s \in W$ a set of individuals is determined by the literals of the form $(\neg)E(n, c)$ and the (negated) identities between elements in $\bar{c}$. The literals $(\neg)\bar{P}(n, c_1, \ldots, c_m)$ determine the valuation function I.

2 Hybrid logic

Hybrid logic is an extension of modal logic in which it is possible to name states and to assert that a formula is true at a named state. Hybrid logic uses three fundamental tools to do this: nominals, satisfaction operators, and the $\downarrow$-binder.

We shall now hybridize quantified modal logic. We do so in two steps. First we add nominals and satisfaction operators and show that this extension gives the modal object language the power to express Robinson diagrams. We then add the $\downarrow$ binder, and show that it enables us to 'bind out' nominals, thus opening the way for the interpolation result.

2.1 Nominals and satisfaction operators

Let NOM be a set of propositional symbols (that is, nullary predicate symbols) distinct from any propositional symbols already in the language. These special propositional symbols are called *nominals* and will be used to name states. Although nominals will be used as states, note that (like any propositional symbol) they are *formulas* of the language (the idea of using 'formulas as terms' is characteristic of hybrid logic and dates back to the pioneering work of Arthur Prior [10]). We make one other syntactic extension:

- if ϕ is a formula and $n \in \mathsf{NOM}$, then $@_n\phi$ is also a formula.

The intuitive meaning of $@_n\phi$ (pronounced: at n, ϕ) is that formula ϕ holds at the state named n.

Nominals name worlds as follows. Recall that an n-place predicate symbol P is interpreted as $I(P) \subseteq W \times \mathsf{D} \times \ldots \times \mathsf{D}$ (n D's). Hence nullary predicate symbols are interpreted by subsets of W (strictly speaking they are interpreted by sets of 1-tuples drawn from W, but it is convenient and harmless to ignore this distinction). Intuitively these are the sets of states at which the propositional symbol is true. And now for the key change: we insist that nominals always be interpreted by a singleton subset of W. Thus nominals 'name' the unique state they are true at. The following terminology will be helpful. In any model $\mathfrak{M}$, for any nominal n, $I(n)$ is a singleton set containing a state s. In what follows we shall call the unique state s in $I(n)$ the *denotation* of n in $\mathfrak{M}$.

As nominals are simply a special sort of propositional symbol constrained in the way they are interpreted, we don't need to add a new clause to the QML satisfaction definition to cover them. To cope with the satisfaction operators $@_n$, we add the following clause:

7. $\mathfrak{M} \models @_n\phi\ [s, A]$ iff $\mathfrak{M} \models \phi[\bar{n}, A]$, where $\bar{n}$ is the denotation of n in $\mathfrak{M}$.

The addition of nominals and satisfaction operators increases the expressive power at our disposal in a fundamental way, for we can now mimic the description of the diagram in the first-order correspondence language inside the hybrid object language (expanded with the appropriate new nominals). We do so as follows:

$$\begin{array}{rcl}
(\mathfrak{M}, \bar{n}, \bar{c}) \models R(n_i, n_j) & \iff & (\mathfrak{M}, \bar{n}, \bar{c}) \models @_{n_i}\Diamond n_j \\
(\mathfrak{M}, \bar{n}, \bar{c}) \models n_i = n_j & \iff & (\mathfrak{M}, \bar{n}, \bar{c}) \models @_{n_i} n_j \\
(\mathfrak{M}, \bar{n}, \bar{c}) \models E(n_i, c_j) & \iff & (\mathfrak{M}, \bar{n}, \bar{c}) \models @_{n_i}\exists x(x = c_j) \\
(\mathfrak{M}, \bar{n}, \bar{c}) \models P(n_i, c_1, \ldots, c_m) & \iff & (\mathfrak{M}, \bar{n}, \bar{c}) \models @_{n_i} P(c_1, \ldots c_m).
\end{array}$$

Thus we have a one-to-one fit between syntax and semantics in the form of Robinson diagrams. However this fit was obtained in a language expanded with extra nominals, thus we have moved to a different signature.

To prove general interpolation results, we need this kind of close fit between the syntax and semantics, but we need it in the same signature. That is, we need to be able to name and refer to states in a signature independent manner. In first-order logic, this is done with the help of variables and the quantifiers.

Now, it is certainly possible to add explicit classical quantification over nominals (and in fact Prior's pioneering systems of hybrid logic allowed this). But an interpolation result based on such an extension would be rather uninteresting, as adding full classical quantication over nominals gives us the power of the first-order correspondence language. We would like a general interpolation result in a *restricted* extension that is modally natural. Such an extension has been the focus of attention in the recent work on hybrid logic. The basic idea is this: instead of making use of full classical quantifiers over states, we add a 'local' binder called $\downarrow$.

2.2 The $\downarrow$ binder

Let SVAR be a set of variables disjoint from the variables we already have. We usually use w, v for elements of SVAR, to indicate that they denote states. We expand the language of quantified modal logic as follows:

- every $w \in \mathsf{SVAR}$ is a formula,
- if ϕ is a formula and $w \in \mathsf{SVAR}$, then $@_w\phi$ is a formula,
- if ϕ is a formula and $w \in \mathsf{SVAR}$, then $\downarrow w.\phi$ is a formula.

The intuitive meaning of $\downarrow w.\phi$ at state s is best explained dynamically: bind the variable w to the current state s and continue evaluating ϕ at s under the new assignment. (That is, $\downarrow$ creates a name for the 'here and now' — it is an intrinsically local binder.) Formally, if $\mathfrak{M}$ is a model and A a sorted assignment which now also assigns variables from SVAR to elements in W, then:

8. $\mathfrak{M} \models w\ [s, A]$ iff $A(w) = s$,
9. $\mathfrak{M} \models @_w\phi\ [s, A]$ iff $\mathfrak{M} \models \phi\ [A(w), A]$,
10. $\mathfrak{M} \models \downarrow w.\phi\ [s, A]$ iff $\mathfrak{M} \models \phi\ [s, A^w_s]$,

where A^w_s is the assignment which differs from A only in that $A^w_s(w) = s$.

For clarity, we provide the extra clauses needed for a standard translation. The clauses show how $\downarrow$ and $@$ function as substitution operators. For $w \in \mathsf{SVAR}$ and $n \in \mathsf{SVAR} \cup \mathsf{NOM}$,

$$\begin{array}{lll} ST(\downarrow w.\varphi) & := & ST(\varphi)[w/s] \\ ST(@_n\varphi) & := & ST(\varphi)[s/n] \\ ST(n) & := & s = n. \end{array}$$

Note that $@_w$ has no binding force and $\downarrow w$ binds variables just as in first-order logic. This is reflected in their standard translations. For instance, $ST(@_w\Diamond w)$ is equivalent to Rww and $ST(\downarrow w.\Diamond w)$ to $\exists w(s = w \wedge Rww)$.

Let the language of quantified hybrid logic (QHL) be the expansion of quantified modal logic with nominals, variables, and the operators $\downarrow$ and $@$. A formula is called a *sentence* if all variables (whether they range over objects or states) are bound. Note that the standard translation of a quantified hybrid logic sentence ϕ need not be a first-order sentence. In general, for ϕ a sentence, $ST(\phi)$ is a first-order formula in one free variable. An exception is when ϕ is a Boolean combination of sentences of the form $@_n\psi$ for $n \in \mathsf{NOM}$. Such sentences are called *closed sentences*. For sentences ϕ, $\mathfrak{M} \models \phi\ [s]$ abbreviates $\mathfrak{M} \models \phi\ [s, A]$ for all A. For closed sentences ϕ, we use $\mathfrak{M} \models \phi$ in a similar manner.

The following result is easy, but crucial for interpolation. It shows that we can use $\downarrow$ to bind out nominals, thus enabling us to capture Robinson diagrams without changing signature.

Lemma 2.1. *Let $\bar{n} = (n_1, \ldots, n_l)$ be a sequence of nominals and let $\phi, \theta(\bar{n})$ be quantified hybrid formulas such that the elements in $\bar{n}$ do not occur in ϕ. Let $\theta(\bar{w})$ be $\theta(\bar{n})$ in which each n_i is replaced throughout by w_i. Then*
(i) $\models \phi \to \theta(\bar{n})$ implies $\models \phi \to \downarrow w_1 \ldots \downarrow w_l.\theta(\bar{w})$, and
(ii) $\models \theta(\bar{n}) \to \phi$ implies $\models \downarrow w_1 \ldots \downarrow w_l.\theta(\bar{w}) \to \phi$.

For (i), if $\phi \to \theta(\bar{n})$ is valid and the n_i do not occur in ϕ, then it is valid for every choice of denotation for the n_i, and in particular when they are all assigned to the same current state of evaluation. This is just what is expressed by $\downarrow w_1 \ldots \downarrow w_l.\theta(\bar{w})$.

For (ii), if $\downarrow w_1 \ldots \downarrow w_l.\theta(\bar{w}) \to \phi$ is not valid, then there exists a model $\mathfrak{M}$, an assignment A, and a state s such that $\downarrow w_1 \ldots \downarrow w_l.\theta(\bar{w})$ holds at s under A, but ϕ does not. Change $\mathfrak{M}$ into $\mathfrak{M}'$ by only changing the valuation of the nominals $\bar{n}$ such that for all n_i, $I'(n_i) = s$. Then $\mathfrak{M}' \models \theta(\bar{n})[s, A]$, and, as the n_i do not occur in ϕ, still $\mathfrak{M}' \not\models \phi[s, A]$. Thus $\mathfrak{M}'$ is the required counterexample to $\models \downarrow w_1 \ldots \downarrow w_l.\theta(\bar{w}) \to \phi$.

3 Fine's counterexamples

We review two of the counterexamples from Fine [5] and show how easy it is to express interpolants and explicit definitions in quantified hybrid logic.

For quantified **S5** with varying domains, Fine provides the following counterexample to interpolation:

$$\begin{array}{rcl} \phi & = & p \wedge \Box \forall x \Box (p \to \mathsf{E}x) \\ \psi & = & q \to \Box \forall x \Diamond (q \wedge \mathsf{E}x). \end{array}$$

Here p and q are propositional constants. Let CON_s abbreviate

$$\text{for any } s' \text{ such that } Rss',\ D_{s'} \subseteq D_s.$$

Thus CON_s expresses that the domains are contracting from s. It is easy to see that under the assumption of a symmetric accessibility relation R, $\mathfrak{M}, s \models \phi$ implies CON_s and that CON_s implies $\mathfrak{M}, s \models \psi$. Thus $\phi \to \psi$ is a theorem in quantified **S5** and the required interpolant should express CON_s in the empty signature. However Fine shows that such an interpolant does not exist in quantified **S5**. On the other hand, the following hybrid formula (in the empty signature) expresses that domains are contracting, and thus is an interpolant:

$$\downarrow w. \Box \forall x @_w \mathsf{E}x.$$

The ability to name the current state using $\downarrow w$, and to refer back to it using $@_w$, lets us find an interpolant in the common language, something which was not possible before.

For quantified **S5** with constant domains, Fine provides this counterexample to the Beth Definability Theorem: let T be the theory consisting of the following two formulas

$$p \to \Diamond\forall x(Fx \to \Box(p \to \neg Fx))$$
$$\neg p \to \Box\exists x(Fx \wedge \Box(\neg p \to Fx)).$$

Let $\mathfrak{M}$ be a model, and s any state in $\mathfrak{M}$. Let $\bar{F}_s$ be $\{d \in \mathsf{D} \mid \mathfrak{M}, s \models F(d)\}$. Fine shows that $\mathfrak{M}, s \models p$ iff $\bar{F}_s$ is disjoint from $\bar{F}_{s'}$ for some s' accessible from s. For this implicit definition of p there does not exist an explicit one in the language of quantified **S5**. But in quantified hybrid logic it is easily expressible. In any $\mathfrak{M}$, $\bar{F}_s$ is disjoint from $\bar{F}_{s'}$ for some s' accessible from s iff

$$\mathfrak{M}, s \models \downarrow w.\Diamond\forall x(Fx \to @_w \neg Fx).$$

Again, naming the state of evaluation and referring back to it enables us to express the implicit definition in the object language.

4 Expressive power of quantified hybrid logic

In order to state the interpolation theorem in its full generality, we need to introduce a few notions.

The *bounded fragment* of first-order logic is the set of formulas in one free variable generated by the following grammar:

$$\varphi = Rxy \mid x = y \mid \neg\varphi \mid \varphi \wedge \varphi' \mid \exists x(Ryx \wedge \varphi) \mid \forall x(Ryx \to \varphi) \text{ (for } x \neq y),$$

where x and y are arbitrary variables.

A class of Kripke frames or skeletons F is said to be *bounded fragment definable* if F consists of all frames (skeletons) satisfying $\forall x Cx$, for Cx some formula in the bounded fragment.

Many well known classes of Kripke frames are bounded fragment definable. For instance, any class defined by a combination of the following properties:

reflexivity	Rxx
symmetry	$\forall y(Rxy \to Ryx)$
transitivity	$\forall y(Rxy \to \forall z(Ryz \to Rxz))$
confluence	$\forall y(Rxy \to \forall z(Rxz \to \exists w(Ryw \wedge Rzw)))$
non-branching	$\forall y(Rxy \to \forall z(Rxz \to (z = y \vee Ryz \vee Rzy)))$.

It turns out that membership of a bounded fragment definable class of skeletons is expressible by means of a hybrid sentence in the empty signature.

Theorem 4.1. *Let F be a bounded fragment definable class of skeletons. Then there exists a quantified hybrid sentence δ_{F} constructed only from logical constants such that for every model $\mathfrak{M} = (W, R, D, I)$,*

$$\mathfrak{M} \models \delta_{\mathsf{F}} \textit{ if and only if } (W, R, D) \textit{ is a member of the class } \mathsf{F}.$$

The theorem follows from Corollary 3.2 in [1]. There it is shown that for every bounded formula Cx it holds that for every model $\mathfrak{M}$ and every state s,

$$\mathfrak{M} \models C(s) \iff \mathfrak{M} \models \downarrow x.HT(Cx)\ [s],$$

in which HT is the following translation function from bounded formulas to propositional hybrid formulas:

$$\begin{array}{lcl} HT(Rxy) & = & @_x \Diamond y. \\ HT(x = y) & = & @_x y. \\ HT(\neg\varphi) & = & \neg HT(\varphi). \\ HT(\varphi \wedge \psi) & = & HT(\varphi) \wedge HT(\psi). \\ HT(\exists x(Ryx \wedge \varphi)) & = & @_y \Diamond \downarrow x.HT(\varphi). \end{array}$$

Note that the translation of a formula from the bounded fragment is a hybrid formula build up from logical constants only, as required.

Not only conditions on the accessibility relation can be enforced by hybrid sentences in the empty signature — important conditions on the domain can be expressed as well. Let F be a class of skeletons. Let $\mathfrak{M}$ be a model. We say that the class F or the model $\mathfrak{M}$ has varying domains if there is no restriction on the domains, expanding domains if Rst implies $D_s \subseteq D_t$, contracting domains if Rst implies $D_t \subseteq D_s$, and constant domains if $D_s = D_t$ for all states s, t.

The following result is easy to prove:

Proposition 4.2. *A model $\mathfrak{M}$ has contracting domains iff*

$$\mathfrak{M} \models \downarrow w.\Box\forall x @_w \exists y(x = y)\ [s],\ \textit{for all } s \in W,$$

and a model $\mathfrak{M}$ has expanding domains iff

$$\mathfrak{M} \models \forall x \Box \exists y(x = y)\ [s],\ \textit{for all } s \in W.$$

The formula defining contracting domains was discussed in Section 3. The expanding domain formula is standard in first-order modal logic. Obviously constant domain models can be defined by the conjunction of these two formulas.

5 Interpolation for first-order hybrid logic

We come to the main result of the paper. Let F be a class of skeletons. We say that the quantified hybrid logic of F enjoys interpolation if for all sentences ϕ, ψ, $\models_{\mathsf{F}} \phi \to \psi$ implies the existence of an interpolant θ such that $\models_{\mathsf{F}} \phi \to \theta$ and $\models_{\mathsf{F}} \theta \to \psi$, and θ is constructed from the constants and predicate symbols occurring both in ϕ and ψ. (Recall that ordinary propositional variables and nominals are nullary predicate symbols, so these are covered by this definition.)

Theorem 5.1 (Interpolation Theorem). *Let F be a bounded fragment definable class of skeletons, with either varying, expanding, contracting or constant domains. Then the quantified hybrid logic of F enjoys interpolation.*

We obtain as a corollary that all counterexamples explicitly mentioned in [5] are repaired. In particular, quantified hybrid **K**, **T**, **D**, **S4**, **S4.3** and **S5** with varying, expanding, contracting or constant domains all enjoy interpolation.

The main work in the proof of the theorem is to show the interpolation theorem for the quantified hybrid logic of the class of *all* skeletons. We shall formulate this result as a lemma, show how the theorem follows from the lemma, and then proceed with a proof of the lemma.

Lemma 5.2. *The interpolation theorem holds for the quantified hybrid logic of the class of all skeletons.*

Proof of Theorem 5.1. Let F be a class of skeletons as in the theorem. By Theorems 4.1 and Proposition 4.2 there exists a quantified hybrid sentence δ_{F} constructed from logical constants such that for every skeleton (W, R, D),

$$(W, R, D) \models \delta_{\mathsf{F}}\ [s] \text{ for all } s \text{ if and only if } (W, R, D) \text{ in the class } \mathsf{F}.$$

Let K denote the class of all skeletons. For sets of sentences $\Gamma \cup \{\phi\}$, $\Gamma \models_{\mathsf{K}} \phi$ denotes the local consequence relation from modal logic. That is, $\Gamma \models_{\mathsf{K}} \phi$ holds if for every model $\mathfrak{M}$ over a skeleton in K, for all states s, $\mathfrak{M} \models \gamma\ [s]$, for all $\gamma \in \Gamma$ only if $\mathfrak{M} \models \phi\ [s]$.

Now assume $\models_{\mathsf{F}} \phi \to \psi$. Then $\{\Box^n \delta_{\mathsf{F}} \mid n \in \omega\} \models_{\mathsf{K}} \phi \to \psi$, with $\Box^n$ denoting a sequence of n boxes. This holds because quantified hybrid formulas are preserved under generated submodels [1]. But then by compactness and the deduction theorem $\models_{\mathsf{K}} (\Delta \wedge \phi) \to \psi$, for Δ a conjunction of sentences from $\{\Box^n \delta_{\mathsf{F}} \mid n \in \omega\}$. Thus by Lemma 5.2, there exists an interpolant θ such that

$$\models_{\mathsf{K}} (\Delta \wedge \phi) \to \theta \text{ and } \models_{\mathsf{K}} \theta \to \psi.$$

Now θ is in the common language of ϕ and ψ, as Δ only contains logical symbols. Obviously $\models_{\mathsf{K}} \theta \to \psi$ implies $\models_{\mathsf{F}} \theta \to \psi$, as F is a subclass of K. On a model $\mathfrak{M}$ based on a skeleton from F, δ_{F} holds at every state, thus Δ holds at every state, and $\models_{\mathsf{K}} (\Delta \wedge \phi) \to \theta$ implies $\models_{\mathsf{F}} \phi \to \theta$. Thus θ is the required interpolant. QED

We will now prove Lemma 5.2. First observe the following. $\models \phi \to \psi$ implies $\models ST(\phi) \to ST(\psi)$, which is a first-order statement. Thus there exists an interpolant θ by the interpolation theorem for first-order logic. Now θ may contain the extra predicates R and E and in general there is no guarantee that $\theta = ST(\theta')$ for some hybrid formula θ'. Thus the standard translation plus the interpolation theorem do not provide sufficient information to obtain our desired result. Instead we will closely inspect the *proof* of the interpolation theorem in first-order logic and see that we added just enough expressive power to quantified modal logic to find the interpolant in the right language.

Proof of Lemma 5.2. Let ϕ and ψ be quantified hybrid sentences. Without loss of generality we may assume that ϕ and ψ are closed sentences. (Suppose

ϕ and ψ are just sentences. Let n be a nominal not occurring in ϕ and ψ. If $\models \phi \to \psi$, then also $\models @_n\phi \to @_n\psi$. Let θ be an interpolant for $@_n\phi \to @_n\psi$. As n does not occur in ϕ or ψ, $\downarrow w.\theta([n/w])$ is an interpolant for $\models \phi \to \psi$.)

We will redo the proof from [4] which builds up a model using fresh constants. As input we take the standard translations of ϕ and ψ and assume that there is no standard translation of a quantified hybrid formula which is an interpolant. To enhance readability however, we will use the easier modal notation instead of the standard translations of hybrid first-order sentences.

We assume that there is no interpolant for $\phi \to \psi$ and show that $\not\models \phi \to \psi$ by constructing a model for $\phi \wedge \neg\psi$. Let $\mathcal{L}_\phi$ ($\mathcal{L}_\psi$) be the language of all symbols occurring in ϕ (ψ) and set

$$\mathcal{L} = \mathcal{L}_\phi \cup \mathcal{L}_\psi \text{ and } \mathcal{L}_{\phi\psi} = \mathcal{L}_\phi \cap \mathcal{L}_\psi.$$

Form an expansion $\mathcal{L}'$ of $\mathcal{L}$ by adding countable sets C and N of new constant symbols and nominals, respectively and define $\mathcal{L}'_\phi$, $\mathcal{L}'_\psi$ and $\mathcal{L}'_{\phi\psi}$ accordingly

A set of closed sentences is called a *theory*. Consider a pair of theories T in $\mathcal{L}'_\phi$ and U in $\mathcal{L}'_\psi$. A closed sentence θ of $\mathcal{L}'_{\phi\psi}$ is said to *separate* T and U if

$$T \models \theta \text{ and } U \models \neg\theta.$$

T and U are said to be *inseparable* if no closed $\mathcal{L}'_{\phi\psi}$ sentence separates them.

The downarrow binder is crucial in the first step of the proof in which it is shown that

$$\{\phi\} \text{ and } \{\neg\psi\} \text{ are inseparable.}$$

Indeed if $\theta(c_1 \dots c_l, n_1 \dots n_k)$ separates $\{\phi\}$ and $\{\neg\psi\}$ and $x_1 \dots x_l$ and $w_1 \dots w_k$ are variables not occurring in θ, then by Lemma 2.1 and standard first-order reasoning, $\downarrow w_1 \dots \downarrow w_k.\forall x_1 \dots \forall x_l \theta(x_1 \dots x_l, w_1 \dots w_k)$ is a Craig interpolant of ϕ and ψ, contrary to our assumption.

Now we construct in a step by step fashion two inseparable Hintikka sets $T_\omega \supseteq \{\phi\}$ and $U_\omega \supseteq \{\neg\psi\}$ containing closed sentences in the quantified hybrid languages $\mathcal{L}'_\phi$ and $\mathcal{L}'_\psi$, respectively. The crucial property of Hintikka sets is that they are closed under Henkin witnesses for the existential quantifiers. If we standard translate quantified hybrid formulas we get two types of existentially quantified formulas $\exists x\sigma(x)$. In the first type, x is an ordinary first order variable. In this case, if $\exists x\sigma(x) \in T_\omega$ we ensure that also $\sigma(c) \in T_\omega$, for some $c \in C$. This can be done as usual in first-order logic. In the second type, x ranges over the state sort. Then it must be the standard translation of the closed sentence $@_n\Diamond\phi$, whence is of the form

$$\exists t(Rnt \wedge \exists s(t = s \wedge ST(\phi))).$$

The novelty is that we can also witness these existential formulas in the hybrid formalism. This existential formula is witnessed in the hybrid formalism using a new nominal m by the closed sentence

$$@_n\Diamond m \wedge @_m\phi.$$

The simultaneous construction of T_ω and U_ω from [4] can be completed in this way.

Now the final argument is standard. As the two theories T_ω and U_ω are both Hintikka sets, they each have a model whose Robinson diagram is part of the respective theory.

The Robinson diagram of the skeleton part of the two models only contains closed sentences of the form $@_n\Diamond m$, $@_n m$ and $@_n \exists x(x = c)$, which are all in the common language $\mathcal{L}'_{\phi\psi}$. Thus by inseparability the skeletons of the two models are isomorphic. Moreover, also by inseparability, the two models agree on the interpretation of all symbols in the common language $\mathcal{L}_{\phi\psi}$. Thus we can expand the model for T_ω to a model for U_ω as well. Since $\phi \in T_\omega$ and $\neg\psi \in U_\omega$, we constructed a model for $\phi \wedge \neg\psi$, as desired. QED

6 Concluding remarks

Fine concludes his paper as follows:

> *Nor is it clear to what extent the failures depend upon the expressive inadequacies of a modal language. 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.*

We think that the machinery of contemporary hybrid logic yields the kind of natural extension of orthodox modal logic that Fine speculated about. The hybrid extension is modally natural because the hybrid propositional sentences are precisely the first-order sentences which are invariant under generated submodels (the proof of this, along with a proof that such sentences are precisely those belonging to the bounded fragment, can be found in [1]). Sentences invariant under generated submodels are precisely the 'local' sentences — their truth value can only be affected by changes in *accessible* information, thus this fragment is intrinsically modal in spirit. Moreover the bounded fragment certainly falls short of the full classical language, thus the hybrid solution is non trivial.

It also seems to us that the hybrid extension is, in a sense, the minimal one which repairs the interpolation theorem. Adding $@_n$ and nominals alone is not sufficient, nor is a fixed finite number of variables, even in the presence of $\downarrow$ (see [1], Theorem 4.7).

Also worth stressing is the intrinsically first-order nature of what we have been doing (this will be apparent to any reader who has followed the essentially classical details of most of our proofs). Many writers on hybrid logic have stressed that it can be viewed as an attempt to marry the best features of modal logic and first-order correspondence language, and the general interpolation results proved in this paper are a testimony to the fruitfulness of the union. It is instructive to compare the strategy of the present paper with one recently explored by Fitting [6]. Fitting's approach to repairing interpolation is to enrich the underlying modal language with explicit quantification over ordinary propositional variables. But while this strategy is superficially similar to hybridization used in the present paper, there is crucial technical difference: propositional quantification is intrinsically second-order (after all, propositional variables can denote arbitrary subsets of frames). Thus Fitting's approach does not offer an easy way to 'modalise' classical model theoretic methods, and Fitting only shows in his paper how to repair interpolation failure in first-order **S5**.

References

[1] C. Areces, P. Blackburn, and M. Marx. Hybrid logics. Characterization, interpolation and complexity. Technical Report PP-1999–07, Institute for Logic, Language and Computation, University of Amsterdam, 1999. To appear in Journal of Symbolic Logic.

[2] P. Blackburn,M. de Rijke, and Y. Venema. *Modal Logic.* Cambridge University Press, 2001.

[3] P. Blackburn and J. Seligman. Hybrid languages. *Journal of Logic, Language and Information*, 4:251–272, 1995.

[4] C. Chang and H. Keisler. *Model Theory.* North-Holland, 1973.

[5] K. Fine. Failures of the interpolation lemma in quantified modal logic. *Journal of Symbolic Logic*, 44(2):201–206, 1979.

[6] M. Fitting. Interpolation for first-order S5. To appear in Journal of Symbolic Logic.

[7] M. Fitting and R. Mendelsohn. *First–Order Modal Logic.* Kluwer Academic Publishers, 1998.

[8] V. Goranko. Hierarchies of modal and temporal logics with reference pointers. *Journal of Logic, Language and Information*, 5:1–24, 1996.

[9] S. Passy and T. Tinchev. An essay in combinatory dynamic logic. *Information and Computation*, 93:263–332, 1991.

[10] A. Prior. *Past, Present and Future.* Oxford University Press, 1967.

[11] J. van Benthem. *Modal Logic and Classical Logic.* Bibliopolis, Naples, 1983.

[12] M. Vardi. Why is modal logic so robustly decidable? In *DIMACS Series in Discrete Mathematics and Theoretical Computer Science 31*, pages 149–184. American Math. Society, 1997.