HAJNAL ANDRÉKA, ISTVÁN NÉMETI and JOHAN VAN BENTHEM

MODAL LANGUAGES AND BOUNDED FRAGMENTS OF PREDICATE LOGIC

1. MODAL LOGIC AND CLASSICAL LOGIC

Modal Logic is traditionally concerned with the intensional operators "possibly" and "necessary", whose intuitive correspondence with the standard quantifiers "there exists" and "for all" comes out clearly in the usual Kripke semantics. This observation underlies the well-known translation from propositional modal logic with operators $\Diamond$ and $\Box$, possibly indexed, into the first-order language over possible worlds models (van Benthem 1976, 1984). In this way, modal formalisms correspond to *fragments* of a full first-order (or sometimes higher-order) language over these models, which are both expressively perspicuous and deductively tractable. In this paper, by the "modal fragment" of predicate logic we understand the set of all first-order formulas obtainable as translations of basic (poly-)modal formulas. As the modal fragment is merely a notational variant of the basic modal language, we will often refer to the two interchangeably. Basic modal logic shares several nice properties with full predicate logic, namely, finite axiomatizability, Craig Interpolation and Beth Definability, as well as model-theoretic preservation results such as the Łoś–Tarski Theorem characterizing those formulas that are preserved under taking submodels. In addition, basic modal logic has some nice properties not shared with predicate logic as a whole: e.g., its axiomatization does not need side conditions on free or bound variables – and most evidently: basic modal logic is *decidable*. We shall concentrate on this list in what follows, in the hope that it forms a representative sample. Our aim is to find natural fragments of predicate logic extending the modal one which inherit the above-mentioned nice properties. This quest has two virtues. It forces us to understand why basic modal logic has these nice properties. And it points the way to new insights concerning predicate logic itself. Note that this study takes place over the universe of all models, without special restrictions on accessibility relations. This is the domain of the "minimal modal logic", which

Journal of Philosophical Logic **27:** 217–274, 1998.

still serves as the "pure paradigm" in a rapidly expanding field of more expressive modal formalisms (Venema 1991, De Rijke 1993). Of course, one can also study the effects of special frame restrictions – but we must leave this issue for further investigation, except for some passing remarks.

What precisely are fragments of classical first-order logic showing "modal" behaviour? Perhaps the most influential answer is that of Gabbay 1981, which identifies them with so-called "finite-variable fragments", using only some fixed finite number of variables (free or bound). This view-point has been endorsed by many authors (cf. van Benthem 1991). We will investigate these fragments, and find that, illuminating and interesting though they are, they lack the required nice behaviour in our sense. (Several new negative results support this claim.) As a counterproposal, then, we define a large fragment of predicate logic characterized by its use of only bounded quantification. This so-called *guarded fragment* enjoys the above nice properties, including decidability, through an effectively bounded finite model property. (These are new results, obtained by generalizing notions and techniques from modal logic.) Moreover, its own internal finite variable hierarchy turns out to work well. Finally, we shall make another move. The above analogy works both ways. Modal operators are like quantifiers, but quantifiers are also like modal operators. This observation inspires a generalized semantics for first-order predicate logic with accessibility constraints on available assignments (cf. Németi 1986, 1992) which moves the earlier quantifier restrictions into the semantics. This provides a fresh look at the landscape of possible predicate logics, including candidates sharing various desirable features with basic modal logic – in particular, its decidability.

The organization of this paper is as follows. In Section 2 we recall the results and methods of basic modal logic which we intend to generalize to "nice" fragments later. We allow modalities of higher ranks too (binary, ternary, etcetera), and define the modal fragment of predicate logic accordingly. Most results in this section are known, whence it has a sketchy character. In Section 3 we study finite variable fragments in the spirit outlined above. In Section 4 we define bounded quantifier fragments and single out the one most central to our purposes. We investigate this "guarded fragment" and prove that it has all the desirable properties. To put this in perspective, in Section 5 we briefly discuss the "semantic" version of our approach, replacing syntactic bounds by restrictions on ranges of assignments in models. The account includes connections with cylindric algebra. Finally, Section 6 presents some further directions.

This paper is the first public version of a longer projected document – whose current working version is Andréka, van Benthem and Németi 1994A. Further off-spring of this Amsterdam–Budapest collaboration in the field of modal logic and universal algebra are Andréka, van Benthem and Németi 1993, Andréka, van Benthem and Németi 1994B.

2. Basic Modal Logic

2.1. *First-Order Translation*

Consider the basic propositional modal logic, in the language with Booleans $\neg$ and $\vee$ and modalities $\Box$ $\Diamond$. The following computable translation takes modal formulas ϕ to first-order formulas $\underline{\phi}$ with one free variable (standing for the "current world" of evaluation) recording their truth conditions on possible worlds models:

$$\begin{array}{llll} \underline{p} & Px & \underline{\neg\phi} & \neg\underline{\phi} \\ \underline{\phi \vee \psi} & \underline{\phi} \vee \underline{\psi} & \underline{\phi \vee \psi} & \underline{\phi}\,\&\,\underline{\psi} \\ \underline{\Diamond\phi} & \exists y(Rxy\,\&\,\underline{\phi}(y)) & \underline{\Box\phi} & \forall y(Rxy \to \underline{\phi}(y)) \end{array}$$

where y is some fresh individual variable in the last two clauses.

Here, that y is a fresh variable means that y does not occur in ϕ, while $\underline{\phi}(y)$ is obtained from $\underline{\phi}$ by replacing all free occurrences of x by y. This translation preserves truth, and so it gives various facts about modal logic for free, namely those properties of first-order logic which are inherited by all its fragments, such as the Löwenheim–Skolem Theorem and Compactness – or by all its decidable fragments, such as recursive enumerability of valid formulas. The embedding gives no specific axiomatization: more detailed analysis is needed for that (see below). Also, we do not get complex meta-properties that make existential claims. For example, consider Interpolation. If a modal formula ϕ implies another modal formula ψ, then, by the translation, some interpolant exists in the first-order language – but there is no guarantee that this interpolant is equivalent to a modal formula: we must work for this (see again below).

We call the above language "basic modal logic" because it contains only the usual unary modalities. Later, in Section 2.9, we consider more than one unary modality $\langle i\rangle$, $[i]$, referring to binary relations R_i, and several polyadic modalities, say binary $\Diamond\phi\psi$ referring to ternary accessibility relations. We will also call this basic modal logic, as the first-order translation is completely obvious from the above schema. Throughout, we shall not impose any constraint on accessibility relations – as is the case in the semantics of predicate logic – so in modal terminology, our sets of modal validities are the "minimal" ones, lying at the bottom of the lattice of modal logics.

2.2. *Invariance for Bisimulation*

The expressive power of the basic modal language with respect to classical logic is measured precisely by the following Invariance Theorem (van Benthem 1976, 1985):

THEOREM 2.2.1. *A first-order formula ϕ with one free variable x is equivalent to the translation of a modal formula iff it is invariant for bisimulation.*

Here, a *bisimulation* is a binary relation between the domains of two first-order models linking points with the same unary predicates P, corresponding to modal proposition letters p, and satisfying two "back-and-forth" or "zigzag clauses" with respect to relational R-successors. (More precisely, if x bisimulates y, and Rxz, then z bisimulates some u with Ryu, and vice versa. This is a kind of unbounded Ehrenfeucht Game with restricted choices of objects in each move – which has a natural generalization to the case with whole families of n-ary accessibility relations.) In the above theorem, the first-order formula may contain any other relation symbols, or equality, too. A formula ϕ with one free variable is *invariant for bisimulations* if, for any bisimulation, ϕ has the same truth value at linked objects in the two models. That modal formulas are invariant in this sense subsumes the usual textbook facts about preservation under generated submodels, disjoint unions and p-morphic images.

Proof of the Theorem. For later use, we sketch a proof of the Invariance Theorem. We say that a first-order formula ψ is modal if it is a translation of a modal formula. Thus, ψ has one free variable x. If $\mathbf{M}$ is a first-order model and a is an element of this model, then $\mathbf{M}, a \models \psi$ says that ψ is true in $\mathbf{M}$ when x is evaluated to a. We then say that $(\mathbf{M}, a)$ is a model for ψ, or that ψ is true in $(\mathbf{M}, a)$. Similarly for a set of modal formulas Σ instead of ψ. Last, we have the usual notion of consequence. $\Sigma \models \psi$ says that for any pair $(\mathbf{M}, a)$, $\mathbf{M}, a \models \Sigma$ implies $\mathbf{M}, a \models \psi$. (This *local version* of modal consequence is used throughout this paper.) Modal formulas are invariant, by a simple induction on their construction. The existential modality is taken care of, precisely, by the two zigzag clauses. Conversely, suppose that ϕ is an invariant first-order formula with one free variable. Let $\mathbf{mod}(\phi)$ be the set of all modal consequences of ϕ, i.e. $\{\psi|\psi$ is modal and $\phi \models \psi\}$. We prove the following:

CLAIM. $\mathbf{mod}(\phi) \models \phi$.

From this, by Compactness, ϕ is easily shown equivalent to some finite conjunction of its modal consequences. The proof of the Claim is as follows. Let $(\mathbf{M}, a)$ be any model for $\mathbf{mod}(\phi)$. Now consider the complete modal theory of a in $\mathbf{M}$ together with $\{\phi\}$. This set of formulas is finitely satisfiable, by a simple argument (using the fact that $\mathbf{mod}(\phi)$ holds at $(\mathbf{M}, a)$). By Compactness, it therefore has some model $(\mathbf{N}, b)$. Now, take any two ω-*saturated* elementary extensions $(\mathbf{M}^+, a)$ and $(\mathbf{N}^+, b)$ of $(\mathbf{M}, a)$ and $(\mathbf{N}, b)$, respectively. (These exist by a slight adaptation of a result from Chang and Keisler 1973.) We call elements u, v of $\mathbf{M}^+$, $\mathbf{N}^+$, respectively, "modally equivalent" if the same modal formulas are true in $(\mathbf{M}^+, u)$ and $(\mathbf{N}^+, v)$.

CLAIM. *The relation of modal equivalence is a bisimulation between the two models* $\mathbf{M}^+$ *and* $\mathbf{N}^+$*, which connects* a *with* b*.*

Here, of course, the key observation lies in the zigzag clauses. If some world u in $\mathbf{M}^+$ is modally equivalent with v in $\mathbf{N}^+$, and Rus holds, then the following set of formulas is finitely satisfiable in $(\mathbf{N}^+, v)$: $\{Rvx\}$ plus the full modal theory of s in $\mathbf{M}^+$. But then, by ω-saturation, some world t must exist satisfying all of this in $(\mathbf{N}^+, v)$: which is the required match for s. The converse argument is symmetric. Having thus proved the second claim, we return to the first, and clinch the argument by "diagram chasing". For a start, $\mathbf{N}, b \models \phi$, and hence $\mathbf{N}^+, b \models \phi$ (by elementary extension), whence $\mathbf{M}^+, a \models \phi$ (by bisimulation invariance), and so $\mathbf{M}, a \models \phi$ (passing to an elementary submodel). ■

This style of argument can be extended in many directions, by modulating the key connection between zigzag clauses and restricted quantifier patterns. More elaborate discussion of this result and its generalizations to richer modal languages is found in van Benthem and Bergstra 1995, De Rijke 1993. (These also provide connections with the work by Hennessy and Milner 1985 on modal process equivalences.)

2.3. *Decidability via Semantic Tableaus*

A pleasant feature of the modal formalism is a simple tableau method checking universal validity. Its has the usual decomposition rules for Boolean operators. (Modal sequents are of the form $\Sigma \Rightarrow \Delta$ with Σ, Δ finite sets of modal formulas. We take validity of sequents in the usual sense, as universal validity of the implication from the conjunction $\& \Sigma$ to the disjunction $\vee\Delta$.) Here are some samples.

$$\Sigma, \neg A \Rightarrow \Delta \quad \text{iff} \quad \Sigma \Rightarrow A, \Delta$$
$$\Sigma \Rightarrow A \,\&\, B, \Delta \quad \text{iff} \quad \Sigma \Rightarrow A, \Delta \quad \text{and} \quad \Sigma \Rightarrow B, \Delta.$$

In modal tableaus, the key rule is that for existential modalities – which are best treated in a bunch, when no further propositional reductions are possible:

true: $\Diamond\phi_1, \ldots, \Diamond\phi_n \quad \bullet_w \quad \Diamond\psi_1, \ldots, \Diamond\psi_m$: false
create new worlds $v_1, \ldots, v_n$ *with* $Rwv_i (1 \leqslant i \leqslant n)$
and start these with sequents $\phi_i \bullet_{vi} \psi_1, \ldots, \psi_m$.

Applying this method, we start out with any sequent, and in a finite number of steps, arrive at a tableau which is either "closed" or "open" in the usual sense. (We omit details of formulation.) This method is adequate for validity in the minimal modal logic.

THEOREM 2.3.1. *A modal sequent is valid iff it has a closed semantic tableau.*

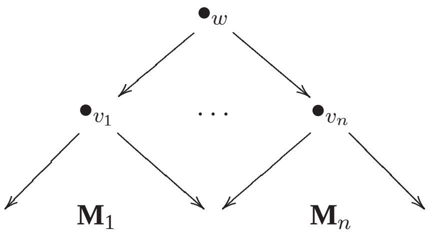
COROLLARY 2.3.2. *Modal universal validity is decidable, and basic modal logic has the finite model property. (That is, a modal formula fails in some model iff it fails in some finite model.)*

That tableaux are sound and complete for validity hinges on the above $\Diamond$-Rule. This is justified by a strong semantic equivalence, which may be proved independently. Let P, Q be disjoint sequences of proposition letters. Then we have:

$P, \Diamond\phi_1, \ldots, \Diamond\phi_n \models Q, \quad \Diamond\psi_1, \ldots, \Diamond\psi_m$ iff
for some $i (1 \leqslant i \leqslant n), \quad \phi_i \models \psi_1, \ldots, \psi_m$.

This assertion is immediate from right to left. The opposite part of its proof is as follows. Suppose that no assertion $\phi_i \models \psi_1, \ldots, \psi_m$ holds. Then, there exist models $\mathbf{M}_i, v_i \models \phi_i \,\&\, \neg\psi_1 \,\&\, \cdots \,\&\, \neg\psi_m (1 \leqslant i \leqslant n)$. Now apply a well-known modal semantic construction of "joint rooting" to produce a counter-example for the left-hand sequent:

any family of models $\mathbf{M}_i,\ v_i$ *with*
$v_i \models \phi_i \,\&\, \neg\psi_1 \,\&\, \cdots \,\&\, \neg\psi_m (1 \leqslant i \leqslant n)$
can be "glued disjointly" under one new common root:



The $\mathbf{M}_i$ lie embedded as generated submodels (the identity relation is a bisimulation), whence no truth values change for modal formulas in their roots (our reduction depends on bisimulation invariance) – so that the new top node w will verify $\Diamond\phi_1 \,\&\, \cdots \,\&\, \Diamond\phi_n \,\&\, \neg\Diamond\psi_1 \,\&\, \cdots \,\&\, \neg\Diamond\psi_m$, thereby refuting the top sequent. Compare this decomposition with the situation in full predicate logic, where no similar reduction via single instantiation of existential quantifiers suffices. We can even say a bit more. The corollary follows since all tableau rules decrease formula complexity of sequents (even though they may temporarily increase the number of parallel tasks). Also, open tableaus give rise to finite countermodels. (Incidentally, the finite model property may also be shown directly via these reduction arguments, without invoking tableaus.) In particular, the above rooting takes finite models to finite models. We shall return to this quantifier decomposition in Section 4, extending these ideas to larger "loose" decidable fragments of predicate logic.

2.4. *Proof Theory via Sequent Calculus*

Another way of describing modal validity is proof-theoretic. Read bottom-up, tableau rules become introduction rules in the "Minimal Modal Logic" consisting of a Gentzen-style calculus of sequents (cf. Fitting 1993), with axioms

$\Sigma \Rightarrow \Delta$ *with* $\Sigma \cap \Delta$ *nonempty.*

The following logical introduction rules are involved:

$$\frac{\Sigma, A \Rightarrow \Delta}{\Sigma \Rightarrow \neg A, \Delta} \qquad \frac{\Sigma \Rightarrow A, \Delta}{\Sigma, \neg A \Rightarrow \Delta}$$

$$\frac{\Sigma, A, B \Rightarrow \Delta}{\Sigma, A \,\&\, B \Rightarrow \Delta} \qquad \frac{\Sigma \Rightarrow A, \Delta \qquad \Sigma \Rightarrow B, \Delta}{\Sigma \Rightarrow A \,\&\, B, \Delta}$$

$$\frac{A \Rightarrow B_1, \ldots, B_m}{\Diamond A \Rightarrow \Diamond B_1, \ldots, \Diamond B_m} \ (\textit{the part "}B_1, \ldots, B_m\textit{" may be emtpy}),$$

the rules for $\vee$ and $\Box$ are analogous, and are omitted here.

Moreover, this calculus has two structural rules of

Permutation from $\Sigma, A, B, \Sigma' \Rightarrow \Delta$ to $\Sigma, B, A, \Sigma' \Rightarrow \Delta$,
from $\Sigma \Rightarrow \Delta, A, B, \Delta'$ to $\Sigma \Rightarrow \Delta, B, A, \Delta'$,

Monotonicity from $\Sigma \Rightarrow \Delta$ to $\Sigma', \Sigma \Rightarrow \Delta, \Delta'$.

These are needed to get the exact correspondence with closed semantic tableaus right. Note that the classical structural rule of *Contraction* is redundant for the completeness proof. (It deduces $\Sigma, A \Rightarrow \Delta$ from $\Sigma, A, A \Rightarrow \Delta$.) In classical tableaus or sequent proofs for predicate logic, this rule ensures that false existential (and true universal) formulas can produce as many substitution instances as are required for the argument. With modal formulas, however, no such unbounded iteration is needed: we did all that is needed in one fell swoop. Thus, our calculus involves no shortening rules, and the proof search space is finite. (In a sense, then, at least as far as quantification is concerned, "linear logic" is already complete for modal fragments of predicate logic.) This observation suggests yet another modal perspective on decidable fragments of predicate logic. For which of these is the standard first-order sequent calculus without Contraction (or with only effectively bounded calls to Contraction) semantically complete? We shall not pursue this proof-theoretic line in our paper, but it would be of interest to understand its systematic relation to our semantic analysis.

2.5. *Interpolation*

Except for its decidability and finite model property, which deviate from classical predicate logic, basic modal logic shares most central meta-properties with the latter. One important example is Interpolation:

THEOREM 2.5.1. *Let $\phi \models \psi$, with ϕ, ψ modal formulas. Then there exists a modal formula α whose proposition letters occur in both ϕ and ψ such that $\phi \models \alpha \models \psi$.*

Proof. We outline two proofs here, illustrating the two perspectives at work.

Proof-theoretic Argument ("Tracing a Sequent Derivation")
Induction on derivations in the Gentzen calculus of Section 2.4. It is convenient to work only with formulas rewritten to the special format $(\neg)$atom, &, $\vee$, $\Diamond$, $\Box$, $\bot$, T (Cf. Schütte 1962, Roorda 1991 for this

technique.) The single axiom case is clear, and one constructs interpolants inductively via the successive rules in a derivation. ■

Model-theoretic Argument ("Amalgamation via a Bisimulation")
Let $L_{\phi\psi}$ be the joint language of ϕ and ψ. Consider the set $\mathbf{cons}_{\phi\psi}(\phi)$ of all modal consequences of ϕ in this language. We prove the following:

CLAIM. $\mathbf{cons}_{\phi\psi}(\phi) \models \psi$.

By Compactness, then, some finite conjunction of formulas in $\mathbf{cons}_{\phi\psi}(\phi)$ implies ψ (and is implied by ϕ). To prove the claim, let $(\mathbf{M}, a)$ be any L_ψ-model verifying $\mathbf{cons}_{\phi\psi}(\phi)$. We must show that $\mathbf{M}, a \models \psi$. First, by a routine argument, the modal $L_{\phi\psi}$-theory of $(\mathbf{M}, a)$ is finitely satisfiable together with $\{\phi\}$. By Compactness again, there is an L_ϕ-model $\mathbf{N}, b \models \phi$ with the same modal $L_{\phi\psi}$-theory as $(\mathbf{M}, a)$. Next, as in the proof of the Invariance Theorem, we can pass to ω-saturated models, without loss of generality. By that earlier argument, there is an $L_{\phi\psi}$-bisimulation $\equiv$ between the two models which connects a to b. (The language subscript reminds us that $\equiv$ only needs to respect proposition letters shared by ϕ and ψ.) Now, we construct a new product model **MN** out of these two bisimulating ones, which will be a kind of joint unraveling under bisimulation. Its worlds are finite sequences of pairs $\langle (a_1, b_1), \ldots, (a_k, b_k) \rangle$, where always $a_i \equiv b_i$, and each world a_{i+1} must be an R-successor of a_i – and likewise for the sequence of worlds b_i – for $1 \leqslant i < k$. Now, consider the two natural projections from the final pairs of such sequences, one going to **M** and the other to **N**. Along these, we can lift the valuation for all proposition letters in L_ψ–$L_{\phi\psi}$ from **M**, and that for L_ϕ–$L_{\phi\psi}$ from **N**. The result is the desired model **MN** for the joint language $L_\phi \cup L_\psi$, whose two projections to **M** and **N** have now become L_{ψ^-} (L_{ϕ^-}) bisimulations. But then, we can argue "clockwise". First, $\mathbf{N}, b \models \phi$, whence $\mathbf{MN}, \langle (a, b) \rangle \models \phi$ (by bisimulation). Next, since $\phi \models \psi$, we also get $\mathbf{MN}, \langle (a, b) \rangle \models \psi$. Then finally, $\mathbf{M}, a \models \psi$ (again by invariance for bisimulation). ■

REMARK. This model-theoretic interpolation argument was stated in correspondence between the authors in 1992. In the meantime, more elaborate versions have appeared independently in Visser *et al.* 1995, Marx 1995, Németi 1994, Madarasz 1995, clarifying its general algebraic and model-theoretic background. The general version of this interpolation theorem for minimal modal logics with an arbitrary number of (polyadic) modalities follows at once from the results in Németi 1985 on the Strong

Amalgamation Property for Boolean algebras with sets of operators of arbitrary ranks.

REMARK. Modal Unraveling and Concrete Representation.
Behind the preceding proof, as well as others to come, lies the well-known fact that each possible worlds model $(\mathbf{M}, a)$ bisimulates with a so-called "unraveled" or "unwound" model consisting of finite sequences of objects (namely, finite sequences of worlds that all R-succeed one another), in which the accessibility relation is end extension by one additional object. This may be regarded as a semantic "normal form" in which abstract accessibility has been replaced by one concrete uniform set-theoretic relation. (Marx 1995 has more elegant concrete representations of this kind for general modal logics, following the program of Henkin–Monk–Tarski 1985.)

2.6. *Model Theory and Preservation*

Much of classical Model Theory holds for the modal fragment. One good example is the Łoś–Tarski Theorem, stated in its upward version. Let **M** be a possible worlds model (i.e., **M** has universe M, a binary accessibility relation R on M, and unary predicates $P \subseteq M$ for proposition letters p). We say that **N** *extends* **M** if **M** is a submodel of **N** in the usual sense. We say that the modal formula ϕ is *preserved under model extensions* if for all models **M**, **N** where **N** extends **M**, if $\mathbf{M}, a \models \phi$ then $\mathbf{N}, a \models \phi$, for all $a \in M$.

THEOREM 2.6.1. *A modal formula is preserved under model extensions iff it can be defined using only propositional atoms and their negations,* $\&, \vee, \Diamond$.

Proof (1). *Original Model-Theoretic Version*
(This proof occurs in correspondence with Albert Visser in 1985. It was published in van Benthem 1995.) Call a modal formula *existential* if it has the form stated in the theorem. By a simple induction, existential formulas are preserved under model extensions. Conversely, let ϕ be so preserved. We prove the following consequence:

$$\mathbf{exist}(\phi) \models \phi, \quad \text{where } \mathbf{exist}(\phi) \stackrel{\text{def}}{=} \{\psi \text{ existential} | \phi \models \psi\}.$$

Then the required existential modal form for ϕ will exist by Compactness, being of the form $\&\Delta$ for some finite set Δ of formulas from $\mathbf{exist}(\phi)$. Now, let $\mathbf{M}, a \models \mathbf{exist}(\phi)$. Without loss of generality, again, we can take the model $(\mathbf{M}, a)$ to be ω-saturated. Next, in the usual manner, we find a second model $(\mathbf{N}, b)$ such that

$\mathbf{N}, b \models \phi$,

$\mathbf{N}, b \models \alpha$ implies that $\mathbf{M}, a \models \alpha$ for all existential modal formulas α.

Next, take the above "modal unraveling" of $(\mathbf{N}, b)$ via finite sequences of worlds, say, $\mathbf{N}_{\text{unrav}}, \langle b\rangle$, which bisimulates $(\mathbf{N}, b)$ via the map sending sequences to their end points. This will yield the following diagram:

$$\begin{tikzcd}
\mathbf{N}, b \arrow[d, shift left] & \Rightarrow_{\textbf{exist}\text{-fragment}} & \mathbf{M}, a \\
\mathbf{N}_{\text{unrav}}, \langle b\rangle \arrow[u, shift left, "\text{bisimulation}"] \arrow[urr, "F"] & &
\end{tikzcd}$$

Now, by induction on the length of sequences in $\mathbf{N}_{\text{unrav}}$, a map F may be defined from $\mathbf{N}_{\text{unrav}}$ to $\mathbf{M}$ sending $\langle b\rangle$ to a which is a homomorphism with respect to R, and which respects atomic facts in worlds. (ω-Saturation of $(\mathbf{M}, a)$ is used here to find suitable R-successors for points already mapped.) Finally, we perform a useful trick. Add a disjoint copy of $(\mathbf{M}, a)$ to $(\mathbf{N}_{\text{unrav}}, \langle b\rangle)$ to obtain a new model $\mathbf{N}_{\text{unrav}}+\mathbf{M}$. Then, extend the relation R between elements of $\mathbf{N}_{\text{unrav}}$ and elements of $\mathbf{M}$ as follows:

for all sequences Y in $\mathbf{N}_{\text{unrav}}$ and for all z in $\mathbf{M}$:
YRz if $F(Y)Rz$.

CLAIM. *F united with the identity on $(\mathbf{M}, a)$ is a bisimulation between the two models $(\mathbf{N}_{\text{unrav}}, \langle b\rangle{+}\mathbf{M}, a)$ and $(\mathbf{M}, a)$.*

Proof. This follows by a simple inspection of cases. The point of using the unraveling $\mathbf{N}_{\text{unrav}}$ instead of $\mathbf{N}$ here is to get unambiguous relationships – while that of adding a copy of $(\mathbf{M}, a)$ to the left is to enforce the backward clause of bisimulation (on top of the already established "forward" homomorphism). ■

To clinch the total argument, we again chase ϕ around the diagram:

$\mathbf{N}, b \models \phi$	(by construction)
$\mathbf{N}_{\text{unrav}}, \langle b\rangle \models \phi$	(bisimulation)
$\mathbf{N}_{\text{unrav}}, \langle b\rangle{+}\mathbf{M}, a \models \phi$	(model extension!)
$\mathbf{M}, a \models \phi$	(bisimulation). ■

Proof (2). *Stream-Lined Modern Version*
In the meantime, simpler proofs of the above result have appeared. One version is essentially due to Dick de Jongh (cf. Visser *et al.* 1995). Here is a sketch of the idea, for the equivalent preservation theorem involving submodels and universal modal forms. Start again from some model $\mathbf{M}, a \models \mathbf{univ}(\phi)$. Unravel this model to a bisimulation equivalent $(\mathbf{M}^*, \langle a \rangle)$ in the form of an intransitive acyclic tree.

CLAIM. *The atomic diagram of* $(\mathbf{M}^*, \langle a \rangle)$ *can be satisfied together with* ϕ.

After this, the usual argument works. From the resulting model $(\mathbf{N}, b)$, our ϕ can be transferred to its submodel (modulo isomorphism) $(\mathbf{M}, a)$. To prove the claim, consider ϕ together with any finite set of (negated) atoms that are true in $(\mathbf{M}^*, \langle a \rangle)$. The worlds mentioned in the latter can be described as a finite subtree, via branches going down all the way to the root $\langle a \rangle$. Now, the resulting structure can be described completely via some (inductively constructed) existential modal formula. Also, ϕ cannot imply the (universal) negation of the latter, given the assumption that $\mathbf{M}^*, \langle a \rangle \models \mathbf{univ}(\phi)$. So ϕ can be satisfied together with this existential description in some model $\mathbf{N}$. By unraveling $\mathbf{N}$ once more, this model can be taken to be an intransitive acyclic tree itself. But then, all atomic (negated) facts that were true in the above finite submodel of $(\mathbf{M}^*, \langle a \rangle)$ must also be true here. (No R-steps will be available except those explicitly demanded, which takes care of all negations.) ■

This second proof is close to the standard model-theoretic one (cf. Chang and Keisler 1973), specialized to the modal fragment of the first-order language, more or less “as is”. We return to this observation below in a more general setting. Further evidence for this analogy may be found with other model-theoretic preservation theorems, using similar methods. One example is the Lyndon homomorphism theorem for positive formulas (cf. van Benthem 1976). Here is a sketch for another classical case.

EXAMPLE. *Preservation Under Unions of Chains*
The first-order formulas preserved under unions of chains of models are precisely those definable by a universal-existential (Π_2) prenex form. In modal logic, the corresponding format must be extended (in the absence of prenex forms), as above. We only allow formulas constructed from atoms and their negations, using $\&, \vee$ as well as $\Diamond, \Box$, provided the former never scope over the latter. (In intuitionistic logic, this is the natural class of formulas with “implication rank” 2.) The classical argument again

starts from a model **M** in which the **univ-exist** consequences of ϕ hold. Then, two models **N**, **K** are found such that (1) **M** is a submodel of **N** and **N** of **K**, (2) **M** is an elementary submodel of **K**, (3) ϕ holds in **N**. Iterating this move yields a model chain in whose union ϕ holds, which then transfers to the elementary submodel **M**. Inspecting the details of this standard argument, and using the above methods, similar triples of models may be constructed for the modal language. ■

2.7. *Analyzing the General Situation: Transfer Results*

The similarities between modal logic and standard first-order logic that have come to light so far call for more general explanation. There must be some general feature in the above arguments that can be isolated, and used to explore the full extent of the analogy. One obvious general point is the pervasive use of bisimulations, which are close to the fundamental notion of "partial isomorphism" $\cong_p$ between first-order models ("cut off" at length 2). This observation may be found in van Benthem 1991, and it has inspired a systematic investigation of model theory for basic poly-modal logic in De Rijke 1993, whose results revolve around the "heuristic equation"

Modal Logic: Bisimulation $=$
Predicate Logic: Partial Isomorphism.

Another approach scrutinizes the above arguments, identifying some key lemmas of "transfer" between modal and classical reasoning. One such result is easily extracted from the earlier proof of the Invariance Theorem. Two models $(\mathbf{M}, a)$ and $(\mathbf{N}, b)$ have the same modal theory iff they possess elementary extensions which bisimulate. (De Rijke 1993 observes that one can choose the latter to be countable ultrapowers.) Here is another result of this kind, which may be of independent interest. It shows how one can "upgrade" modal equivalence to full elementary equivalence, up to bisimulation:

LEMMA 2.7.1. *Two models* $(\mathbf{M}, a)$ *and* $(\mathbf{N}, b)$ *have the same modal theory iff they possess bisimulations with two models* $(\mathbf{M}^+, a)$ *and* $(\mathbf{N}^+, b)$ *(respectively) which are elementarily equivalent.*

Proof. Upwards, the assertion is immediate. Consider the downward direction. The required models are constructed using the above *Unraveling* by finite sequences of the form $(u =)u_1, u_2, \ldots, u_k$, where each u_{i+1} is an R-successor of u_i $(1 \leqslant i < k)$, having "immediate succession" for their accessibility relation, and bisimulating with the original model via their last elements. This unravels to intransitive acyclic trees. In addition, we perform *Multiplication*, making sure that each node except the root a gets copied infinitely many times. This can be done as follows, while maintaining a bisimulation at each stage. First, copy each successor of a at level 1 countably many times, and attach these (disjoint) copies to a. There is an obvious bisimulation here, identifying copies with originals. Next, consider successors at level 2 on all branches of the previous stage, and perform the same copying process at all level-1 worlds. Again, there is an obvious bisimulation with the original model. Iterating this process through all finite levels yields our intended models $(\mathbf{M}^+, a)$ and $(\mathbf{N}^+, b)$.

CLAIM. $(\mathbf{M}^+, a)$ *and* $(\mathbf{N}^+, b)$ *are elementarily equivalent*

Proof. We use Ehrenfeucht Games. It suffices to show, for any finite n, how the Similarity Player can win in a game over n rounds between these structures. What we know at the outset is that the two roots a, b satisfy the same modal formulas. In fact, as we shall prove separately, they even satisfy the same *tense-logical* formulas. This observation will be used to describe the proper invariant for the game. Assume that in round i of the game, a match $\equiv$ has been established already between certain finite groups of worlds in the two models which satisfies three conditions:

- if $u \equiv v$, then $(\mathbf{M}^+, u)$ is equivalent with $(\mathbf{N}^+, v)$ for all tense-logical formulas up to operator depth 2^{n-i}
- if $u \equiv v$ and $u' \equiv v'$, and the distance between u and u' is at most 2^{n-i}, then the distance between v and v' is the same on the other side, and it runs via an isomorphic path, all of whose points have been matched at this stage.

Here, *distance* is measured as follows: "go from node u to node s descending the minimal distance needed to climb up to v again". (The possible backward movement forces us to use two-sided tense-logical formulas in the description of the invariant.)

- if the distance between u and u' is greater than 2^{n-i}, then on the other side, v and v' have distance greater than 2^{n-i}, too.

The upshot of all this is a number of "matched islands" on both sides, all lying a distance of more than 2^{n-i} steps apart. Now, we have to show that this invariant can be maintained in the next step by the Similarity Player, whatever world the Difference Player chooses. Let the next choice be some point P in either tree.

Case 1. P has distance $\leqslant 2^{N-i-1}$ to some point Q that was already matched at the previous stage, say to some point Q'.

Consider the unique path of length k (say) between P and Q, and attach complete tense-logical descriptions δ to its nodes up to operator depth 2^{N-i-1}. This path may then be described, from the perspective of Q, by a tense-logical formula of the form

PAST (δ_1 & PAST ($\cdots$ & PAST (δ_i and FUT (δ_{i+1} & $\cdots$ & FUT (δ_k)),
where δ_k is the full 2^{N-i-1}-description in tense logic of the point P.

The total operator depth of this formula is at most 2^{N-i-1} (being the length of the path) $+2^{N-i-1}$ (the quality of the descriptions at its nodes), which is at most 2^{N-i}. Now, at the previous stage i, Q and Q' agreed on tense-logical formulas up to the latter depth. Hence this path description is also true at Q', and we can find corresponding worlds on the other side, making the two paths isomorphic as required by our invariant, while also achieving the right degree of tense-logical equivalence.

Case 2. P lies at distance $> 2^{N-i-1}$ from all previously matched points. Take the unique path from the root to P. Describe it completely as before, with node descriptions up to level 2^{N-i-1}. The resulting tense-logical formula may be of high complexity (there is no bound on the path length), but since the two roots agree on *all* tense-logical formulas, there must be a similar path on the other side, whose end-point is an appropriate match for P. This path can be chosen so as to remain at a suitable distance from all nodes in already matched regions, by the Multiplication of nodes (we use this feature only here). Thus, again, the above invariant is maintained.

Finally, after n rounds, this invariant produces a partial isomorphism, which is a win for the Similarity Player. (For a concrete feel for the strategy, compare two modally equivalent trees where one has an infinite branch and the other does not. This also shows that we cannot improve our Lemma to the existence of a bisimulation between the unraveled multiplied models.) To wrap things up, we prove the announced

SUBLEMMA. *If the roots of two unraveled modal models have the same modal theory, then they also have the same tense-logical theory (in the basic modal language extended with a backward modal operator for "past").*

Proof. It suffices to observe a number of tense-logical validities on our trees. First:

$$\text{FUT (PAST } \alpha \,\&\, \beta) \leftrightarrow \alpha \,\&\, \text{FUT } \beta$$

$$\text{FUT } (\neg\text{PAST } \alpha \,\&\, \beta) \leftrightarrow \neg\alpha \,\&\, \text{FUT } \beta.$$

As a result, using some standard modal manipulations, every formula is equivalent to one without future operators scoping over past ones. This just leaves compounds of "pure future" (i.e., modal) formulas combined using $\neg$, & and PAST. The latter can still be simplified using two more valid equivalences:

$$\text{PAST } (\alpha \,\&\, \beta) \leftrightarrow \text{PAST } \alpha \,\&\, \text{PAST } \beta$$

$$\text{PAST } \neg\alpha \leftrightarrow \neg\text{PAST } \alpha \,\&\, \text{PAST } \textit{true}.$$

As a result, every formula is equivalent to a Boolean combination of formulas PAST^i ϕ (with i repetitions) where ϕ is purely modal. But then, the roots must agree on all tense-logical formulas. They already agreed on all modal formulas, and they will both reject any PAST formula (lacking predecessors). ■

The preceding Lemma (together with its proof) has the following consequence.

COROLLARY 2.7.2. *Two models* $(\mathbf{M}, a)$ *and* $(\mathbf{N}, b)$ *admit of a bisimulation between* a, b *iff their unraveled multiplied versions are partially isomorphic.*

This observation gives us new switches between modal logic and first-order logic. In particular, it may be used to replace the second Claim in our earlier proof of the Modal Invariance Theorem 2.2.1, thereby providing a new derivation of this result.

2.8. *Analyzing the General Situation: Predictions*

Finally, as to the full analogy between modal logic and classical logic, let us risk a bold generalization. The set of all predicate-logical formulas may be viewed as the domain of a ("meta-")model which carries some natural structure. For instance, meta-theorems like Interpolation are themselves (Π_2) first-order statements about this model, in a similarity type having one binary relation of "semantic consequence", and another of "vocabulary inclusion". (A closely related meta-model was investigated in Mason 1985. The complete first-order meta-theory of propositional logic turned out to be effectively equivalent to True Arithmetic – thereby saving the logical profession from rapid extinction.) Similar observations hold for other preservation theorems. For example, the Łoś–Tarski Theorem states an equivalence between (1) $\phi \models (\phi)^A$ (i.e., ϕ implies its own relativization to some new unary predicate A) and (2) the existence of some universal formula equivalent to ϕ. Both assertions involve slight expansion of the above meta-model to include further predicates encoding "elementary syntax" into the similarity type. Thus, the Loś–Tarski theorem becomes a Π_2-sentence, too. Now, the modal fragment is a submodel of at least the first of these predicate-logical meta-models. (With the second, we must be more careful. The result needs to be restated due to the lack of modal prenex forms – though not of modal relativizations.) In this perspective, here is a guess which would explain why one always seems able to "witness" existential quantifiers over formulas inside the modal fragment:

CONJECTURE. *The modal fragment is an* elementary submodel *of full predicate logic in the first similarity type given above.*

With results like this, one could decide transfer of meta-theorems between first-order logic and modal logic by merely inspecting their syntactic form.

2.9. *Poly-Modal Generalizations*

One test for the naturalness of the above results for the basic modal language is how they survive generalization. At least, things work very smoothly for the practically important case of poly-modal languages with families of unary modalities $\Diamond_i (i \in I)$, each with their corresponding accessibility relation R_i. One can virtually literally transcribe the above theory, putting in appropriate indices. Nevertheless, subtleties do arise occasionally. For instance, in the Interpolation Theorem, one can now also talk about the shared modalities of the two original formulas, and an interpolant should contain only these. But then, the above proof is incorrect as it stands. For, the amalgamation defined in Section 2.5 only yields bisimulating projections for (relations corresponding to) the shared modalities. In order to make the amalgamation bisimulate with the two separate models in their full language (as is required by the final argument), one has to add copies to the amalgam **MN** of those parts of **M** and **N** that branch off via non-shared successor relations, and extend the projection via the identity map on the new parts (cf. van Benthem 1994B). This is not a serious departure from the basic modal case, but it is not totally trivial either. (For earlier proofs of these results, we refer to Németi 1985, Kracht and Wolter 1991. Cf. also Sain 1989, Marx 1995, and the strong generalizations given in Madarasz 1995.)

Next, we consider a more serious generalization, namely to *polyadic modalities*. Here, one needs $(k+1)$-ary accessibility relations for each k-ary existential modality:

$$\Diamond\phi_1 \ldots \phi_k \text{ translates into}$$
$$\exists y_1 \ldots y_k (R^{k+1}x, y_1 \ldots y_k \,\&\, \&_{1 \leqslant i \leqslant k} \underline{\phi}_i(y_i)).$$

We give a quick run-down of basic results, showing what remains the same, and where cosmetic changes are needed. Under translation, modal formalisms of this kind end up in what we call provisionally the *restricted fragment* of first-order logic:

- start with all unary atomic formulas Px, and allow
- closure under Boolean operations for compounds with the same variable
- closure under existential quantifiers of the form

$$\exists y_1 \ldots y_n (R^{n+1}x, y_1 \ldots y_n \,\&\, \phi_1(y_1) \,\&\, \cdots \,\&\, \phi_n(y_n)).$$

Any set of restricting predicates R is allowed in the first-order language. The restricted fragment inherits all attractive properties of the original modal one in the obvious way (cf. van Benthem 1991A, Chapter 17, de Rijke 1993, Chapter 6).

1. The "restricted formulas" are precisely those first-order formulas $\phi(x)$ which are invariant for bisimulation with respect to the new extended set of relations. (One now has to find back-and-forth matches for triples R^3x, y_1y_2, etcetera.) The earlier model-theoretic proof goes through with mere notational changes.
2. There is an adequate semantic tableau method which establishes decidability.
3. There is a complete sequent calculus axiomatization for universal validity, whose principles may be read off from closed tableaus.

Practical complexity may increase in this system. For example, the introduction rule for a binary existential modality that emerges from the tableau calculus reads as follows:

$$\frac{\&_{X \subseteq \{1,\ldots,k\}} (\alpha \vdash \{\gamma_i | i \in X\} \text{ or } \beta \vdash \{\delta_i | i \in \{1,\ldots,k\} - X\})}{\Diamond\alpha\beta \vdash \Diamond\gamma_1\delta_1, \ldots, \Diamond\gamma_k\delta_k}$$

4. Craig Interpolation holds, either. constructively via sequent proofs, or model-theoretically. (Cf. earlier references on this topic, in particular, the algebraic superamalgamation methods of Németi 1985 and Madarasz 1995.) There is also a strong version where interpolants

have only shared modalities between antecedent and consequent (cf. Marx 1995, van Benthem 1994B).

5. The Łoś–Tarski Theorem holds by essentially the earlier argument with bisimulation invariance and copying. This requires a notion of "unraveling" via suitable finite sequences for many relations at the same time.

The latter case again requires some care in the formulation of results. For instance, in unraveling a triple Ra, bc, one has to keep track of the whole ternary configuration, indicating that the step from a to b went via the triple (abc), in order to distinguish this from a possible other situation Ra, bd. There are several notational solutions to problems like this, which we do not elaborate here. Our results in Section 4 are further generalizations of this line of thinking to still larger first-order fragments.

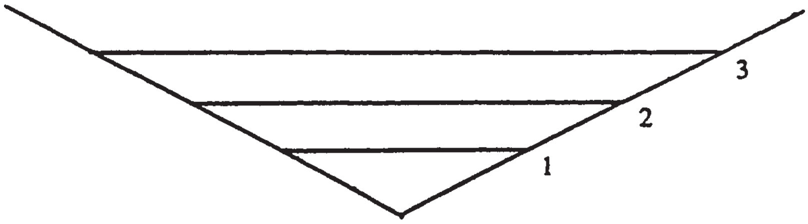
3. Finite variable fragments

3.1. *The Finite Variable Hierarchy*

Finite-variable fragments of first-order logic may be traced back to the 19th century logicians Peirce and Schroeder. They were used by Tarski in the 1950's (cf. also Henkin–Tarski 1961). Probably their first systematical study is in Henkin 1967, as a tool in trying to see what new insights cylindric algebra can provide for "pure" logic. These fragments consist of all formulas using only some fixed finite set of variables (free or bound): say $\{x\}$, $\{x, y\}$, $\{x, y, z\}$, etcetera – but otherwise allowing arbitrary combinations of quantifiers and connectives. An important connection with modal logic was pointed out in Gabbay 1981. Finite operator sets generate modal languages whose translations into first-order logic involve only some fixed finite number of variables (free and bound). For instance, the basic modal language can make do with *two* world variables only (as may be seen by further analysis of the earlier translation, judiciously "recycling" just x, y). This may be seen by changing the translation function in Section 2.1 above to produce two variants $\underline{\phi}^x$ and $\underline{\phi}^y$, having free variables x, y, respectively. The key clause for the existential modality now reads:

$$\underline{\Diamond\phi}^x \qquad \exists y(Rxy \,\&\, \underline{\phi}^y)$$
$$\underline{\Diamond\phi}^y \qquad \exists x(Ryx \,\&\, \underline{\phi}^x).$$

Similarly, e.g., the well-known temporal language with "Since" and "Until" uses essentially *three* variables, being the minimal syntactic apparatus for translating these two operators. Thus, there is a natural division into a *Finite Variable Hierarchy:*



inside whose levels one finds modal logics of ascending expressive strength. Unless otherwise stated, by the k*-variable fragment* we mean that fragment of predicate logic which uses only the variables $x_1, \ldots, x_k$ (an initial segment of some fixed enumeration), with equality and without constant or function symbols. Like the basic modal language itself, these successive levels can be characterized via a semantic invariance property (van Benthem 1991, p. 260; Barwise 1975). To state the result, recall the usual convention that $\phi(x_1, \ldots, x_k)$ denotes a formula whose free variables are all among $\{x_1, \ldots, x_k\}$. Moreover, by k*-variable formulas* we shall mean first-order formulas all of whose variables (whether free or bound) are among $\{x_1, \ldots, x_k\}$.

THEOREM 3.1.1. *A first-order formula* $\phi(x_1, \ldots, x_k)$ *is equivalent to a* k*-variable formula iff it is invariant for* k*-partial isomorphism.*

Here, k*-partial isomorphism* is a cut-off version of the well-known notion of "partial isomorphism" from Abstract Model Theory. These are nonempty families $\mathbf{I}$ of partial isomorphisms between models $\mathbf{M}$ and $\mathbf{N}$, closed under taking restrictions to smaller domains, and satisfying the usual Back-and-Forth properties for extension with objects on either side – restricted to apply only to partial isomorphisms of size at most k. "Invariance for k-partial isomorphism" means having the same truth value at tuples of objects in any two models that are connected by a partial isomorphism in such a set. The precise sense of this is spelt out in the following proof.

Proof. (Outline.) k-variable formulas are preserved under partial isomorphism, by a simple induction. More precisely, one proves, for any assignment A and any partial isomorphism $I \in \mathbf{I}$ which is defined on the A-values for all variables $x_1, \ldots, x_k$, that

$$\mathbf{M}, A \models \phi \quad \text{iff} \quad \mathbf{N}, I \circ A \models \phi.$$

The crucial step in the induction is the quantifier case. Quantified variables are irrelevant to the assignment, so that the relevant partial isomorphism can be restricted to size at most $k - 1$, whence a matching choice for the witness can be made on the opposite side. This proves "only if". Next, "if" is analogous to that for the earlier Invariance Theorem (cf. the proof of Theorem 2.2.1). One shows that any invariant formula $\phi(x_1, \ldots, x_k)$ (in the above sense) is implied by the set of all its k-variable consequences. The key step in this argument goes as before. We find two models which are elementarily equivalent for all k-variable formulas. These then possess ω-saturated elementary extensions – for which the relation of k-elementary equivalence between tuples of objects itself defines a family of partial isomorphisms up to length k, which satisfies all the above requirements for k-partial isomorphism.

3.2. *Positive and Negative Properties*

Below we summarize some of the known properties of the finite variable hierarchy. This will help us to see the strengths and weaknesses of this classification, while also pointing the way to our eventual "nice" fragments.

Positive properties

(1) Theorem 3.1.1 above provides a natural semantic characterization.

(2) Gabbay's Functional Completeness Theorem (Gabbay 1981) also shows how, conversely, for each finite-variable level, a finite set of modal operators can be constructed effectively whose modal language yields precisely that fragment.

(3) Variables are "semantic registers" providing a natural fine-structure of expressive complexity – and hence complexity hierarchies for checking first-order assertions (Immerman 1981, 1982, Hodkinson 1994, Andréka, Düntsch and Németi 1995).

(4) The controlled use of more variables, free and bound, in these fragments suggests natural "many-dimensional completions" of general modal languages, as well as links with algebraic logic (their expressive power and effective axiomatization are investigated extensively in Venema 1991, Venema 1995A, 1995B).

(5) Finite-variable fragments emerge naturally in other areas of mathematical logic, such as Relational, Polyadic and Cylindric Algebra (Németi 1991), and they are crucially involved in the formalization of set theory of Tarski and Givant 1987.

(6) Finite-variable fragments provide natural query languages supporting fixed-point operators in Finite Model Theory (cf. Kolaitis and Väänänen 1992).
(7) Finite-variable fragments admit a preservation theorem (via so-called "partial embeddings") characterizing the universal formulas (cf. Section 3.5 below).

But there is also mounting evidence as to

Negative properties (Throughout the following list $k > 2$, unless stated otherwise.)

(1) Finite-variable fragments have a poor proof theory. No finitely axiomatized Hilbert-style system exists (Monk 1969), and the complexity of the necessary axiom schemes is inevitably high (Andréka 1991). Moreover, adding new logical connectives to fragments L_k does not seem to help (Andréka and Németi 1994 discuss "hereditary nonaxiomatizability" of finite variable fragments).
(2) Finite-variable fragments are undecidable (Henkin, Monk and Tarski 1985).
(3) Even for $k > 1$, Craig Interpolation and Beth Definability fail (Sain 1990, Sain and Simon 1993, Andréka, van Benthem and Németi 1993). Cf. Section 3.5 below.
(4) Finally, here is a new type of problem. The Loś–Tarski submodel preservation theorem fails, as is demonstrated in Section 3.3.

Our contribution is both critical and constructive. Some negative items in the above list are proved in the present section (with the rest supported by references). But later on, in Section 4, we show that the negative properties all go away if we replace full predicate logic by its "guarded fragment". Alternatively, in Section 5, finite-variable fragments regain their positive properties when interpreted, not on standard models but on suitably generalized models having restrictions on available assignments.

3.3. *Failure of the Submodel Preservation Theorem*

The Loś–Tarski Theorem trivially fails for finite-variable fragments. The 1-variable formula $\forall x Ax \vee \forall x Bx$ is preserved under submodels, but lacks a prenex form with one variable (two are needed). The more natural conjecture is this: a k-variable formula is preserved under submodels iff it is equivalent to a k-universal formula. Here we define *k-universal formulas* as all those constructible in the k-variable fragment using atoms and their negations, &, $\vee$ and $\forall$. (The above formula is 1-universal as it stands.) But this result, too, fails in finite-variable fragments.

THEOREM 3.3.1. *For each $k \geqslant 3$, the k-variable fragment contains formulas that are preserved under submodels, while lacking any k-universal equivalent.*

Proof. We do the case $k = 3$ for an illustration. The general case is completely analogous. Let R be some 3-place relation. Define $\delta(x) := \exists yz\, Rxyz$ ("x is in the head of R"). Consider the following first-order formula, where $|\delta| \leqslant 2$ is a formula expressing that "there are at most two objects satisfying δ" and similarly for $|\delta| \leqslant 1$:

$$\phi: \ |\delta| \leqslant 2 \ \& \ (|\delta| \leqslant 1 \vee \forall xx'yz((Rxyz \,\&\, Rx'yz) \to x = x')).$$

This formula can be written as a purely universal prenex form (even as a 4-universal formula, by proper variable management.) So, ϕ is preserved under submodels. Now, consider the following variant Φ of ϕ (involving existential quantifiers):

$$\Phi: \ |\delta| \leqslant 2 \ \& \ (|\delta| \leqslant 1 \vee \forall xyz(Rxyz \to \exists x(\delta(x) \,\& \neg Rxyz))).$$

CLAIM 1. *ϕ and Φ are logically equivalent.*

Proof. This is a simple computation. In both directions, it suffices to consider the case where there are exactly two objects satisfying δ. "From ϕ to Φ". Assume that $Rxyz$. Then there must be some $x' \neq x$ in δ (by $|\delta| = 2$), which cannot have $Rx'yz$, by ϕ. This is the required witness for the existential quantifier. "From Φ to ϕ". Assume that $Rxyz$, $Rx'yz$. Since $\delta(x)$, there must be some x'' in δ with $\neg Rx''yz$. But then, x, x'' are different, and hence, x' must be equal to one of them. The only option here is $x' = x$, since $Rx'yz, \neg Rx''yz$ rules out $x' = x''$. ■

CLAIM 2. *Φ is in the 3-variable fragment.*

Proof. We only have to show that $|\delta| = 2$ and $|\delta| = 1$ can be expressed with 3 variables. This is a simple syntactic calculation. ■

We introduce two special models $\mathbf{M} = (D, Z)$, $\mathbf{N} = (D, U)$ for this language, where

$$D = \{0, 1, 2, 3, 4, 5\},$$
$$U = \{0, 1\} \times \{2, 3\} \times \{4, 5\},$$
$$Z = \{(i, j, k) \in U | i + j + k \text{ is even}\}.$$

CLAIM 3. *$\mathbf{M} \models \Phi$, but not $\mathbf{N} \models \Phi$.*

Proof. By direct inspection. In both cases, there are precisely two objects in δ. In particular, note how the parity in the definition of Z ensures that there will be another object in δ which does not have the same "tail". ■

Now, the counter-example is complete once we prove our final.

CLAIM 4. *Every 3-universal formula which holds in* **M** *also holds in* **N**.

Proof. One proof uses one-sided Ehrenfeucht games with three pebbles (Immermann and Kozen 1987). Here we take another road, that is suggestive for later developments. Consider the set **3PI** of all partial isomorphisms f with size at most 3 between **N** and **M**, satisfying the additional restriction that, whenever $f(a) = b$, then a and b lie in the same component $\{0, 1\}$, $\{2, 3\}$ or $\{4, 5\}$. We show that this family has the "Forth" property, from **N** to **M**. More precisely, let $f(a) = b$ and let c be any object in D. We find an object d such that the map $(f - \{(a, b)\}) \cup \{(c, d)\}$ is again in **3PI**. It suffices to consider the case where $c \neq a$. We distinguish cases for the remaining f-arguments (after removal of the object a).

Case 1. "c equals some existing f-argument different from a". Then its mate d is the corresponding f-value. (This must yield a partial isomorphism of the right kind.)

Case 2. "c differs from all existing f-arguments".

Case 2.1. Suppose that c is in the same component as some existing f-argument. Then let d be the remaining possibility in this component. (In this case, no Z- or U-relation can hold on either side of the partial isomorphism.)

Case 2.2. Suppose c is in a different component from the existing f-arguments. This case, too, will yield a partial isomorphism.

Case 2.2.1. The other f-arguments lie in the same component, and hence no Z- or U-relations can hold: then let c be its own image.

Case 2.2.2. These arguments lie in different components. Then, the choice of an image for c in its component may be made according to the parity of the sum of the values assigned to the other arguments. (One of the two available options will always do.) Finally, an easy induction on 3-formulas shows that

> If $f \in$ **3PI**, A is some assignment whose values are in the domain of f, and α is some 3-existential statement such that **N**, $A \models \alpha$, then **M**, $A \circ f \models \alpha$.

This shows that all true 3-existential statements in **N** are also true in **M**, from which the required assertion about 3-universal statements follows by duality. ■

For further information, as well as details missing from Section 3.4 below, we refer to Andréka, van Benthem and Németi 1994B, which presents a more elaborate version of the above argument, that extends to first-order languages without equality. Moreover, related ideas are used in Andréka, van Benthem and Németi 1993 to construct uniform failures of Interpolation in finite-variable fragments. For completeness, we note that an open problem concerning uniform failures of the Łoś–Tarski Theorem formulated in earlier versions of this paper was subsequently solved in Rosen and Weinstein 1995.

Some Remaining Questions

(1) Does the Łoś–Tarski Theorem hold for the 2-variable fragment?
(2) k-variable formulas that are preserved under submodels must have universal equivalents somewhere in the finite-variable hierarchy, by the ordinary Łoś–Tarski Theorem. Is there a *recursive* function f of k such that every k-variable formula preserved under submodels has an $f(k)$-variable universal equivalent? In particular, does the choice $f(k) = k + 1$ work? And what about a similar function defined over formulas?

3.4. *Modified Preservation Theorems*

The above negative argument suggests a positive result. (For convenience, we shift to a dual existential formulation.) There exists a Łoś–Tarski Theorem for k-variable fragments characterizing an appropriate syntactic notion of "k-existential definability" (by atoms and their negations, &, $\vee$ $\exists$) via preservation under "k-partial embeddings", being restriction-closed nonempty families of k-partial isomorphisms which satisfy the Forth-condition only (cf. van Benthem 1991). Partial embeddings generalize submodels as partial isomorphisms generalize isomorphisms. Here is a model-theoretic characterization for existential formulas in the k-variable fragment:

THEOREM 3.4.1. *A k-variable formula is equivalent to a k-existentially definable formula iff it is preserved under k-partial embeddings.*

Proof. That existential formulas are preserved follows by a straightforward induction. Next, let ϕ be preserved under k-partial embeddings. Then earlier arguments apply (see Theorems 2.6.1, 3.1.1). It suffices to show that ϕ is a consequence of the set k-**exist**(ϕ) of all k-existential logical consequences of ϕ. Let $\mathbf{M}, A \models k$-**exist**(ϕ). By familiar reasoning, we find a model $(\mathbf{N}, B)$ for ϕ, each of whose k-existential formulas is true in $(\mathbf{M}, A)$. Without loss of generality, we may take $(\mathbf{M}, A)$ to be ω-saturated. But then, the following defines a k-partial embedding from $\mathbf{N}$ into $\mathbf{M}$:

> all partial isomorphisms f from $\mathbf{N}$ to $\mathbf{M}$ of size at most k, such that, for all k-existential formulas α, if $\mathbf{N}, A \models \alpha$, then $\mathbf{M}, A \circ f \models \alpha$.

In proving the "Forth" clause here, one has finite approximations of the new element by means of k-existential formulas, and then finds a simultaneous witness via Saturation. In particular, our embedding sends the sequence of B-values on our k-variables to the corresponding A-values, whence ϕ must hold in $(\mathbf{M}, A)$, too. ∎

Similar finite-variable modifications exist for other classical model-theoretic results.

3.5. *Failure of the Interpolation Theorem*

Other classical key properties may fail, too. Here is a simple result to this effect. It improves a theorem in Pigozzi 1971 to interpolation failure with monadic predicates.

THEOREM 3.5.1. *Craig Interpolation fails in all k-variable fragments* $(k \geqslant 2)$, *even if the language has only unary predicate symbols.*

Proof. Consider any k-variable fragment L_k. Take k unary predicates $A_1, \ldots, A_k$. Let the first-order formula ϕ^k say that (i) each A_i holds for exactly one object (this needs two variables), (ii) all A_i are disjoint (one variable) and (iii) every object satisfies at least one A_i (again, one variable). ϕ^k holds only in domains of size k. In a similar way, construct a formula ψ^{k+1}, with new unary predicates $B_1, \ldots, B_{k+1}$, which is true only in domains of size $k+1$. Clearly $\phi^k \models \neg\psi^{k+1}$, with both formulas from L_k. By Interpolation, there must be a k-variable formula α in only identity with $\phi^k \models \alpha \models \neg\psi^{k+1}$. This is a contradiction. For, pure identity formulas using only k variables cannot distinguish between domains with k and with $k+1$ objects. ∎

The counter-example generalizes to first-order k-variable languages without identity, replacing $=$ with a suitable equivalence relation. Also, Interpolation still fails with one binary and two unary relation symbols (Andréka, van Benthem and Németi 1993). With only two nonlogical symbols, the question is open. (Madarasz 1995 shows that logics without interpolation may have a weaker two-predicate "relevance property".)

3.6. *Conclusion*

The preceding analysis shows that finite-variable fragments do not explain all there is to modal logic. Hence, as they stand, they cannot serve as the nice fragments of first-order logic sought in our Introduction. But we can turn this observation into a further requirement. Nice fragments of first-order logic should support a finite-variable hierarchy which does have the desired meta-poperties. With this new desideratum in mind, we now turn to an alternative analysis, whose focus is restriction of quantifiers.

4. BOUNDED QUANTIFIER FRAGMENTS

4.1. *The Basic Restriction Schema*

The basic modal fragment is only a subset of the full two-variable fragment, since its syntax satisfies additional constraints. In particular, all quantifiers in translations of modal formulas occur "restricted" or "bounded", in the forms

$$\exists y(Rxy \,\&\, \phi(y)),\ \forall y(Rxy \rightarrow \phi(y)).$$

Semantically, the latter form correlates with the earlier definition of bisimulation, explaining its particular zigzag clauses. This observation suggests another classification. What we are dealing with are *quantifier restrictions*, which may be varied along various dimensions. The general schema here is as follows:

$$\exists \mathbf{y}(R\mathbf{xy} \,\&\, \phi(\mathbf{x}, \mathbf{y}, \mathbf{z}))$$

where **x**, **y**, **z** are finite sequences of variables.

And the question is how much can be allowed as to variable occurrences without losing the attractive features of the basic modal logic, in particular, its decidability. We shall take this perspective as our point of departure in a hierarchy of "bounded" fragments of predicate logic. Its initial stages already occurred in Section 2.9 above. For instance, polymodal logic shows that families of different restricting predicates R_i are admissible. And polyadic modal logics showed that one can allow restrictions of the special form $\exists \mathbf{y}(Rx, \mathbf{y} \,\&\, \&_i \phi_i(y_i))$ without major changes in theory and practice. We shall be concerned mainly with the following schemata in what follows:

Fragment 1	$\exists \mathbf{y}(R\mathbf{yx} \,\&\, \phi(\mathbf{y}))$
Fragment 2	$\exists \mathbf{y}(R\mathbf{yx} \,\&\, \phi(\mathbf{x}, \mathbf{y}))$
Fragment 3	$\exists \mathbf{y}(R\mathbf{yx} \,\&\, \phi(\mathbf{x}, \mathbf{y}, \mathbf{z}))$.

Restricted formulas play a crucial role in absoluteness in Set Theory ("Δ_0-formulas"). Let us be more precise. These fragments of a standard first-order language start with arbitrary atomic formulas, and allow further constructions with Boolean operators and the above bounded quantifiers, where R can be any relation symbol (this atomic formula is called the "guard" of the formula), whose variables may appear in any order and multiplicity. Identity atoms $u = v$ are allowed, but not as guards. There are other plausible bounded fragments – but the present ones will do. In particular, Fragment 2, dubbed the *Guarded Fragment* of predicate logic, displays nice modal behaviour in the sense of the earlier sections. To understand these, it is helpful to consider a restricted version of Fragment 2, closer to the basic modal language, which involves special relations R not occurring anywhere except as a guard, with a fixed argument order **x**, **y**. Crucial for these fragments is the atomic nature of guards: Boolean combinations of atomic formulas are not permitted. For example, symmetry of a relation is in Fragment 2, but transitivity is not. These fragments may be understood in various ways. Model-theoretically, one can extend the earlier modal *bisimulations* to describe them (Section 4.2 below), which shows that we have a genuine upward hierarchy of expressive strength. We use a corresponding model unraveling method to prove a Loś–Tarski theorem. Next, we take a more combinatorial approach, focusing on their "looseness" and decidability (cf. Section 2.3 above). Fragment 1 is decidable, being close to modal logic. By contrast, Fragment 3 is easily shown undecidable. Our main result is that the powerful intermediate Guarded Fragment is decidable. (Indeed, as we shall show elsewhere, it has a uniform finite model property.) This decidability theorem generalizes several existing results from modal and algebraic logic.

4.2. *Bounded Fragments and Bisimulation*

Bounded fragments may be analyzed semantically in terms of modal bisimulations. For this purpose, one can fix the earlier modal Invariance Theorem as a target, and use its proof as a heuristic for generating appropriate notions of semantic simulation. This style of analysis uses the following notions. By a *partial isomorphism*, we mean a finite one-to-one partial map between models which preserves relations both ways. In any model **M**, we call a set X of objects *guarded* if there exists a relation symbol R, say k-ary, and objects $a_1, \ldots, a_k \in M$ (possibly with repetitions) such that $R^{\mathbf{M}}(a_1, \ldots, a_k)$ and $X = a_1, \ldots, a_k$. Here, we merely formulate appropriate bisimulations for the Guarded Fragment 2, the other fragments involve simple and obvious variations.

DEFINITION. Guarded Bisimulations
A *guarded bisimulation* is a nonempty set **F** of finite partial isomorphisms between two models **M** and **N** which satisfies the following back-and-forth conditions. Given any $f: X \to Y$ in **F**,

(i) for any guarded $Z \subseteq M$ there is a $g \in \mathbf{F}$ with domain Z such that g and f agree on the intersection $X \cap Z$
(ii) for any guarded $W \subseteq N$ there is a $g \in \mathbf{F}$ with range W such that the inverses g^{-1} and f^{-1} agree on $Y \cap W$.

A *guarded k-bisimulation* is a set F as above, except that the partial isomorphism and the mentioned guarded subsets are all required to be of size at most k.

The point of this definition shows in semantic invariance for guarded bisimulation, proved by straightforward induction on the construction of F2-formulas (members of Fragment 2). The zigzag conditions take care of the bounded existential quantifiers.

PROPOSITION 4.2.1. *Let* **F** *be a guarded bisimulation between models* **M** *and* **N** *with* $f \in \mathbf{F}$. *For all guarded formulas* ϕ *and all variable assignments* α *into the domain of* f*, we have* $\mathbf{M}, \alpha \models \phi$ *iff* $\mathbf{N}, f \circ \alpha \models \phi$.

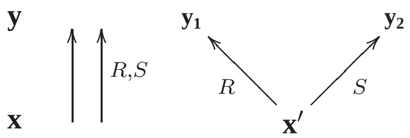
Moreover, a straightforward analogue of the proof of Theorem 2.2.1 yields a full model-theoretic preservation theorem here. Also as before, we can "cut off" guarded bisimulations at any finite length – and the proof of Theorem 3.1.1 then carries over, too – producing a semantic characterization of guarded finite variable fragments.

THEOREM 4.2.2. *Let* ϕ *be any first-order formula.* ϕ *is invariant for guarded bisimulations iff* ϕ *is equivalent to an F2 formula.* ϕ *is invariant for guarded* k*-bisimulations iff* ϕ *is equivalent to a formula in the* k*-variable subfragment of F2.*

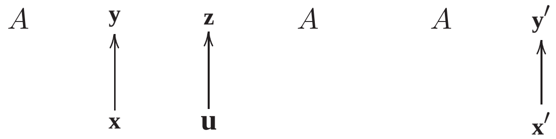
We can adapt these notions to Fragments 1 and 3 in a straightforward manner. The modified back-and-forth clauses will now match the relevant quantifier restrictions. We do not spell them out here. Here is an application of this semantic analysis.

THEOREM 4.2.3. *The three fragments form a properly ascending hierarchy.*

Proof. We sketch the gist of the counter-examples. (1) The formula $\exists y(Rxy \,\&\, Sxy)$ is in Fragment 2, but it is not equivalent to any formula in Fragment 1. Namely, the following two models have different truth values for this formula – even though there runs a Fragment-1 bisimulation between them, which consists of the following three matches between single objects: (x, x'), (y, y_1), (y, y_2).



(2) The formula $\exists y(Ay \,\&\neg Rxy)$ is in Fragment 3, but without being equivalent to one in Fragment 2. For, it can distinguish between the following two models, even though they admit of an obvious Fragment-2 bisimulation consisting of the partial isomorphisms $\{(x, x')\}$, $\{(z, y')\}$, $\{(x, x'), (y, y')\}$, $\{(u, x'), (z, y')\}$:



(3) Finally, Fragment 3 is still somewhat poorer than predicate logic as a whole. For instance, the formula $\forall x Ax$ is beyond it. This may be shown by the Fragment-3 bisimulation (x, x) between the following two models:

y

x A **x** A ∎

4.3. *Unraveling Models*

As another useful application of guarded bisimulations, we generalize the unraveling construction of Section 2. In the next section, we will apply this technique to obtain the Łoś–Tarski Theorem – while it will also serve as an inspiration for later decidability arguments. For a start, let **M** be any ordinary model for predicate logic.

DEFINITION. The *unraveling* $\mathbf{M}^u$ of **M** has for its objects all pairs (π, d) – where the "path" π is a finite sequence of guarded sets, and the M-object d is "new" in π: i.e., it occurs in the final set of π but not in the one before that. The interpretation of predicate symbols Q is as follows. $I(Q)$ holds for a finite sequence of objects $\langle(\pi_i, d_i)\rangle_{1 \leqslant i \leqslant k}$ iff $I^{\mathbf{M}}(Q)\langle d_i\rangle_{1 \leqslant i \leqslant k}$ *and* there is some maximal path π^* among those listed of which all other π_i are initial segments in such a way that their new objects d_i remain present in each set until the end of π^*. Finally, we let $\mathbf{F}^u$ be the family of all restrictions of the finite maps sending (π_i, d_i) to d_i for all guarded finite domains in $\mathbf{M}^u$.

PROPOSITION 4.3.1. $\mathbf{F}^u$ *is a guarded bisimulation from* $\mathbf{M}^u$ *to* **M**.

Proof. (i) All maps in $\mathbf{F}^u$ are partial isomorphisms. Preservation of atomic relations is obvious in going from $\mathbf{M}^u$ to **M**. Vice versa, the fact that relations are preserved backwards, as well as the required injectivity, depends heavily on the assumption that the domain is guarded. Let $f \in \mathbf{F}^u$. Then $S = \mathrm{Dom}(f)$ is guarded, which implies that there is a path π^* such that, if $(\pi, d) \in S$, then π is an initial segment of π^* such that d remains in the members of π^* all the way up. This immediately ensures that relations are preserved backwards, too. To prove injectivity, it suffices to show that, if (π, d), $(\pi', d) \in S$, then $\pi = \pi'$. Indeed, since both π and π' are initial segments of π^*, one of them must be larger than the other. If π were larger than (say) π', the object d would persist from π upwards – and hence it would not be new in π'. (ii) Next, we check the two back-and-forth properties. From $\mathbf{M}^u$ to **M**, this is easy, using the fact that our finite maps are submaps of a single relation-preserving map. Going in the opposite direction, consider $f\colon X \to Y$ with some guarded set Z in **M**. Since X is guarded, there is some maximal path π^* among its objects. Therefore, there is also a maximal path π^+ among the objects which are mapped by f onto the intersection $Y \cap Z$. We either use this path, or extend it by the guarded set Z (in case Z contains objects that are not in the last set of π^+). Now, the latter path gives us an obvious sequence of objects satisfying the guard of Z (some of them new at the end, others suitably introduced in initial sequences). The induced partial isomorphism is the one we are looking for. ■

The above construction can be slightly generalized. Given any finite set Y of objects in **M** (guarded or not), we can define a *parametrized unraveling*

$\mathbf{M}^u(Y)$ whose paths all start from Y (continuing with guarded sets only), and whose objects (π, d) are defined just as above. By a straightforward adaptation of the above argument, the obvious restriction map $\mathbf{F}^u(Y)$ is a guarded bisimulation from $\mathbf{M}^u(Y)$ to $\mathbf{M}$.

Also, we can put finite fine-structure into the construction, which will serve to specialize later results to k-variable fragments. In particular, the *k-unraveling* $\mathbf{M}^u_k$ is that submodel of the above $\mathbf{M}^u$ whose paths contain only sets of size at most k. The above proof then easily shows that the restriction map $\mathbf{F}^u$ is even a guarded k-bisimulation (Section 4.2). Finite thresholds may be combined with parametrization, to define models $\mathbf{M}^u(Y)_k$ with the obvious properties. Finally, one can also restrict the *length*, rather than the "width", of paths in unraveled models. This would connect up with languages of restricted quantifier or modal operator depth in ways familiar from the first-order and modal literature. The latter direction will not be pursued here.

Finally, we state another useful auxiliary result for the Loś–Tarski Theorem to follow. (This generalizes the modal construction in the proof of Theorem 2.6.1.) Let $\mathbf{M}$, $\mathbf{N}$ be two models and let $f : Y \to Z$ be a finite map between them such that guarded existential formulas are preserved along f. That is, $\mathbf{M}, \alpha \models \psi$ implies $\mathbf{N}, f \circ \alpha \models \psi$ for all guarded existential formulas ψ and all evaluations α of the free variables of ψ into Y. We say that f *preserves guarded existential formulas from* $\mathbf{M}$ *to* $\mathbf{N}$.

PROPOSITION 4.3.2. *Let the finite map $f : Y \to X$ preserve guarded existential formulas from $\mathbf{M}$ to $\mathbf{N}$, and let $\mathbf{N}$ be ω-saturated. Then there is an extension F of f which maps $\mathbf{M}^u(Y)$ into $\mathbf{N}$ such that F is an isomorphism on all guarded subsets of $\mathbf{M}^u(Y)$.*

We note that though, strictly speaking, Y is not a subset of $\mathbf{M}^u(Y)$, it can be identified with $\{(\langle Y\rangle, d) | d \in Y\} \subseteq \mathbf{M}^u(Y)$.

Proof. For notation's sake, we set $\mathbf{M}' \stackrel{\text{def}}{=} \mathbf{M}^u(Y)$. If S is a subset of $\mathbf{M}'$ and if $F : S \to N$, then we say that F *locally preserves guarded existential formulas* if the restriction maps $F|G$ preserve guarded existential formulas from $\mathbf{M}'$ to $\mathbf{N}$, for all guarded subsets G of $\mathbf{M}'$ and for $G = Y$. For all natural numbers $k \geqslant 1$ we define

$$D_k = \{(\pi, d) \in \mathbf{M}' | |\pi| \leqslant k\}.$$

Then $D_1 = \{(\langle Y\rangle, d) | d \in Y\}$, so D_1 is the set we already identified with Y. Now define a map $F_1 : D_1 \to \mathbf{N}$ by

$$F_1((\langle Y\rangle, d)) \stackrel{\text{def}}{=} f(d) \quad \text{for all } d \in Y.$$

By assumption, F_1 locally preserves guarded existential formulas. Now, assume that $F_k : D_k \to \mathbf{N}$ has been defined, subject to our preservation condition. We define an extension to D_{k+1} with the same property. For all paths π of length $k+1$, we set

$$Z_\pi \stackrel{\text{def}}{=} \{(\pi, d) | (\pi, d) \in \mathbf{M}'\}.$$

By the construction of $\mathbf{M}'$ then, we have the following

(1) There is a greatest guarded subset G_π of D_{k+1} containing Z_π.
(2) $G_\pi - Z_\pi$ is contained in a guarded subset G' of D_k if $k > 1$.
(3) Each guarded subset of D_{k+1} is contained in some G_π or is contained in D_k.
(4) $D_{k+1} - D_k$ is the disjoint union of the Z_π's.

Indeed, G_π can be constructed from the last element of π, and G' from the last but one element of π. (4) is trivial, (3) comes from the definition of relations in $\mathbf{M}'$. We now define F separately on all Z_π's. Let π be fixed. For simplicity, we will consider the elements of G_π also as variables. Then the identity map Id evaluates these variables in $\mathbf{M}'$, and F_k is an evaluation of the variables $G_\pi - Z_\pi$ in $\mathbf{N}$. Set

$$\Sigma \stackrel{\text{def}}{=} \{\psi | \psi \text{ is a guarded existential formula with free variables in } G_\pi \text{ such that } \mathbf{M}', \text{Id} \models \psi\}.$$

Clearly, $\mathbf{M}'$, Id $\models \Sigma$. We want to show that

$$\mathbf{N}, F_k \models \exists Z_\pi \& \Sigma.$$

By G_π's being guarded, there is a relation symbol R and an enumeration $\mathbf{g}$ of G_π such that $R(\mathbf{g})$ holds in $\mathbf{M}'$. Let $C_\pi \stackrel{\text{def}}{=} G_\pi - Z_\pi$, and let $D \subseteq \Sigma$ be finite. Then

$$\mathbf{M}', \text{Id}|C_\pi \models \exists Z_\pi (R(\mathbf{g}) \& \& \Delta),$$

so by (2) and the inductive hypothesis, we have

$$\mathbf{N}, F_k \models \exists Z_\pi (R(\mathbf{g}) \& \& \Delta).$$

By ω-saturation of $\mathbf{N}$ then:

$$\mathbf{N}, F_k \models \exists Z_\pi \& \Sigma.$$

Let $Z'_\pi = \{d' | d \in Z_\pi\}$ be such elements in $\mathbf{N}$ and define

$$F_{k+1}(d) \stackrel{\text{def}}{=} d' \quad \text{for all } d \in Z_\pi.$$

Doing this for all π, by (4) we defined

$$F_{k+1}\colon D_{k+1} \to \mathbf{N}.$$

Now $F_{k+1}|G$ preserves guarded existential formulas for all $G = G_\pi$ – by our construction – so by (3), F_{k+1} locally preserves guarded existential formulas. Define

$$F \stackrel{\text{def}}{=} \cup\{F_k | k \geqslant 1\}.$$

Clearly $F : \mathbf{M}' \to \mathbf{N}$, F is an extension of f, and F locally preserves guarded existential formulas. It remains to show F is an isomorphism on all guarded subsets. Thus, let G be a guarded subset of $\mathbf{M}'$. Since F locally preserves guarded existential formulas ψ, for all these we have

$$\mathbf{M}', \mathrm{Id}|G \models \psi \quad \textit{implies } \mathbf{N}, F \models \psi$$

if the free variables of ψ are among G. Taking ψ to be a suitable conjunction of atomic formulas $\{\neg u = v, R(\mathbf{g}), \neg R(\mathbf{g})\}$ – with $u, v \in G$, $u \neq v$, sequences $\mathbf{g}$ of variables from G and relation symbols R – we get that F is one-one on G and $F|G$ preserves forward and backward relations: i.e., F is an isomorphism on G. ∎

Again, this result may be specialized in various ways. Most importantly, if we only have preservation of guarded existential k-formulas, the extension homomorphism F can be defined on the k-unraveling $\mathbf{M}^u(Y)_k$, provided that $|Y| \leqslant k$.

4.4. *Preservation Under Submodels*

One of our recurring semantic tests for being a “nice fragment” of predicate logic was validity of the Łoś–Tarski Theorem characterizing those formulas which are preserved under submodels. And indeed, we have the following result here.

THEOREM 4.4.1. *A formula ϕ in the Guarded Fragment is preserved under submodels iff it is equivalent to an F2-formula constructed from atomic formulas and negated atomic formulas by the use of* $\vee$, & *and* $\forall$.

Proof. From universal definition to submodel preservation, the assertion is obvious. The converse can be proved along the lines of the first proof for Theorem 2.6.1 above. We merely outline the main steps of the construction. The reader is invited to check back with the original argument for motivation. The final aim is again to show that $\mathbf{univ}(\phi) \models \phi$, where $\mathbf{univ}(\phi)$ consists of all universal consequences of ϕ (with the same set $\mathbf{x}$

of free variables). Consider any model and assignment $\mathbf{M}, \alpha \models \mathbf{univ}(\phi)$. Without loss of generality, we may suppose that this model is unraveled. (By the results of Section 4.2, valid consequence for guarded formulas is witnessed entirely by unraveled models.) Moreover, this model may be taken to start from a one-step path formed by the finite set Y of objects in **M** assigned to the free variables of ϕ. This is the parametrized unraveling $\mathbf{M}^u(Y)$. Next, entirely by previous reasoning, there exists some ω-saturated model $(\mathbf{N}, \beta)$ for ϕ such that all existential formulas (constructed as in the Theorem – but now using $\exists$ instead of $\forall$) that are true in $(\mathbf{M}, \alpha)$ are also true in $(\mathbf{N}, \beta)$. Thus, the finite partial isomorphism sending the objects $\alpha(\mathbf{x})$ to the $\beta(\mathbf{x})$ preserves all guarded existential formulas going from **M** to **N**. By Proposition 4.3.2, then, there exists a mapping **F** wich is a guarded analogue of the earlier "partial embeddings" (cf. Section 3.4) – whence, at this stage, we have a complete proof for Łoś–Tarski with respect to preservation under the latter notion. But we can improve matters as in the basic modal case. We extend the model **M** as follows to a new model $\mathbf{M}^+$. To each of the guarded subsets X in **M**, which was mapped by its representative in **F** onto a corresponding guarded subset Y of **N**, we attach a unique copy of **N**, which identifies X with Y. Now, it is easy to see that the identity maps on guarded subsets in all these added parts, united with the partial embedding **F** already constructed is a guarded bisimulation between **N** and $\mathbf{M}^+$. By invariance for guarded bisimulations, the original formula ϕ is true in $\mathbf{M}^+$ at assignment α – whence it holds at $(\mathbf{M}, \alpha)$ by its preservation under submodels. ■

Finally, we note that, exercising some care in syntactic details of the above argument, our Łoś–Tarski Theorem may be specialized to completely characterize submodel preservation for all *finite-variable fragments* of the Guarded Fragment.

THEOREM 4.4.2. *An F2-formula ψ is preserved under submodels iff it is equivalent to an F2-existential formula with the same variables.*

4.5. *Decidability via Decomposition of Universal Validity*

To complete our analysis of bounded fragments, we now turn to their decidability. One way of understanding this typical phenomenon extends the earlier semantic tableaus. Their crucial point was to find some decomposition rule for validity of a sequent with only existential quantifers (plus atoms) on both sides. This is not the eventual outcome achieved here (cf. Section 4.5 – we do not know if the Guarded Fragment has a simple reduction like this). But as a warm-up, it is useful to see how far one can push direct modal arguments. First, here is a simple earlier observation:

FACT 4.5.1.

$$\exists y_1(Rxy_1 \,\&\, \phi_1(y_1)), \ldots, \exists y_k(Rxy_k \,\&\, \phi_k(y_k)) \models$$
$$\exists y_1(Rxy_1 \,\&\, \psi_1(u_1)), \ldots, \exists y_m(Rxy_m \,\&\, \psi_m(y_m))$$
$$\text{iff} \quad \textit{for some}\, i \ (1 \leqslant i \leqslant k)\, \phi_i \models \psi_1, \ldots, \psi_m.$$

This fact depended on bisimulation invariance for this language – or rather, invariance of these formulas for the generated submodels in the rooting construction. There is another convenient reduction for modal formulas involving different "current worlds" (cf. Kracht 1993 for this generalization inside modal logic itself):

FACT 4.5.2.

$$\phi_1(x_1), \ldots, \phi_k(x_k) \models \psi_1(x_1), \ldots, \psi_k(x_k)$$
$$\text{iff} \quad \textit{for some} \ \ i \ (1 \leqslant i \leqslant k) \ \phi_i(x_i) \models \psi_i(x_i).$$

This may be proved like Fact 4.4.1, by mere disjoint union of counter-examples to the lower sequents. Thus, Fact 4.4.2 holds for all first-order formulas that are invariant for disjoint unions (van Benthem 1985 characterizes these). Next, we take this further.

THEOREM 4.5.3. *Validity of formulas in Fragment 1 is decidable.*

Proof (Outline). First, we perform all possible propositional reductions in a sequent, so that only atoms and existential quantifiers remain on both sides. Then we prove a reduction to matrix formulas like above. Again, we glue together counter-examples for sequents below where the quantifiers have been stripped off, so as to refute the original sequent with quantifiers. The longer sequent arguments **x**, **y** do not make an essential difference to this modal construction. And neither does the presence of arbitrary arguments **y** in the matrix formula, provided that the former have the required invariance property. Indeed, even when starting with "mixed" initial formulas like $\exists y(Rx_1x_2, y \,\&\, \phi_1(y))$, $\exists y(Rx_3x_1, y \,\&\, \phi_2(y))$ and $\exists y(Rx_2, y \,\&\, \phi_3(y))$ followed by similar heterogeneous conclusions, one just matches up premise/conclusion pairs with identical sequences of **x**-parameters, as no semantic dependencies hold between behaviour of R-successors for sequences and their subsequences. (In special model classes with extra "frame conditions" on R, this would have to be re-checked.) ■

The preceding type of argument establishes more than was stated. It is easy to see that all counter-examples constructed for non-valid formulas may be taken to be *finite*:

COROLLARY 4.5.4. *Fragment 1 has the Finite Model Property.*

Things become more difficult for Fragment 2, where the parameters **x** of the guard can also occur in the matrix formula. For instance, as for direct reductions, we do have the valid consequence $\exists y(Rxy \,\&\, \neg Rxy) \models \exists y(Rxy \,\&\, \perp)$, but we do not have $\neg Rxy \models \perp$ or any obvious variant thereof. Nevertheless, we can prove a modal-style direct reduction to deal with the quantifier restriction schema $\exists \mathbf{y}(R\mathbf{x}, \mathbf{y} \,\&\, \phi(\mathbf{x}, \mathbf{y}))$, provided that we forbid occurrences of the relation R outside of quantifier guards. Also, we need to have the variables **x**, **y** in the guards occurring in the order stated. (Thus, even basic tense logic falls outside the scope of the following argument.)

Although the following argument is not our final contribution, we present it for its independent interest – while it also highlights some difficulties to be resolved. For a start, we note a useful normal form for the Guarded Fragment.

FACT 4.5.5. Every formula is equivalent to one whose immediate quantifier scope jumps, being of the form $[\alpha{=}]\exists \mathbf{y}(R\mathbf{x}, \mathbf{y} \,\&\, \ldots [\beta{=}]\exists \mathbf{z}(R\mathbf{u}, \mathbf{z} \,\&\, \ldots) \ldots)$, always have at least one variable y from **y** among the parameters **u**.

Proof. If this normal form fails somewhere, then repair this, working inside out, by removing inner formulas outside of the scope of the outer ones – using the valid logical equivalence $\alpha \leftrightarrow (\beta \,\&\, [T/\beta]\alpha) \vee (\neg\beta \,\&\, [\perp/\beta]\alpha)$. ■

The resulting formulas with linked chains of successive guards resemble the “secure” formulas of second-order logic (cf. van Benthem 1986). Next, we need a simplified version of a notion in Section 4.2. The *unraveling* $\mathbf{M}_{\text{unrav}}$ of a first-order model **M** consists of all finite sequences of objects in **M**, with predicates defined as follows:

$$\begin{aligned} &RX_1 \ldots X_k Y_1 \ldots Y_m \text{ iff} \\ &\quad \exists d_1 \ldots d_m : R_{\mathbf{M}} \text{ last } (X_1) \ldots \text{ last } (X_k) d_1 \ldots d_m \\ &\quad \& \, Y_i = X_1\hat{}\ldots\hat{}X_k\hat{}\langle d_i \rangle \, (1 \leqslant i \leqslant m) \end{aligned}$$

and for all other Q,

$$QX_1 \ldots X_n \quad \text{iff} \quad Q_{\mathbf{M}} \text{ last } (X_1) \ldots \text{ last } (X_n).$$

Here is the key semantic fact concerning this construction.

LEMMA 4.5.6. *For all formulas $\phi = \phi(x_1, \ldots, x_n)$ in the current fragment,*

$$\begin{aligned}&\mathbf{M}_{\text{unrav}} \models \phi[X_1, \ldots, X_n]\\ &\quad \text{iff} \quad \mathbf{M} \models \phi[\text{last}(X_1), \ldots, \ \text{last}(X_n)].\end{aligned}$$

Combining the Unraveling Lemma with the earlier normal forms, we see that, for normal forms $\exists\mathbf{y}(R\mathbf{x}, \mathbf{y} \,\&\, \phi(\mathbf{x}, \mathbf{y}))$, evaluation of the part $\phi(\mathbf{x}, \mathbf{y})$ "moves upward". Its truth value depends only on objects reachable through a finite chain of R-steps, starting from a tuple containing some y in $\mathbf{y}$. In particular, immediate R-successors of $\mathbf{x}$ are never encountered in the process of evaluation. Now, we are ready to describe the desired general reduction. Consider any first-order consequence schema in the current impoverished version of Fragment 2, which is of the form

$$\begin{aligned}\#\qquad &\&\{\text{non } R\text{-atoms}, \ \exists\mathbf{y}(R\mathbf{x}, \mathbf{y} \,\&\, \phi(\mathbf{x}, \mathbf{y}))\} \quad \models\\ &\vee\{\text{non } R\text{-atoms}, \ \exists\mathbf{u}(R\mathbf{z}, \mathbf{u} \,\&\, \psi(\mathbf{z}, \mathbf{u}))\}.\end{aligned}$$

Without loss of generality, we assume that no atom occurs on both sides, and that each "parameter group" $\mathbf{x}$ occurs on both sides. (The latter can always be achieved by inserting "inert" formulas with a constant "true" or "false" matrix for these parameters). There can be more than one formula to the left (or right) for each parameter group.

PROPOSITION 4.5.7. *A consequence # holds iff for some totally disjoint choice of variables* $\mathbf{y}$*, we have a valid schema of the following form* £*:*

$$\begin{aligned}&\&_{\text{all parameters } \mathbf{x}}\{\text{non } R\text{-atoms}, \ \phi(\mathbf{x}, \mathbf{y})\} \quad \models\\ &\vee_{\text{all parameters } \mathbf{x}} \vee_{\text{all left-hand } \phi(\mathbf{x}, \mathbf{y})} \{\text{non } R\text{-atoms}, \ \psi(\mathbf{x}, \mathbf{y})\}.\end{aligned}$$

EXAMPLE. Let the schema # to be reduced have the form displayed below:

$$\begin{aligned}&\exists y_1 y_2(Rx_1x_2y_1y_2 \,\&\, \phi_1(x_1, x_2, y_1, y_2))\\ &\& \exists y_3 y_4(Rx_1x_2y_3y_4 \,\&\, \phi_2(x_1, x_2, y_3, y_4))\\ &\& \exists y_5 y_6(Rx_1y_5y_6 \,\&\, \phi_3(x_1, y_5, y_6))\\ &\models \quad \exists y_7 y_8(Rx_1x_2y_7y_8 \,\&\, \psi_1(x_1, x_2, y_7, y_8))\\ &\qquad \vee \exists y_9 y_{10}(Rx_1y_9y_{10} \,\&\, \psi_2(x_1, y_9, y_{10})).\end{aligned}$$

Its reducing schema £ described in the above Proposition looks as follows:

$$\begin{array}{lcl} \phi_1(x_1, x_2, y_1, y_2) & & \psi_1(x_1, x_2, y_1, y_2) \\ \& \, \phi_2(x_1, x_2, y_3, y_4) & \models & \vee \psi_1(x_1, x_2, y_3, y_4) \\ \& \, \phi_3(x_1, y_5, y_6) & & \vee \psi_2(x_1, y_5, y_6). \end{array}$$

Outline of a Proof for the Proposition. From £ to #, a simple inspection suffices. Next, from # to £, suppose that the reducing sequent is not valid. Then it has a counterexample **M** with some assignment verifying its antecedent, while falsifying every disjunct in its consequent. Unravel **M**, and choose sequence-objects for the various y in the parameter groups **y** on the left making them all incomparable. In particular, then, their hereditary R-successors (recall the above remark about upward evaluation) will all be different. This gives us freedom for the following stipulation. For each of the parameters **x**, let its only R-successors be the vectors of objects for its associated **y** in the list of formulas to the left of #. By previous observations, this does not affect truth values in **M** for matrix formulas ϕ. Thus, this slightly modified model verifies all restricted formulas $\exists \mathbf{y}(R\mathbf{x}, \mathbf{y} \,\&\, \phi(\mathbf{x}, \mathbf{y}))$ to the left of schema #, and it falsifies all formulas $\exists \mathbf{u}(R\mathbf{z}, \mathbf{u} \,\&\, \psi(\mathbf{z}, \mathbf{u}))$ on its right. ■

This analysis also provides constructive information on the *complexity* of decidability. The argument given here depends crucially on the syntactic restrictions of the current version of Fragment 2. We shall follow another route in the following Section. But to round off our present discussion, we conclude with an easy result about Fragment 3. The latter turns out to be *undecidable*. For, the general parametrized quantifier restriction schema $\exists y(R\mathbf{xy} \,\&\, \phi(\mathbf{x}, \mathbf{y}, \mathbf{z}))$ is as powerful as predicate logic itself.

FACT 4.5.8. Predicate-logical satisfiability is effectively reducible to satisfiability in the parametrized restriction language F3. Hence, universal validity of F3-formulas is undecidable.

Proof. The reduction takes any predicate-logical sentence ϕ to its relativization $\rho(\phi)$ to some unary predicate U not occurring in ϕ. $\rho(\phi)$ lies inside the parametrized restriction language. It is easy to see that ϕ is satisfiable if and only if $\rho(\phi)$ is. ■

4.6. *Decidability of the Guarded Fragment*

In this Section, we present a proof of decidability for the Guarded Fragment without identity. It is inspired by the above unraveling method, but even more by earlier proof techniques using so-called "mosaics" developed originally for "generalized assignment models" (which validate the logic Crs to be introduced in Section 5). Németi 1994, Andréka and Németi 1994 have a full presentation, plus origins in cylindric algebra.

THEOREM 4.6.1. *The Guarded Fragment is Decidable*

Proof. Our strategy is as follows. We show that any guarded satisfiable formula ϕ has a finite "quasi-model" (described below) of size effectively computable from that of ϕ, which can be used conversely to generate a model for ϕ. (This is like modal filtration: Goldblatt 1987, van Benthem 1996 – but we do not present the procedure as defining a model for ϕ.) Thus, the question whether a guarded formula is satisfiable is equivalent to whether it has a finite quasi-model – and from the effective description below, it is easily seen that the existence of such a structure is decidable.

From Standard Models to Finite Quasi-Models

Suppose that a formula ϕ is satisfiable in some standard model **M**. Let V be the set of variables occurring in ϕ (free or bound). Henceforth, we restrict attention to the finite set Sub_ϕ consisting of ϕ and all its subformulas (and closed under alphabetic variants using only variables in V, as explained below). Each variable assignment realizes a "type" Δ consisting of finitely many formulas from this set. Types satisfy some closure conditions, which will emerge in due course in the proofs that follow. Our quasi-model has a universe consisting of the finitely many types realized in **M**. Furthermore note that, for each guarded formula $\exists \mathbf{y}(Q\mathbf{xy} \,\&\, \psi(\mathbf{x}, \mathbf{y})) \in \Delta$ (no special order for the variables intended), there exists a type Δ' with (i) $Q\mathbf{xy}, \psi(\mathbf{x}, \mathbf{y}) \in \Delta'$ and (ii) Δ, Δ' agree on all "unaffected" formulas with free variables contained in $\mathbf{x}$. We sum this up in the following notion, which may be viewed as an abstract version of the "mosaics" of Németi 1992, Andréka and Németi 1994. Assume henceforth that a guarded F2-formula ϕ is given, where V is the set of variables occurring in ϕ.

DEFINITION. (i) Let F denote the set of all guarded formulas of length $\leqslant |\phi|$ that use only variables from V. Note that $\phi \in F$ and F is closed under taking subformulas as well as "alphabetic variants". Also, F is finite. (ii) An F-*type* is a subset Δ of F for which (a), (b), (c) below hold:

(a)	$\neg\psi \in \Delta$	iff	*not* $\psi \in \Delta$	whenever $\neg\psi \in F$
(b)	$\psi \,\&\, \xi \in \Delta$	iff	$\psi \in \Delta$ *not* $\xi \in \Delta$	whenever $\psi \,\&\, \xi \in F$
(c)	$[\mathbf{u}/\mathbf{y}]\psi$	implies	$\exists y\psi \in \Delta$	whenever $\exists y\psi \in F$.

Here $[\mathbf{u}/\mathbf{y}]\psi$ is the formula obtained from ψ by replacing each free variable in $\mathbf{y}$ with the corresponding variable in $\mathbf{u}$, simultaneously. (iii) Let $\mathbf{y}$ be a sequence of variables, and let Δ, Δ' be types. We say that Δ and Δ' are $\mathbf{y}$-*close*, in symbols, $\Delta =_{\mathbf{y}} \Delta'$, if Δ and Δ' have the same formulas with free variables disjoint from $\mathbf{y}$. (iv) A *quasimodel* is a set of F-types S such that, for each $\Delta \in S$ and each guarded formula $\exists\mathbf{y}(Q\mathbf{xy}\,\&\,\psi) \in \Delta$, there is a type $\Delta' \in S$ with $Q\mathbf{xy}$ and $\psi(\mathbf{x},\mathbf{y})$ in Δ' and $\Delta =_{\mathbf{y}} \Delta'$. We say that ϕ *holds in a quasi-model* if $\phi \in \Delta$ for some Δ in this model.

Clearly, if ϕ is satisfied by some model, then ϕ also holds in some quasi-model. The converse holds as well:

From Quasi-Models to Standard Models

Given a quasi-model $\mathbf{M}$ of the above kind, we can again define a standard model $\mathbf{N}$. We say that π is a *path* if $\pi = \langle \Delta_1, \phi_1, \ldots, \Delta_n, \phi_n, \Delta_{n+1}\rangle$ where Δ_1, Δ_{n+1} are types in $\mathbf{M}$, each formula ϕ_i is of the form $\exists\mathbf{y}(Q\mathbf{xy}\& \psi) \in \Delta_i$ and Δ_{i+1} is an alternative type as described above (i.e., $Q\mathbf{xy}$ and $\psi(\mathbf{x},\mathbf{y})$ in Δ_{i+1} and $\Delta_{i+1} =_{\mathbf{y}} \Delta_i$). We say that the variables in $\mathbf{y}$ *changed their values* from Δ_i to Δ_{i+1}, whereas the others did not. Finally, a variable z is called *new in a path* π if either $|\pi| = 1$ or z's value was changed at the last round in π. Now we are ready to define our model $\mathbf{N}$.

Objects in $\mathbf{N}$ are all pairs (π, z) where π is a path, and z is new in π. Next, we *interpret predicates* over these objects. We say that $I(Q)$ holds of the sequence of objects $\langle(\pi_j, x_j)\rangle_{j\in J}$ iff the paths π_j fit into one linear sequence under inclusion, with a maximal path π^* such that (i) the atom $Q\langle x_j\rangle_{j\in J} \in \Delta^*$ (the last type on π^*) and for no (π_j, x_j) does x_j change its value on the further path to the end of π^*. Finally, we also define a *canonical assignment* s_π for each path. We set $s_\pi(x) \stackrel{\text{def}}{=} (\pi', x)$ where π' is the unique subpath of π at whose end x was new, while it remained unchanged afterwards. Now, we formulate correctness of this construction via the following assertion – where last(π) is the last type on the path π.

LEMMA 4.6.2. *For all paths π in* $\mathbf{N}$, *and all formulas $\psi \in F$,*

$$\mathbf{N}, s_\pi \models \psi \quad \textit{iff} \quad \psi \in \text{last}(\pi).$$

Proof. Induction on relevant formulas y. ***Boolean cases*** are immediate, using the standard closure conditions for $\neg$ and & on types. ***Atoms***: here we do an example. (i) Suppose that $Qxyx \in$ last(π). The objects $s_\pi(x), s_\pi(y)$ were introduced at some maximal subsequence π^* of π.

Note that no x- or y-values changed on π after their introduction. Therefore, by the above transfer of "unaffected" formulas across successor types in these sequences, $Qxyx \in \text{last}(\pi^*)$. Then by the above definition, $I(Q)$ holds for the objects $s_{\pi^*}(x)$, $s_{\pi^*}(y)$, $s_{\pi^*}(x) = s_\pi(x)$, $s_\pi(y)$, $s_\pi(x)$. This means that $\mathbf{N}, s_\pi \models Qxyx$. (ii) Next, suppose that $\mathbf{N}, s_\pi \models Qxyx$. This means that $I(Q)$ holds for the objects $s_\pi(x)$, $s_\pi(y)$, $s_\pi(x) = s_{\pi^*}(x)$, $s_{\pi^*}(y)$, $s_{\pi*}(x)$. The picture is the same as in the previous case. By the definition again, we have that $Qxyx \in \text{last}(\pi^*)$. And once again by transfer of unaffected formulas, $Qxyx \in \text{last}(\pi)$. [In the presence of identity, an argument like this would have to be complicated – allowing for the same object to be marked by different variables on different paths.]

Finally, consider the case of bounded ***Existential Quantifiers*** $\exists\mathbf{y}(Q\mathbf{xy} \,\&\, \psi(\mathbf{x}, \mathbf{y}))$. (i) First, suppose that $\exists\mathbf{y}(Q\mathbf{xy} \,\&\, \psi(\mathbf{x}, \mathbf{y})) \in \text{last}(\pi)$. Then there is an extended path $\pi^+ \stackrel{\text{def}}{=} \pi$ *concatenated with* $\langle \exists\mathbf{y}(Q\mathbf{xy} \,\&\, \psi(\mathbf{x}, \mathbf{y})), \Delta' \rangle$, where Δ' is a successor type for Δ chosen as above with $Q\mathbf{xy}, \psi(\mathbf{x}, \mathbf{y}) \in \Delta'$ (and satisfying the transfer condition for unaffected formulas with free variables $\mathbf{x}$). All objects (π^+, y_i) with y_i in $\mathbf{y}$ are new here. By definition, the atomic guard predicate $I(Q)$ holds for the object tuples $s_{\pi+}(\mathbf{y})$, $s_{\pi+}(\mathbf{x}) (= s_\pi(\mathbf{x}))$. By the inductive hypothesis, we must have $\mathbf{N}, s_{\pi+} \models \psi(\mathbf{x}, \mathbf{y})$. Therefore, $\mathbf{N}, s_{\pi+} \models \exists\mathbf{y}(Q\mathbf{xy} \,\&\, \psi(\mathbf{x}, \mathbf{y}))$. From this we see, by $\mathbf{x}$-invariance in the standard model $\mathbf{N}$, that indeed $\mathbf{N}, s_\pi \models \exists\mathbf{y}(Q\mathbf{xy} \,\&\, \psi(\mathbf{x}, \mathbf{y}))$. (ii) Conversely, suppose that $\mathbf{N}$, $s_\pi \models \exists\mathbf{y}(Q\mathbf{xy} \,\&\, \psi(\mathbf{x}, \mathbf{y}))$. By the truth definition, there exist objects $d_i = (\pi_i, u_i)$ such that $\mathbf{N}, s_\pi{}^{\mathbf{y}}{}_{\mathbf{d}} \models Q\mathbf{xy} \,\&\, \psi(\mathbf{x}, \mathbf{y})$. (Here, $s_\pi{}^{\mathbf{y}}{}_{\mathbf{d}}$ is the assignment which is like s_π except for setting all y_i to d_i.) In particular, $I(Q)$ holds of the objects $s_\pi(\mathbf{x}), \mathbf{d}_i$. This leads to a simple picture of forking paths. The objects $s_\pi(\mathbf{x})$ were all introduced by stage π^* inside π, and then the objects $\mathbf{d}_i$ were (either interpolated among them or) added to form a maximal sequence π^+ where the true atom $Q\mathbf{xy}$ holds at the end. The fork is such that $\mathbf{x}$-values do not change any more from π^* onward, whether toward π or π^+. (That this is the only relevant situation is where the atomic guard on our quantifier comes in essentially.) Now, a minor complication. Note that the variables u_i do not have to be the y_i. (We can be sure that they are not x_i, though, as the *original objects* $s_\pi(\mathbf{x}) = s_{\pi^*}(\mathbf{x})$ were involved in the true atom $Q\mathbf{xy}$.) Also, π^+ is such that $s_{\pi+}(u_i) = (\pi_i,\, u_i) = d_i$. Thus, the two assignments $s_\pi{}^{\mathbf{y}}{}_{\mathbf{d}}$ and $s_{\pi+}$ agree on $\mathbf{x}$, and for all $y_i \in \mathbf{y}$ we have $s_\pi{}^{\mathbf{y}}{}_{\mathbf{d}}(y_i) = d_i = s_{\pi+}(u_i)$. Then, by $\mathbf{N}, s_\pi{}^{\mathbf{y}}{}_{\mathbf{d}} \models Q\mathbf{xy} \,\&\, \psi$ and the above observations, we have $\mathbf{N}, s_{\pi+} \models [\mathbf{u}/\mathbf{y}]Q\mathbf{xy}, \mathbf{N}, s_{\pi+} \models [\mathbf{u}/\mathbf{y}]\psi$. By the inductive hypothesis then, $[\mathbf{u}/\mathbf{y}]\psi \in \text{last}(\pi^+)$. (Here we assume that our set of relevant formulas is closed under simultaneous substitutions, that do not increase syntactic complexity. For a proof, see the Remark below.) Also, from the initial description of π^+, we see at once that $[\mathbf{u}/\mathbf{y}]Q\mathbf{xy} \in \mathrm{last}(\pi^+)$ (by the interpretation of atomic predicates). By closure conditions (b), (c) in the definition of a type, one gets $\exists\mathbf{y}(Q\mathbf{xy}\,\&\,\psi(\mathbf{x},\ \mathbf{y})) \in \mathrm{last}(\pi^+)$. Finally, since no changes in **x**-values occurred on the fork from π^*, the transfer condition on successor types along paths ensures that this same formula is in $\mathrm{last}(\pi)$. ∎

REMARK. *Finite Variable Fragments are closed under Simultaneous Substitutions.*
Our proof assumes the finite set of relevant formulas is closed under simultaneous substitutions – without enlarging the set V of relevant variables. To show this, consider any substitution $[\mathbf{x} := f(\mathbf{x})]\phi$ in a k-variable fragment with variables $\mathbf{x} = x_1, \ldots, x_k$. Atomic replacements are straightforward. Also, we can push substitutions inside over Booleans. The only interesting case is when we encounter an existential quantifier: $[\mathbf{x} := f(\mathbf{x})]\exists x_j\psi$. Then, the assignment clause $x_j := f(x_j)$ has no effect, and so it can be omitted. Hence, in the remaining substitution σ, at least one variable x_k is not used at all on the right-hand side in any assignment. But then, the following formula is easily shown to be equivalent to the original one: $\exists x_k[x_j := x_k, \sigma]\psi$. This gives a simple recursive algorithm computing substitutions inside our fragment. (With function symbols, the result fails: witness the case of $[x := fxy]\exists yRxy$.) ∎

From the Lemma, the Theorem is immediate. Our original formula ϕ is satisfied in a standard model iff it has a quasi-model, and it is decidable whether ϕ has the latter. ∎

This proof may be extended to deal with the Guarded Fragment with *equality*. We also have an alternative route to the latter result, using the earlier-mentioned mosaics, and an excursion into generalized assignment models (cf. Andréka and Németi 1994). What both proof methods leave open is the *Finite Model Property*. As a by-product, we do obtain completeness with respect to finite modal-filtration-like abstract models, but not for standard models. In the follow-up publication Andréka, van Benthem and Németi 1996, we shall prove the full Finite Model Property of the Guarded Fragment, combining modal-style filtration methods with additional combinatorial analysis.

4.7. *Meta-Properties of Bounded Fragments*

Other earlier modal techniques may be generalized to these fragments just as well. We merely discuss a short list of results obtainable along the above lines. Detailed statements and proofs may be found in Andréka, van Benthem and Németi 1994A.

	Gentzeniz-ability	Decidability	Finite Model Property	Łoś–Tarski	Inter-polation
Fragment 1	+	+	+	+	+
Fragment 2	+	+	+	+	+
Fragment 3	–	–	–	+	+

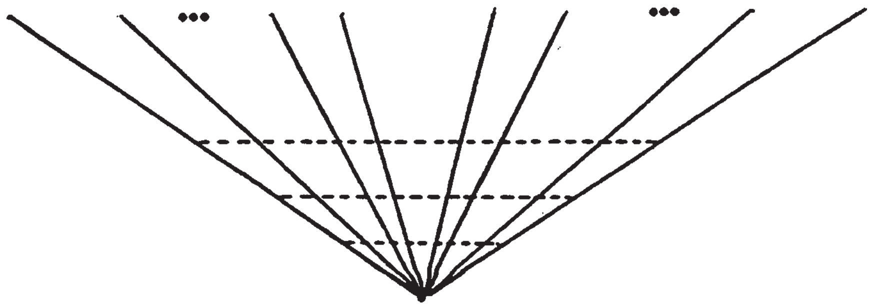
Here, the first negative outcome for Fragment 3 is true in a very strong sense – since the nonfinite-axiomatizability arguments of Andréka 1991 generalize to this case. The table also suggests further issues of interest. In particular, in Section 3 above, we formulated a desideratum to the effect that "finite-variable fragmentation works well". Indeed, we also have proofs for complete finite axiomatizability, Craig Interpolation and Łoś–Tarski for all k-variable levels of Fragments 1 and 2. By contrast, one can obtain negative counterparts to the above negative outcomes for all finite-variable levels of Fragment 3 (for all $k > 1$). All this may be proved by suitably "relativized" versions of earlier results. These negative findings even carry over to all finite-variable fragments F3.k $(k > 1)$. We illustrate this relativization method by one example.

EXAMPLE. Failure of Submodel Preservation in F3.3.
Consider the counter-example ϕ to the Submodel Preservation Theorem from Section 3. Its relativization ϕ^U to some new unary predicate U is in Fragment 3.

It is easy to see that ϕ^U, too, is preserved under submodels. Now, suppose that it had a universal equivalent α in the three-variable fragment of Fragment 3. This cannot happen, because we can now compare the two models constructed in the proof of Claim 4 in Section 3.4, both having U as the universal predicate. There is a 3-simulation from one to the other, which refutes the adequacy of α just as before. Similar arguments turn out to work for most other meta-properties that we have considered.

4.8. *A General Picture*

More generally, the bounded fragments serve as a point of departure for a new hierarchy in predicate logic, orthogonal to the finite-variable levels:



What is the natural layering here? In addition to the degrees of freedom in the above restriction schema, one can vary the format of the restricting predicates themselves. For example, Temporal Logic typically involves "Betweenness": $\exists z(Rxz \,\&\, Rzy \,\&\, \phi(z))$. Or, Dynamic Logic has modal predicate operations: $\exists z(\mathbf{O}(R,...)xz \,\&\, \phi(z))$. Some operations **O** produce formulas inside our restricted fragments (e.g., sequential composition and choice), others lead outside of them (e.g., predicate intersection and complement violate "bisimulation safety": van Benthem 1993). Such more expressive fragments, too, can be analyzed via our previous modal techniques. To conclude, however, we list one particular simple problem where our analysis has failed so far. The Guarded Fragment used only atomic guards – and to obtain decidability, this requirement cannot be relaxed in general (a counter-example is to follow in Section 5.2). But the minimal logic of the temporal operators "Since" and "Until" is known to be decidable – and yet, their natural first-order translations involve *composite* atomic *guards* for betweenness. Thus, our decidability theorem for the Guarded Fragment does not explain this important modal fact. Thus, we are left with a new question. Which Boolean combinations of guards are harmless for decidability? (Some new developments here are reported in van Benthem 1997.)

5. Generalized first-order semantics

5.1. *A Modal View of Tarski Semantics*

In our discussion so far, the main impact of modal techniques on standard logic has been the identification of first-order *fragments* that behave well over standard Tarski models. But one can also turn the tables, and interpret the full language of first-order predicate logic over generalized first-order models – where assignments (or objects) may only be "available" subject to certain constraints (regulated by certain accessibility relations R), retaining the core of Tarski's truth definition. This proposal originates in algebraic logic (cf. Németi 1981, 1990, 1992 on relativizations of representable algebras of logics, with applications to the field of logic itself). It has been pursued since with modal techniques as well (cf. Venema 1995A, Marx 1995, van Benthem 1994B). This second research program is not the subject of this paper, but it is closely related in motivation and content. Here is a sketch of its main features and outcomes.

Modal first-order models are triples of the form $\mathbf{M} = (S, \{R_x\}_{x\in \mathrm{VAR}}, I)$ where S is a set of "states", R_x a binary relation between states for each variable x, and I is an "interpretation function" giving a truth value to all atomic formulas $Px, Rxy, Ryx, \ldots,$ in each state α. This abstract modal format turns out to be all that is needed to set up the standard inductive truth definition for first-order logic:

$$\begin{array}{lll} \mathbf{M}, \alpha \models P\mathbf{x} & \text{iff} & I(\alpha, P\mathbf{x}) \\ \text{Boolean connectives} & & \textit{as usual} \\ \mathbf{M}, \alpha \models \exists x\phi & \text{iff} & \textit{for some } \beta : R_x\alpha\beta \textit{ and } \mathbf{M}, \beta \models \phi. \end{array}$$

Thus, predicate logic becomes a poly-modal logic with $\exists x$ as an existential modality. In this full generality, models may deviate considerably from the standard paradigm. Notably, the interpretation of intuitively related atoms like Rxy and Ryx may become completely independent. And the same holds for such related formulas as Px and $\exists yPx$. Nevertheless, one can easily enforce such desired behaviour by additional stipulations (cf. Andréka, Gergely and Németi 1977, Németi 1986: section on "NA", Németi 1992: Sections 7, 12, or Németi 1994: Section 8). In particular, also, one might insist that the binary relations R_x be *equivalence relations*, as they are in standard Tarski models. This happens in a natural "half-way house", in between modal first-order models and standard Tarski semantics, called *generalized assigment models*. Here S is some family of ordinary assignments (not necessarily the full function space D^{VAR}), and the accessibilities R_x are the standard relations $=_x$ of identity up to x-values. The truth condition for the existential quantifier then runs as follows:

$$\begin{array}{l} \mathbf{M}, S, \alpha \models \exists x\phi \quad \text{iff} \\ \textit{for some } \beta \in S : \alpha =_x \beta \textit{ and } \ \mathbf{M}, \beta \models \phi. \end{array}$$

The possible assignment gaps in generalized assignment models have positive virtues. They model the natural phenomenon of "dependencies" between variables: which occurs when changes in value for one variable x may induce, or be correlated with, changes in value for another variable y. (Examples in natural deduction and probabilistic reasoning are in Fine 1985, van Lambalgen 1991.) Dependence cannot be modeled in standard Tarskian semantics, wich modifies values for variables completely arbitrarily. Finally, to get an even closer approximation to the standard first-order language, one must introduce substitutions in the models (see below).

There is a growing literature on this generalized semantics. Modal first-order models validate a "minimal predicate logic", which is really just the minimal poly-modal logic, with all the positive properties studied in this paper (including decidability). On top of that lies a landscape of further calculi, all the way up to full predicate logic: which now becomes the particular (undecidable) mathematical theory of full function-space assignment models. The modal logic of the above generalized assignment models is an interesting intermediate possibility (called "cylindric-relativized-set algebras" in the algebraic literature), which is decidable and has positive meta-properties (Loś–Tarski, Craig Interpolation). Natural extensions arise by imposing constraints on admissible assignments, such as "local squareness" or the "patchwork property" (cf. Németi 1992). Results on completeness, correspondence and interpolation for modal first-order logics in this sub-classical landscape, as well as representation theorems for its abstract models, may be found in Németi 1992 and other cited papers (e.g., van Benthem 1994B, Marx 1995.) One novel feature of this approach is the introduction of *new vocabulary*, reflecting distinctions not usually found in first-order logic. Examples are irreducibly *polyadic quantifiers* $\exists \mathbf{y}$ binding tuples of variables $\mathbf{y}$, or explicit modal calculi of *substitutions* (cf. van Lambalgen and Simon 1994, Andréka and Németi 1994, Németi 1994, Venema 1995A). Issues may be subtle. For example, whether Interpolation holds depends on the choice of vocabulary. (Madarasz 1995 shows that in some generalized model classes, Interpolation demands introduction of a "universal modality" ranging over all states, accessible or not. Likewise, Andréka and Németi 1994 show that polyadic quantifiers are essential to obtain finite axiomatizability.)

ILLUSTRATION. *Modal Analysis of Substitutions*

For each variable x, modal first-order models had an accessibility relation R_x, corresponding to what is called "random assignment" in dynamic logic. Next, we can introduce "determinate assignment" mirroring the syntactic operation of substitution. Here is a way of doing this. For each pair of variables x, y, introduce a new unary modality S_{xy} saying that, for each formula ϕ of predicate logic, $S_{xy}\phi$ is equivalent to the formula $[y/x]\phi$ with all free occurrences of x replaced by y in the usual way.

That is, $S_{xy}\phi$ is equivalent with $\exists x(x = y \,\&\, \phi)$. Modal first-order models now carry extra accessibility relations $A_{x,y}$ that can be subjected to constraints "corresponding to"' substitution axioms in the modal sense (van Benthem 1984). (1) $A_{x,y}$ is a function (this reflects commutation of S_{xy} with the Booleans). (2) $A_{x,y}$ is contained in R_x (axiom $S_{xy}\phi \to \exists x\phi$). (3) As a function, $A_{x,y}A_{x,y} = A_{x,y}$: this reflects $S_{xy}S_{xy}\phi \leftrightarrow S_{xy}\phi$. (4) Likewise, $S_{xy}S_{yx}\phi \leftrightarrow S_{xy}\phi$ reflects $A_{x,y}A_{y,x} = A_{x,y}$. (5) Finally, the interpretation function I can be restricted to satisfy atomic substitution laws like $S_{xy}Rxz \leftrightarrow Ryz$. This modal logic displays all positive properties of the basic one. For generalized assignment models, these definitions become more concrete.

5.2. *Back-and-Forth between Modal Logic and Predicate Logic*

Comparing the main thrust of this paper and the program outlined in Section 5.1, two main approaches emerge towards "taming" classical first-order logic: i.e., localizing what may be called a well-behaved decidable "core part". One can either use standard semantics over nonstandard language fragments, or use nonstandard generalized semantics over the full standard first-order language. The former approach is more "syntactical" in nature, the latter more "semantical". (Eventually, as so often in logic, this distinction is relative. For instance, one can also translate "semantic" modal discourse about the above modal first-order models into a restricted syntactic fragment of a *two-sorted* first-order logic, with direct reference to both "individuals" and "states". But also conversely,..., etcetera.) More specifically, evident technical analogies exist between existing proof methods for generalized semantics in the sense of Section 5.1 and those of the present paper. We feel that there is a mathematical duality lurking in the background here, largely unexplored – which we illustrate by some simple observations. In particular, our earlier analysis of bounded first-order fragments may be used to derive results about generalized assignment semantics, or equivalently, about relativized cylindric algebras (i.e., Crs-models). These new applications of our results provide a uniform perspective on the earlier literature.

From Bounded Fragments to Cylindric Algebra and Generalized Models

Consider any k-variable language $L\{x_1,\ldots,x_k\}$. Let R be a new k-ary predicate. Here is a translation tr_g from k-variable formulas to guarded first-order formulas:

> Global Relativization
> $\mathrm{tr}_g(\phi)$ arises from ϕ by relativization of all its quantifiers
> to the same atom $Rx_1\ldots x_k$.

Next, we define a corresponding operation on models. Let **M** be any generalized assignment model for $L\{x_1, \ldots, x_k\}$ (as yet without the new predicate symbol R).

> Restricted Standard Models
> The standard model $\mathbf{M}_{\text{rest}}$ is **M**, viewed as a standard model, and expanded with the following interpretation for the new predicate: $R(d_1, \ldots, d_k)$ iff the assignment $x_i := d_i (1 \leqslant i \leqslant k)$ is available in **M**.

The purpose of this construction shows in the following fact.

PROPOSITION 5.2.1. *For all available assignments α in* **M**, *and all formulas ϕ,*

$$\mathbf{M}, \alpha \models \phi \quad \text{iff} \quad \mathbf{M}_{\text{rest}}, \alpha \models \text{tr}_g(\phi).$$

Proof. Induction on first-order formulas. The crucial case is the existential quantifier. In particular, suppose that $\mathbf{M}_{\text{rest}}, \alpha \models \text{tr}_g(\exists x_i \phi) = \exists x_i (R x_1 \ldots x_k \,\&\, \text{tr}_g(\phi))$. Then, there exists a satisfying k-tuple of objects in R for $\text{tr}_g(\phi)$, which corresponds to an available assignment in **M** which is an i-variant of α. That is, $\mathbf{M}, \alpha \models \exists x_i \phi$. ■

As a consequence, one can effectively reduce universal validity over all generalized assignment models (i.e., in Crs) to standard validity in Fragment 2.

COROLLARY 5.2.2. $\models_{\text{gen'd}} \phi$ *iff* $\models_{\text{standard}} R x_1 \ldots x_k \to \text{tr}_g(\phi)$.

Proof. "Only if". If ϕ has a generalized counter-example $(\mathbf{M}, \alpha)$, then the above model $\mathbf{M}_{\text{rest}}$ falsifies $R x_1 \ldots x_k \to \text{tr}_g(\phi)$. "If". Suppose, conversely, that the latter formula has a standard counter-example $(\mathbf{M}, \alpha)$. Now define a corresponding generalized model $\mathbf{M}_g$ by retaining only those assignments whose values for $x_1, \ldots, x_k$ stand in the relation $R_{\mathbf{M}}$ (in particular, the falsifying assignment α itself remains available). Then ϕ is falsified in $\mathbf{M}_g$ by α as above. ■

This result provides a new "modal" proof for the following theorem (cf. Nemeti 1992).

THEOREM 5.2.3. *Validity in Crs is decidable, and Crs has the finite model property.*

Proof. This follows from the corresponding results for Fragment 2. ■

This translation is easily extended to a predicate logic with *polyadic quantifiers* over generalized assignment models (referring to assignments changed at some tuple of arguments simultaneously). One translates a polyadic formula $\exists x_{i1} \ldots x_{im}\phi$ to the guarded formula $\exists x_{i1} \ldots x_{im}$ $(Rx_1 \ldots x_k$ & *translation* $(\phi))$, where R is again our new predicate symbol. (This reduction highlights the fact that the Guarded Fragment *itself* has a polyadic quantification schema, not reducible to single existential steps.) Also, the translation works for the full first-order language at once, by a slightly modified model construction. (One assigns a dummy to all but finitely many variables. Cf. Andréka, van Benthem and Németi 1994.) There is still more to the above analysis. Special classes of generalized assignment models have arisen by imposing more specific *constraints* on admissible assignments. The first-order theory of such classes, too, will be decidable, provided their additional conditions can be stated in first-order forms translatable *into the Guarded Fragment*. This applies, e.g., to the "D_α" or "G_α" of Németi 1992. In particular, we get a new proof for a result from Németi 1986, 1992 (Marx 1995 has a modal proof) – where *locally square* models are closed under permutations and substitutions of values in the range of any available assignment:

THEOREM 5.2.4. *Universal validity is decidable on the class of generalized assignment models which are locally square.*

Proof. The reason is that the requirements for being locally square are all expressible inside Fragment 2. Here is an example of the relevant kind of formula:

$$\forall xy(Rxy \rightarrow (Ryx \,\&\, Rxx \,\&\, Ryy)). \qquad \blacksquare$$

By contrast, we know validity is undecidable over generalized assignment models satisfying the so-called "Patchwork Property". Again, this checks out. In first-order form, the latter constraint involves statements like

$$\forall xyzuv((Rxyz \,\&\, Ruyv) \rightarrow (Rxyv \,\&\, Ruyz)).$$

These are not in Fragment 2: variable inclusion holds from matrix to restriction, but the latter is not one single atom. (Thus, we cannot allow free Boolean combinations of guards in decidability results.) This style of analysis is quite powerful, and it can be used to predict decidability of many other combinations of algebraic axioms on top of Crs, as long as their complete frame properties fall inside Fragment 2. We conclude with a natural converse question. Can one also derive the behaviour of our modal bounded fragments from algebraic results about generalized assignment models? In particular, is there a *converse reduction* going from standard validity of formulas ψ in Fragment 2 to generalized validity of suitable formulas $\mathrm{red}(\psi)$ over generalized assignment models? At least, the identical translation does not work. The following Fragment 2 formula is valid, but it is not in Crs:

$$\exists x(Ax \,\&\, \exists y(Rxy \,\&\, Ay)) \to \exists y(Ay \,\&\, \exists x(Rxy \,\&\, Ax)).$$

We do have some partial converse results, that work for suitably "uniformly relativized" formulas in Fragment 2, and one also finds striking similarities in proofs for meta-properties of the two calculi – but we shall leave this matter open here.

Digression on Dependency Semantics

The preceding analysis suggests an analogy between generalized assignment models and the "dependency models" for generalized quantifiers $Qx \bullet \phi$ proposed in van Lambalgen 1991, Alechina and van Benthem 1993. These quantifiers are read there as stating the existence of some object "depending" on the range of the assignment so far. The two semantics are evidently related in spirit, but not isomorphic. Generalized assignment semantics validates unrestricted Monotonicity for existential quantifiers (i.e., $\exists x\phi \to \exists x(\phi \vee \psi)$), whereas dependency semantics does not. (It only has Monotonicity and Distribution for suitably "balanced" variables.) On the other hand, dependency semantics validates the unrestricted axiom $\exists x\phi \to \phi$ (x not free in ϕ), which does not hold on all generalized assignment models. We explain the situation. Dependency semantics arises from first-order logic through a "local translation" tr_l much like the above "global translation" tr_g, but with a delicate difference. At each subformula $\exists x_i \psi$, one only relativizes to an atom $R\mathbf{x}$ where $\mathbf{x}$ enumerates all free variables of the local context ψ. This difference explains all the deviant behaviour. For example, consider the effect of the two translations on Monotonicity. The global one makes this principle valid in Fragment 2, whereas the local one does not, witness:

$$\forall y(\forall x(Ax \to Bxy) \to (\exists x Ax \to \exists x Bxy))$$

translation tr_g

$$Rxy \to \forall y(Rxy \to (\forall x(Rxy \to (Ax \to Bxy)) \to$$
$$\to (\exists x(Rxy \,\&\, Ax) \to \exists x(Rxy \,\&\, Bxy))))$$

translation tr_l

$$\forall y(Ry \to (\forall x(Rxy \to (Ax \to Bxy)) \to$$
$$\to (\exists x(Rx \,\&\, Ax) \to \exists x(Rxy \,\&\, Bxy)))).$$

Here, our general results apply. Tr_l, like tr_g, takes first-order formulas to formulas inside the Guarded Fragment. Thus, we derive the decidability results of Alechina 1995, and we predict decidability for stronger dependency logics having characteristic frame conditions inside Fragment 2 (cf. Alechina and van Lambalgen 1995).

Bisimulation and Ehrenfeucht Games

Next, consider basic model equivalences in the two domains. As pointed out in van Benthem 1991, de Rijke 1993, *bisimulation* stands to modal logic as Ehrenfeucht games (or rather, "partial isomorphism") to standard first-order logic. As before, comparisons in the present setting are promising, though not yet conclusive. A modal bisimulation $\equiv$ between generalized assignment models which relates assignments α, β only if they have the same domain, induces an obvious relation PI between (those tuples of objects that form) the ranges of α and β. PI is a family of partial isomorphisms. It satisfies the usual back-and-forth extension conditions for "partial isomorphism" iff our generalized first-order model satisfies an Update Postulate: "For any object, each assignment has an extension assigning that object to some fresh variable". Then, the bisimulation clause for the latter variable will do the job. Conversely, any partial isomorphism PI between two models induces a modal bisimulation $\equiv$ between partial assignments over them by checking whether their ranges are a matching pair of object tuples in PI. But the more interesting comparison may lie in the differences. Generalized assignment models suggest a *change* in the standard model-theoretic notion of partial isomorphism, using a finer-grained match between partial assignments (rather than flat sequences of objects) in the two models.

We conclude that the mathematical analogies uncovered so far between nonstandard generalized assignment models in Cylindric Algebra and standard possible worlds semantics for generalized modal logics have already proved of evident benefit.

6. Further Directions

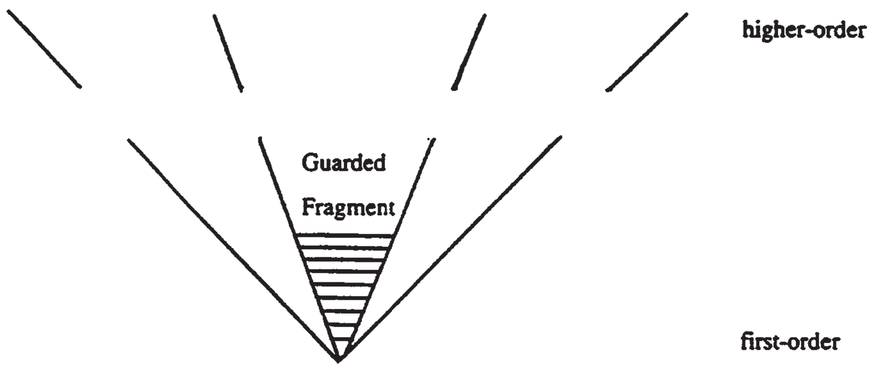
At various places so far, we noted open research questions. These concerned both technical elaboration within our framework (details of Łoś–Tarski Theorems in Section 3, meta-properties of further first-order fragments: cf. Section 4) and extensions to a broader environment (e.g., mathematical connections with generalized assignment or dependency semantics: cf. Section 5). We briefly mention some further directions.

6.1. *Special Frame Constraints*

Perhaps the first question that will occur to modal logicians is this. Basic modal logic is usually enriched with special frame conditions, as with S4, S5 or other calculi, imposing reflexivity, symmetry, transitivity or other natural frame conditions. From this viewpoint, we have been concerned merely with "minimal modal logics" of our frame classes. (Of course, we hope to have shown how attractive minimal logics are, when viewed as bounded fragments. . .) Note, e.g., that the generalized assignment semantics of Section 5 employs S5-like frame conditions. How are our results affected by imposing special frame constraints? Little is known. The results of Alechina 1995 suggest that *permutation* of arguments in guards is unproblematic for decidability. Also, we observed that those frame constraints which can be expressed as guarded formulas do not endanger decidability. Examples are reflexivity and symmetry. But already transitivity is not F2-definable, and hence we lack an explanation so far for its unproblematic behaviour in the frame theory of basic modal logic.

6.2. *Infinitary Extensions*

"Restriction" works just as well in fragments of *higher-order* languages, such as $L_{\omega 1 \omega}$ or $L_{\infty\omega}$ or second-order logic. We can transfer our "modal hierarchy" up to here:



Possible analogies to be explored lie in second-order logic (cf. Gallin 1975 on the good behaviour of restricted "extensional fragments") or admissible set theory (cf. Barwise 1975). Possibly significant here is the simple folklore characterization of bisimulation between models via their elementary equivalence in the $L_{\infty\omega}$-version of modal logic with arbitrary set conjunctions and disjunctions (van Benthem and Bergstra 1993). (De Rijke 1993 present more sophisticated results, e.g., inside $L_{\omega 1 \omega}$. Cf. also van Benthem 1993, Barwise and Moss 1995 exploring infinitary modal logic.) Can we also generalize other results from the above, such as the Łoś–Tarski Theorem?

6.3. *Extended Modal Logics*

We have ignored enriched modal operator formalisms in the style of Gargov, Passy and Tinchev 1987, de Rijke 1993, which allow both modest additions (e.g., a “difference modality”) and strong extensions in expressive power (such as the temporal logics of “Since” and “Until”). It would be of interest to extend our analysis in this direction (Section 4.7). This would fit in with the move towards richer vocabularies in generalized assignment semantics, including the polyadic quantifiers and substitutions of Section 5. Thus, a concern with fragments by no means implies logical poverty.

6.4. *Alternative Semantics*

Further options in modelling first-order predicate logic may be relevant. For example, starting from a computational motivation in “dynamic semantics”, Hollenberg and Vermeulen 1994 propose a stack-based account of first-order logic which makes the latter’s two-variable fragment as powerful as the whole language. (Thus, two variables over finite sequences are as good as arbitrary finite numbers of variables over single objects.) It remains to be seen how our considerations fare in such a semantics. Other first-order translations reflect yet different semantics (cf. the “path formulas” of Ohlbach 1991).

Acknowledgements

Over the years, participants at the Amsterdam Modal Logic Colloquium (organized by Maarten de Rijke and Natasha Alechina) provided useful comments on successive versions of this work, including Peter van Emde Boas and Tim Fernando. We were also encouraged by the lively participants at the FoLLI Summer School, Lisbon 1993. Finally, we wish to thank a group of inspiring colleagues and students who responded to this work – in particular, Natasha Alechina, Jaap van der Does, Marco Hollenberg, Dick de Jongh, Michiel van Lambalgen, Judith Madarasz, Maarten Marx, Szabolcs Mikulas, Eric Rosen, Maarten de Rijke, Ildiko Sain, Andras Simon, Yde Venema, Kees Vermeulen, Albert Visser and Scott Weinstein.

The research of Andréka and Németi was supported by the Hungarian National Fund for Scientific Research, OTKA grants No. T16448 and T7255.

REFERENCES

Alechina, N., On a decidable generalized quantifier logic corresponding to a decidable fragment of predicate logic, *Journal of Logic, Language and Information* **4**(3) (1995), 177–189.

Alechina, N. and van Benthem, J., *Modal Logics of Structured Domains*, Report ML-93-02, Institute for Logic, Language and Computation, University of Amsterdam. To appear in M. de Rijke (ed.), *Advances in Intensional Logic*, Kluwer, Dordrecht, 1997.

Alechina, N. and van Lambalgen, M., Correspondence and completeness for generalized quantifiers, *Bulletin of the Interest Group in Pure and Applied Logic* **3** (1995), 167–190.

Andréka, H., *Complexity of Equations Valid in Algebras of Relations*, D. Sc. Dissertation, Hungarian Academy of Sciences, Budapest, 1991. To appear in *Annals of Pure and Applied Logic*.

Andréka, H., van Benthem, J. and Németi, I., *What Makes Modal Logic Tick?*, manuscript, Institute for Logic, Language and Computation, University of Amsterdam, 1992.

Andréka, H., van Benthem, J. and Németi, I., *Interpolation and Definability Properties of Finite Variable Fragments*, Mathematical Institute, Hungarian Academy of Sciences, Budapest, 1993.

Andréka, H., van Benthem, J. and Németi, I., *Modal Logic and the Finite Variable Hierarchy*, manuscript, Institute for Logic, Language and Computation, University of Amsterdam, Mathematical Institute, Hungarian Academy of Sciences, Budapest, 1994.

Andréka, H., van Benthem, J. and Németi, I., Back and forth between modal logic and classical logic, *Bulletin of the Interest Group in Pure and Applied Logics* **3** (1995A), 685–720, London and Saarbruecken.

Andréka, H., van Benthem, J. and Németi, I., Submodel preservation theorems in finite variable fragments, in A. Ponse, M. de Rijke and Y. Venema (eds), *Modal Logic and Process Algebra*, CSLI Lecture Notes 53, CSLI Publications, Stanford, 1995B, pp. 1–11.

Andréka, H., van Benthem, J. and Németi, I., Strong decidability and other positive results for algebras of sets of sequences, in preparation, 1996.

Andréka, H., Düntsch, I. and Németi, I., Expressibility of properties of relations, *Journal of Symbolic Logic* **60**(3) (1995), 970–991.

Andréka, H., Gergely, T. and Németi, I., On universal algebraic construction of logics, *Studia Logica* **36** (1977), 9–47.

Andréka, H. and Németi, I., *Crs With, and Without Substitutions*, manuscript, Mathematical Institute, Hungarian Academy of Sciences, Budapest, 1994.

Barwise, J., *Admissible Sets and Structures*, Springer Verlag, Berlin, 1975.

Barwise, J. and Moss, L., *Vicious Circles: On the Mathematics of Non-Wellfounded Phenomena*, CSLI Publications, Stanford, 1996.

van Benthem, J., *Modal Correspondence Theory*, dissertation, Mathematical Institute, University of Amsterdam, 1976.

van Benthem, J., Correspondence theory, in D. Gabbay and F. Guenthner (eds), *Handbook of Philosophical Logic*, vol. II, Reidel, Dordrecht, 1984.

van Benthem, J., *Modal Logic and Classical Logic*, Bibliopolis, Napoli, 1985.

van Benthem, J., *Language in Action*, North-Holland, Amsterdam, 1991.

van Benthem, J., *Programming Operations that are Safe for Bisimulation*, CSLI Report 93-179, Center for the Study of Language and Information, Stanford, 1993. To appear in *Studia Logica*, 1997.

van Benthem, J., *A Modal Perspective on Process Operations*, manuscript, Institute for Logic, Language and Computation, University of Amsterdam, 1994A.

van Benthem, J., *Modal Foundations for Predicate Logic*, CSLI Research Report 94-191, Center for the Study of Language and Information, Stanford University. To appear in E. Orlowska (ed.), *Logic at Work*, Memorial Volume for Helena Rasiowa, Kluwer Academic Publishers, Dordrecht, 1994B.

van Benthem, J., Temporal logic, in D. Gabbay, Ch. Hogger and J. Robinson (eds), *Handbook of Logic in AI and Logic Programming*, Vol. 4, Oxford University Press, Oxford, 1995, pp. 241–350.

van Benthem, J., Exploring Logical Dynamics, in *Studies in Logic, Language and Information*, CSLI Publications, Stanford, 1996.

van Benthem, J., Dynamic Bits and Pieces, Report LP-97-01, ILLC, University of Amsterdam.

van Benthem, J. and Bergstra, J., Logic of transition systems, *Journal of Logic, Language and Information* **3**(4) (1995), 247–283.

Chang, C. and Keisler, H., *Model Theory*, North-Holland, Amsterdam, 1973.

Dawar, A., *Feasible Computation Through Model Theory*, dissertation, Computer Science Department, University College of Swansea, 1993.

Fine, K., Natural deduction and arbitrary objects, *Journal of Philosophical Logic* **14** (1985), 57–107.

Fitting, M., Basic modal logic, in D. Gabbay, Ch. Hogger and J. Robinson (eds), *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. I, Oxford University Press, Oxford, 1993.

Gabbay, D., Expressive functional completeness in tense logic, in U. Mönnich (ed.), *Aspects of Philosophical Logic*, Reidel, Dordrecht, 1981, pp. 91–117.

Gallin, D., *Intensional and Higher-Order Modal Logic*, North-Holland, Mathematics Studies 19, Amsterdam, 1975.

Gargov, G., Passy, S. and Tinchev, T., Modal environment for Boolean speculations, in D. Skordev (ed.), *Mathematical Logic and Applications*, Plenum Press, New York, 1987, pp. 253–263.

Henkin, L., *Logical Systems Containing Only a Finite Number of Symbols*, Séminaire de Mathématiques Supérieures 21, Les Presses de l'Université de Montréal, Montréal, 1967, 48 pp.

Henkin, L., Monk, J. D. and Tarski, A., *Cylindric Algebras, Part I*, North-Holland, Amsterdam, 1971.

Henkin, L., Monk, J. D. and Tarski, A., *Cylindric Algebras, Part II*, North-Holland, Amsterdam, 1985.

Hennessy, M. and Milner, R., Algebraic laws for nondeterminism and concurrency, *Journal of the ACM* **32** (1985), 137–161.

Hodkinson, I., Finite H-dimension does not imply expressive completeness, *Journal of Philosophical Logic* **23** (1994), 535–573.

Hollenberg, M. and Vermeulen, K., *Counting Variables in a Dynamic Setting*, Logic Group Preprint Series 125, Filosofisch Instituut, Rijksuniversiteit, Utrecht, 1994.

Immerman, N., Number of quantifiers is better than number of tape cells, *Journal of Computer and Systems Sciences* **22** (1981), 65–72.

Immerman, N., Relational queries computable in polynomial time, *Proceedings 14th Annual ACM Symposium on the Theory of Computing*, 1982, pp. 147–152.

Immermann, N. and Kozen, D., Definability with bounded number of bound variables, *Proceedings 2d IEEE Symposium on Logic in Computer Science*, 1987, pp. 236–244.

Kolaitis, F. and Väänänen, J., Generalized quantifiers and pebble games on finite structures, *Proceedings of the 7th IEEE Symposium on Logic in Computer Science*, 1992, pp. 348–359. (Full version to appear in *Annals of Pure and Applied Logic*.)

Kracht, M., How completeness and correspondence theory got married, in M. de Rijke (ed.), *Diamonds and Defaults*, Kluwer, Dordrecht, 1993, pp. 175–214.

Kracht, M. and Wolter, F., Properties of independently axiomatizable bimodal logics, *Journal of Symbolic Logic* **56** (1991), 1469–1485.

van Lambalgen, M., Natural deduction for generalized quantifiers, in J. van der Does and J. van Eyck (eds), *Generalized Quantifiers, Theory and Applications*, Dutch PhD Network for Language, Logic and Information, Amsterdam, 1991, pp. 143–154. (Appeared with CSLI Publications, Stanford, 1996.)

van Lambalgen, M. and Simon, A., *Axiomatization of Cylindric-Relativized Set Algebras with Polyadic Quantifiers*, Institute for Logic, Language and Computation, Amsterdam, and Mathematical Institute, Hungarian Academy of Sciences, Budapest, 1994.

Madaràsz, J., *Craig Interpolation in Algebraizable Logics*, submitted. Mathematical Institute, Budapest, 1994, preprint.

Marx, M., *Arrow Logic and Relativized Algebras of Relations*, PhD dissertation, CCSOM and ILLC, University of Amsterdam, 1995.

Mason, I., The meta-theory of the classical propositional calculus is not axiomatizable, *Journal of Symbolic Logic* **50** (1985), 451–457.

Monk, J. D., Nonfinitizability of classes of representable cylindric algebras, *Journal of Symbolic Logic* **34** (1969), 331–343.

Németi, I., Connections between cylindric algebras and initial algebra semantics of CF languages, in B. Domolki and T. Gergely (eds), *Mathematical Logic in Computer Science* (Proceedings Conference Salgotarjan 1978), Colloq. Math. Soc. J. Bolyai, Vol. 26, North-Holland, Amsterdam, 1981, pp. 561–605.

Németi, I., Cylindric-relativized set algebras have strong amalgamation, *Journal of Symbolic Logic* **50**(3) (1985), 689–700.

Németi, I., *Free Algebras and Decidability in Algebraic Logic*, Third Doctoral Dissertation, Mathematical Institute, Hungarian Academy of Sciences, Budapest, 1986.

Németi, I., On cylindric algebraic model theory, in C. Bergman, R. Maddux and D. Pigozzi (eds), *Algebraic Logic and Universal Algebra in Computer Science* (Proceedings Conference Ames 1988), Lecture Notes in Computer Science, Vol. 425, Springer Verlag, Berlin, 1990, pp. 37–76.

Németi, I., Algebraizations of quantifier logics: an introductory overview, *Studia Logica* **50**(3–4) (1991), 485–569. (See also Németi 1994.)

Németi, I., Decidability of weakened versions of first-order logic. First version in *Pre-Proceedings Logic at Work,* CCSOM, Amsterdam. Appeared as “Decidable Versions of First-Order Logic and Cylindric-Relativized Set Algebras”, in L. Csirmaz, D. Gabbay and M. de Rijke, eds., 1995, *Logic Colloquium 92.Veszprem, Hungary*, Studies in Logic, Language and Information, CSLI Publications, Stanford, 1992, pp. 177–241.

Németi, I., *Algebraizations of Quantifier Logics*, extended version, to appear in *Proceedings Summer School on Algebraic Logic*, Budapest, 1994.

Ohlbach, H.-J., Semantic-based translation methods for modal logics, *Journal of Logic and Computation* **1**(5) (1991), 691–746.

Pigozzi, D., Amalgamation, congruence extension, and interpolation properties in algebras, *Algebra Universalis* **1** (1971), 269–349.

Roorda, D., *Resource Logics. A Proof-Theoretic Investigation*, dissertation, Institute for Logic, Language and Computation, University of Amsterdam, 1991.

Rosen, E. and Weinstein, S., *Preservation Theorems in Finite Model Theory*, The Institute for Research in Cognitive Science, University of Pennsylvania, Philadelphia, 1995.

de Rijke, M., *Extending Modal Logic*, dissertation, Institute for Logic, Language and Computation, University of Amsterdam, 1993.

Sain, I., Beth's and Craig's properties via epimorphisms and amalgamation in algebraic logic, in C. Bergman, R. Maddux and D. Pigozzi (eds), *Algebraic Logic and Universal Algebra in Computer Science*, Lecture Notes in Computer Science 425, Springer Verlag, Berlin, 1990, pp. 209–226.

Sain, I. and Simon, A., *Negative and Positive Results on Definability Properties in Finite Variable Fragments*, manuscript, Mathematical Institute, Hungarian Academy of Sciences, Budapest, 1993.

Sain, I. and Thompson, R. J., *Strictly Finite Schema Axiomatization of Quasi-Polyadic Algebras*, Algebraic Logic, North-Holland, 1990, pp. 539–572.

Schütte, K., Der Interpolationssatz der Intuitionistischen Prädikatenlogik, *Mathematische Annalen* **148** (1962), 192–200.

Tarski, A. and Givant, S. R., *A Formalization of Set Theory Without Variables*, AMS Colloquium Publications, Vol. 41, Providence, Rhode Island, 1987.

Venema, Y., *Many-Dimensional Modal Logic*, dissertation, Mathematical Institute, University of Amsterdam, 1991.

Venema, Y., A modal logic for quantification and substitution, in L. Csirmaz, D. Gabbay and M. de Rijke (eds), 1995, *Logic Colloquium 92.Veszprem, Hungary*, Studies in Logic, Language and Information, CSLI Publications, Stanford, 1995A, pp. 293–309.

Venema, Y., Cylindric modal logic, *Journal of Symbolic Logic* **60** (1995B), 591–623.

Venema, Y. and Marx, M., *Multi-Dimensional Modal Logic*, Kluwer, Dordrecht, 1996.

Visser, A., de Jongh, D., Renardel, G. and van Benthem, J., NNIL, a study in intuitionistic propositional logic, in A. Ponse, M. de Rijke and Y. Venema (eds), *Modal Logic and Process Algebra*, CSLI Lecture Notes, Stanford, 1995, pp. 289–326.

HAJNAL ANDRÉKA and ISTVÁN NÉMETI
Institute of Mathematics,
Hungarian Academy of Sciences,
Budapest, Hungary
(email: andreka@sgiserv.rmki.kfki.hu,
nemeti@math-inst.hu)

JOHAN VAN BENTHEM
Institute for Logic, Language and Computation,
University of Amsterdam, The Netherlands
(email: johan@wins.uva.nl)