MINIMAL PREDICATES, FIXED-POINTS, AND DEFINABILITY

Johan van Benthem, Amsterdam & Stanford

Revised version, October 2004

Abstract

Minimal predicates P satisfying a given first-order description $\phi(P)$ occur widely in mathematical logic and computer science. We give an explicit first-order syntax for special first-order '*PIA* conditions' $\phi(P)$ which guarantees unique existence of such minimal predicates. Our main technical result is a preservation theorem showing *PIA*-conditions to be expressively complete for all those first-order formulas that are preserved under a natural model-theoretic operation of 'predicate intersection'. Next, we show how iterated predicate minimization on *PIA*-conditions yields a language *MIN(FO)* equal in expressive power to *LFP(FO)*, first-order logic closed under smallest fixed-points for monotone operations. As a concrete illustration of these notions, we show how our sort of predicate minimization extends the usual frame correspondence theory of modal logic, leading to a proper hierarchy of modal axioms: first-order-definable, first-order fixed-point definable, and beyond.

1 First-order logic with predicate minimization

One often defines a predicate uniquely in a model $\boldsymbol{M}$ as the smallest P satisfying a certain first-order description $\phi(P, \boldsymbol{Q})$, where $\boldsymbol{Q}$ is some tuple of given predicates. Our aim is to define a formalism allowing this device in a natural and useful fashion.

Before we get to general definitions, let us consider some motivating examples.

Example 1 A straightforward case of minimization.
The minimal predicate P satisfying the first-order formula

$$\phi(P, \boldsymbol{Q}) = \forall x\, (Qx \rightarrow Px)$$

exists in any model $\boldsymbol{M}$, and it is of course the predicate Q itself. ♣

In this case, the minimal predicate P is explicitly first-order definable in terms of the given predicates $\boldsymbol{Q}$. Such facts are widely used, e.g., in modal frame correspondence theory (van Benthem 1983, Blackburn, de Rijke & Venema 2001), whose high-lights include first-order definability for suitable monadic second-order sentences.

Example 2 Computing a first-order modal frame correspondence.

A basic correspondence connects the modal *K4*-axiom $[]p \rightarrow [][]p$ with transitivity $\forall x \forall y\ (Rxy \rightarrow \forall z\ (Ryz \rightarrow Rxz))$ of the accessibility relation *R*. The standard proof here takes a minimal predicate *P* satisfying the antecedent $[]p$ – i.e., the first-order formula $\forall y\ (Rxy \rightarrow Py)$ – at any given point *x*. The minimal *P* satisfying the formula

$$\phi(P, \mathbf{Q}) = \forall u\ (Rxu \rightarrow Pu)$$

is the first-order predicate $Pv := Rxv$. In the correspondence proof, the latter predicate is then substituted for all occurrences of *P* in the consequent $[][]p$ – i.e., $\forall y\ (Rxy \rightarrow \forall z\ (Ryz \rightarrow Pz))$ – to get transitivity at *x*: $\forall y\ (Rxy \rightarrow \forall z\ (Ryz \rightarrow Rxz))$. ♣

Example 2 is not to be confused with the fixed-point formula $\mu p \bullet []p$ in the modal *μ–calculus* (Stirling 1999), whose meaning is much more complex, witness Section 3 below. We will analyze the modal frame correspondence procedure in more detail in Section 4. But not all natural results of predicate minimization are first-order. A more general use is found in logic programs, where new predicates are introduced through recursive rules referring to 'minimal Herbrand models' (Doets 1994).

Example 3 Computing recursive Horn-clause definitions.

Consider a recursive description like

$$\phi(P, \mathbf{R}) = Ps \wedge \forall x \forall y\ ((Px \wedge Rxy) \rightarrow Py)$$

The minimal predicate here is a transitive closure, describing all points reachable from *s* in some finite number (*0* or more) of *R*-steps. This property is not first-order, but it can be defined in the well-known formalism *LFP(FO)* of first-order logic extended with fixed-point operators (cf. Ebbinghaus & Flum 1995). ♣

A minimal Herbrand model for a logic program is the term model for the language where all predicates defined by program clauses have their minimal extensions. The purely universal Horn-clause syntax guarantees the existence of such models for sets of clauses – but this is not necessary for the existence of minimal predicates. More general minimal predicates occur with 'predicate circumscription' in AI (McCarthy 1980). Predicate-circumscriptive consequence, as opposed to standard logical consequence, only requires truth of the conclusion in all predicate-minimal models of premises. This is a widely used formalism in so-called 'default reasoning'.

Example 4 Predicate circumscription.
Let $\alpha(R)$ say that R is a discrete linear order with an initial point but with no final or limit points. Consider the following description of a new predicate P, where the crucial second conjunct is not universal Horn, as its antecedent is not atomic:

$$\phi(P, \mathbf{R}) \ = \ \alpha(R) \wedge \forall x\, (\forall y\, (Ryx \to Py) \to Px).$$

$\phi(P, \mathbf{R})$ has P-minimal models over any domain, where the denotation of P is the initial segment of the R-order consisting of the standard natural numbers. This expressive power beyond first-order logic explains the high complexity of circumscription, which can define standard models categorically. ♣

This example makes P the so-called 'well-founded part' of the given binary order, which is also computable by a standard inductive definition (Aczel 1977). We will return to this particular connection in Sections 3, 4 below.

These four examples suggest a general semantic scheme for new predicates:

$MIN\ P \bullet \phi(P, \mathbf{Q})$ *the minimal predicate P such that* $\phi(P, \mathbf{Q})$,
with $\phi(P, \mathbf{Q})$ a first-order formula in a language with predicates P, $\mathbf{Q}$.

One way of stating its meaning more precisely follows predicate circumscription. The notation is well-defined in all models for the following second-order condition:

$$\exists P \bullet (\phi(P, \mathbf{Q}) \wedge \forall P'\, ((\phi(P', \mathbf{Q}) \to \forall x\, (Px \to P'x)).$$

But sometimes, such minimal conditions of use are not the most informative notion. Inspecting the above examples, we actually see a more concrete model-theoretic criterion that explains the unique existence of the minimal predicates. The following formulation of this criterion involves some harmless abuse of notation.

Definition 1 Intersection Property.
A first-order formula $\phi(P, \mathbf{Q})$ has the *intersection property for P* ('*IP*' for short) if, in any model $\mathbf{M}$, whenever $\mathbf{M}, \mathbf{P}_i \models \phi(P, \mathbf{Q})$ for all predicates in a family $\{\mathbf{P}_i \mid i \in I\}$, ϕ also holds for their intersection: that is, $\mathbf{M}, \cap\mathbf{P}_i \models \phi(P, \mathbf{Q})$. ♣

Applied to the extreme case of an empty family $\{\mathbf{P}_i \mid i \in I\}$, this says that $\phi(P, \mathbf{Q})$ holds for the intersection of the empty set, being the whole domain of the model. That is, $\phi(T, \mathbf{Q})$ is universally valid for formulas with *IP*.

All examples so far had defining clauses satisfying *IP* – as is easy to check by a direct set-theoretic argument. Also, their intended minimal predicates *P* are evidently the intersections of all predicates satisfying $\phi(P, \mathbf{Q})$ in the given model. Thus, *IP* justifies the phrasing 'minimal predicate satisfying the given description'. It is even a little bit stronger, as it also quantifies over smaller families of predicates satisfying $\phi(P, \mathbf{Q})$. We will retain this slight over-kill henceforth for technical convenience.

But our examples also suggest a concrete syntactic format behind this behaviour. The following definition introduces a sort of generalized Horn clauses, allowing non-atomic antecedents in the format '*P*-positive antecedent implies *P*-atom': The clause $\forall x\ (\forall y\ (Rxy \to Py) \to Px)$ in Example 4 is a typical illustration:

Definition 2 A first-order formula with identity is a *PIA condition* if it has the syntactic form $\forall \mathbf{x}\ (\psi(P, \mathbf{Q}, \mathbf{x}) \to P\mathbf{x})$, with *P* occurring only positively in the antecedent formula $\psi(P, \mathbf{Q}, \mathbf{x})$. Here $\mathbf{Q}$ is again a tuple of predicate letters in the base vocabulary, and $\mathbf{x}$ a tuple of individual variables. ♣

Conjunctions of *PIA* conditions can be rewritten to single ones by taking disjunctive antecedents. Here is the major semantic property of these special formulas.

Proposition 1 All *PIA* conditions $\phi(P, \mathbf{Q})$ have the Intersection Property.

Proof Suppose – with some harmless abuse of notation – that $\phi(\mathbf{P}_i, \mathbf{Q})$ holds in some model for all $i \in I$. Now let the antecedent $\psi(P, \mathbf{Q}, \mathbf{x})$ of $\phi(P, \mathbf{Q})$ hold for some tuple of objects $\mathbf{d}$ with *P* as the intersection $\cap \mathbf{P}_i$. By the positive occurrence of *P*, that antecedent $\psi(P, \mathbf{Q}, \mathbf{x})$ then also holds for each separate $\mathbf{P}_i$. But then $\mathbf{P}_i\ \mathbf{d}$ holds because of the truth of $\phi(\mathbf{P}_i, \mathbf{Q})$ – and hence $\cap \mathbf{P}_i$ holds for $\mathbf{d}$. ♣

It follows that the single-step format $MIN\ P \bullet \phi(P, \mathbf{Q})$ of our four examples so far, with $\phi(P, \mathbf{Q})$ a first-order *PIA* condition, defines unique predicate minimizations. There is also an obvious dual *MAX* of *MIN* for *maximal* predicates satisfying a given first-order description, but we will stick with minimization here. In Section 3, we will generalize this minimization format to an extension *MIN(FO)* of first-order logic closed under nested applications of predicate minimization. But for now, we continue with the model-theoretic analysis of first-order *PIA* conditions.

2 A preservation theorem for intersectivity

The main technical result of this paper is a model-theoretic preservation theorem stating the extent to which the syntactic *PIA*-format is expressively complete. But before proving this result, we state a simpler proposition, whose proof is a warm-up version for the more complex argument to follow. First, we restrict the *PIA*-format.

Definition 3 A *universal Horn formula w.r.t. P* is a first-order implication of the form $\forall \mathbf{x}\ (\phi(P, \mathbf{Q}, \mathbf{x}) \rightarrow P\mathbf{x})$ whose antecedent is constructed from arbitrary $\mathbf{Q}$-atoms and their negations, positive *P*-atoms, conjunction and disjunction only. ♣

This restricted first-order format suffices for many computational purposes, such as logic programming, or specifying abstract data-types. The following preservation theorem and its proof come from van Benthem 1985. But the result is already implicit in the discussion of reduced products and submodels in Chang & Keisler 1973, which refers to general results by Weinstein 1965 and Malcev 1971. Moreover, a related semantic take on universal Horn clauses in computer science is found in Mahr & Makowsky 1983. For convenience, we consider *unary* predicates *P* only in the arguments to follow, merely to save on tuple notation.

Theorem 1 The following are equivalent for all first-order formulas $\phi(P, \mathbf{Q})$:

(a) $\phi(P, \mathbf{Q})$ is definable by a universal Horn formula w.r.t. *P*

(b) $\phi(P, \mathbf{Q})$ has the Intersection Property w.r.t. predicate *P*, and it is also preserved under the formation of submodels.

Proof The implication from (a) to (b) is straightforward. *IP* follows from the *PIA*-form of $\phi(P, \mathbf{Q})$ – or alternatively, this property can easily be shown directly. And preservation under submodels follows by ϕ's universal syntactic form.

Conversely, assume that condition (b) holds. We must find a universal Horn definition for $\phi(P, \mathbf{Q})$. Consider the set of all universal Horn consequences of ϕ:

$$UH\text{-}Cons(\phi) = \{\psi \text{ universal Horn w.r.t. } P \mid \phi \models \psi\}$$

Here is the main observation needed to derive condition (a) above:

Lemma 1 (Consequence Lemma) $UH\text{-}Cons(\phi) \models \phi$

Once this is proved, by the Compactness Theorem, ϕ is implied by some finite conjunction of its own universal Horn consequences w.r.t. *P*, and hence it is equivalent to this conjunction. Clause (a) of the Theorem then follows because any such conjunction is equivalent to a single universal Horn condition.

Proof of the Consequence Lemma We need two major steps: one of model construction, and one of truth transfer using the given preservation properties of ϕ. For a start, let $\boldsymbol{M}$ be any model for $UH\text{-}Cons(\phi)$. First, we dispose of a special case. Let the predicate P hold for every object in $\boldsymbol{M}$. Then ϕ holds automatically in $\boldsymbol{M}$, by the earlier-noted fact that $\phi(T, \boldsymbol{Q})$ is universally valid if $\phi(P, \boldsymbol{Q})$ has IP. Next, let some objects d in $\boldsymbol{M}$ lack P. For all such d, we create the following situation:

Lemma 2 (Set-Up Lemma) There exists a model $\boldsymbol{N}_d$ for ϕ, together with a map f_d from $\boldsymbol{M}$ to $\boldsymbol{N}_d$ which is a $\boldsymbol{Q}$-isomorphic embedding and a P-homomorphism.

Proof Extend the given first-order language $L(P, \boldsymbol{Q})$ with new constant names $\underline{e}$ for each object $e \in \boldsymbol{M}$. Then for each object $d \in \boldsymbol{M}$ which lacks the property P:

(#) The following set Σ of formulas is finitely satisfiable:
$\Sigma = \{\phi\} \cup \{\neg P\underline{d}\} \cup$ the $P^+, \boldsymbol{Q}$-atomic diagram of $(\boldsymbol{M}, M)$,
where the latter set consists of all $\boldsymbol{Q}$-atoms and their negations
that are true in $\boldsymbol{M}$, plus all positive P-atoms that are true in $\boldsymbol{M}$.

Proof of (#) Suppose otherwise. Then there is some finite conjunction $\alpha(\underline{\boldsymbol{d}}, \underline{d})$ of formulas from the P^+, $\boldsymbol{Q}$-atomic diagram of $(\boldsymbol{M}, M)$, with the tuple of names $\underline{\boldsymbol{d}}, \underline{d}$ referring to objects $\boldsymbol{d}, d$ in $\boldsymbol{M}$, such that $\phi \wedge \alpha(\underline{\boldsymbol{d}}, \underline{d})$ implies $P\underline{d}$. Since the individual names $\underline{\boldsymbol{d}}, \underline{d}$ do not occur in the formula ϕ, this means that ϕ implies the universal Horn condition $\forall \boldsymbol{x}\ \forall x\ (\alpha(\boldsymbol{x}, x) \rightarrow Px)$. But the latter's evident falsity in $\boldsymbol{M}$ contradicts the assumption that $\boldsymbol{M} \models UH\text{-}Cons(\phi)$. ♣

Now, applying the Compactness Theorem to (#), the whole set Σ is satisfiable. So, there is a model $\boldsymbol{N}_d$ for all of $\{\phi\} \cup \{\neg P\underline{d}\} \cup$ the P^+, $\boldsymbol{Q}$-atomic diagram of $(\boldsymbol{M}, M)$.

Now, consider the map f_d from $\boldsymbol{M}$ into $\boldsymbol{N}_d$ sending the object e to the interpretation of its name $\underline{e}^{\boldsymbol{N}d}$. This is a $\boldsymbol{Q}$-isomorphic embedding as well as a P-homomorphism. E.g., it is a $\boldsymbol{Q}$-isomorphism as $\boldsymbol{N}$ verifies all $\boldsymbol{M}$-true $\boldsymbol{Q}$-literals, including negated identity atoms. Thus, $\boldsymbol{M}$ is $\boldsymbol{Q}$-isomorphic to the submodel $\boldsymbol{N}_d(\boldsymbol{M})$ of $\boldsymbol{N}_d$ whose domain consists of the interpretations in $\boldsymbol{N}_d$ of all names $\underline{e}$. ♣

It remains to use the preservation assumptions in clause (b) of the Theorem to get the desired conclusion for the Consequence Lemma.

Lemma 3 (Transfer Lemma) $\boldsymbol{M} \models \phi$.

Proof First apply the given preservation of ϕ under submodels to the above fact that $\boldsymbol{N}_d \models \phi$. In the submodel $\boldsymbol{N}_d(\boldsymbol{M})$ corresponding to $\boldsymbol{M}$, it follows that $\boldsymbol{N}_d(\boldsymbol{M}) \models \phi$.

Then we can use f_d to copy the interpretation of P in $\boldsymbol{N}_d(\boldsymbol{M})$ back into $\boldsymbol{M}$ to obtain a model $\boldsymbol{M}_d$ which coincides with $\boldsymbol{M}$ on $\boldsymbol{Q}$-predicates, while verifying

(a) ϕ, (b) all true P-atoms from $\boldsymbol{M}$, and (c) $\neg Pd$.

Now is the time to apply the Intersection Property of ϕ to the family of all models $\boldsymbol{M}_d$. The result is that ϕ must also hold on the model $(\boldsymbol{M}, \boldsymbol{Q}, P^*)$ with P^* the intersection of all predicates P_d in the separate models $\boldsymbol{M}_d$. But the latter is just the original predicate P on $\boldsymbol{M}$ itself! This final zeroing in on $P^{\boldsymbol{M}}$ via an intersection of ϕ-models is the main point of the whole elaborate construction of the family of models $\boldsymbol{N}_d$ in this proof. In other words, $(\boldsymbol{M}, \boldsymbol{Q}, P^*) = \boldsymbol{M}$ – and hence $\boldsymbol{M} |= \phi$. ♣

Now we come to the main result of this section. This appears to be new – but again, there is some history. Chang & Keisler 1993, Chapter 6, mentions a syntactic format like *PIA*, but with wholly positive antecedents in all predicate letters. Also, Papalaskari & Weinstein 1990 characterize the intersection property of Section 2 syntactically in the setting of propositional logic.

Theorem 2 The following are equivalent for all first-order formulas $\phi(P, \boldsymbol{Q})$:

(a) $\phi(P, \boldsymbol{Q})$ has the Intersection Property w.r.t. predicate P

(b) $\phi(P, \boldsymbol{Q})$ is definable by means of a *PIA* formula w.r.t. P.

Proof The argument has the same three major steps as the proof of Theorem 1, but there are some complications due to the absence of the shortcut via submodels.

From (b) to (a), the result is just Proposition 1. Next, assume condition (a). Again, we consider just a unary predicate P to avoid cumbersome tuple notation for objects. For a start, define the following set of syntactic consequences of ϕ:

$$PIA\text{-}Cons(\phi) = \{\psi \; PIA \text{ w.r.t. } P \mid \phi |= \psi\}$$

Lemma 4 (Consequence Lemma) $PIA\text{-}Cons(\phi) |= \phi$

If we can show this, then we are done, since the syntactic definability condition (b) will follow by the Compactness Theorem, plus the earlier observation that conjunctions of *PIA*-formulas are equivalent to single ones.

Proof of the Consequence Lemma Let $\boldsymbol{M}$ be any model for the language $L(P, \boldsymbol{Q})$ satisfying $PIA\text{-}Cons(\phi)$. As before, if $\boldsymbol{M} |= \forall x\, Px$, then $\phi(P, \boldsymbol{Q})$ already holds in $\boldsymbol{M}$, by *IP* for ϕ. For the remainder of this proof, we will assume that $\boldsymbol{M} |= \neg \forall x\, Px$. To fix notation, let $L(\boldsymbol{Q})$ be our first-order language with base predicates $\boldsymbol{Q}$ only.

Using a series of auxiliary results on model and map extensions, we will now construct a final situation as described in the following statement:

Lemma 5 (Set-Up Lemma) There exists an elementary extension $\boldsymbol{M}^*$ of $\boldsymbol{M}$ plus, for each $d \in \boldsymbol{M}^*$ lacking P, a model $\boldsymbol{N}_d$ and a map f_d from $\boldsymbol{M}^*$ to $\boldsymbol{N}_d$ such that

(a) ϕ is true in $\boldsymbol{N}_d$

(b) $P(f_d(d))$ is false in $\boldsymbol{N}_d$

(c) f_d is an $L(\boldsymbol{Q})$-isomorphism and a P-homomorphism from $\boldsymbol{M}^*$ onto $\boldsymbol{N}_d$

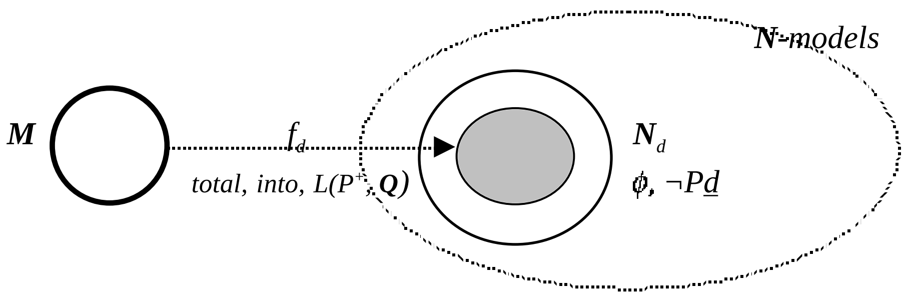
Proof of the Set-Up Lemma We state at the outset that all models in the following argument are *countable*, and so is the totality of all models used in the construction.

As before, $(\boldsymbol{M}, M)$ is the model $\boldsymbol{M}$ expanded to a model for the first-order language $L(P, \boldsymbol{Q})(\boldsymbol{M})$, which is the original $L(P, \boldsymbol{Q})$ enriched with new individual names $\underline{e}$ for each object e in $\boldsymbol{M}$ (whether e satisfies the predicate P or not). First, we find a set of models witnessing all P-failures in $\boldsymbol{M}$. This is much like the argument for claim (#) in the proof of Theorem 1. Fix any d in $\boldsymbol{M}$ with $\neg P^{\boldsymbol{M}}d$. We have that

(##) The following set of formulas is finitely satisfiable:
(i) ϕ, (ii) $\neg P\underline{d}$, plus (iii) $Th(P^+, \boldsymbol{Q})(\boldsymbol{M})$: the complete first-order theory of $(\boldsymbol{M}, M)$ in $L(P^+, \boldsymbol{Q})(\boldsymbol{M})$: i.e., $L(P, \boldsymbol{Q})(\boldsymbol{M})$ with only those formulas having all occurrences of P positive.

Proof of (##) If finite satisfiability fails, then ϕ implies some formula $\alpha(P, \boldsymbol{Q}, \underline{d}, \underline{\boldsymbol{e}}) \to P\underline{d}$ with P occurring only positively in α, and new object constants $\underline{\boldsymbol{e}}$ (one or more) and $\underline{d}$. But then the universal closure $\forall x\ \forall \boldsymbol{y}\ (\alpha(P, \boldsymbol{Q}, \boldsymbol{y}, x) \to Px)$ is a *PIA*-consequence of ϕ, which would therefore have to hold in $\boldsymbol{M}$: *quod non*. ♣

Now, by the Compactness Theorem, take any model $\boldsymbol{N}_d$ for the whole set of formulas in (#). It makes ϕ true, as well as $\neg P\underline{d}$. Moreover, the function f_d from $\boldsymbol{M}$ into $\boldsymbol{N}_d$ sending objects e in $\boldsymbol{M}$ to objects $\underline{e}^{\boldsymbol{N}d}$ preserves all $\boldsymbol{M}$-true first-order formulas of $L(P^+, \boldsymbol{Q})(\boldsymbol{M})$. (In particular, since f_d preserves all true non-identities, it is *1-1*.) We can do this for any object in $\boldsymbol{M}$ lacking the property P, and the result is a countable family of models $\boldsymbol{N}_d$ with maps f_d from $\boldsymbol{M}$ into them. In a picture:



This is the start for a procedure constructing elementary chains of models on the '***M***-side' and the '***N***-side'. There will always be one current model elementarily extending ***M***, while the family of models $\boldsymbol{N}_{\mathrm{d}}$ is both modified by elementary extensions of existing ones and addition of new ones. The inductive step of this construction actually needs a bit less than the above, as the maps f_d need not be total:

Inductive step $(k \to k+1)$ Let $\boldsymbol{M}^k$ be the current model elementarily extending ***M***, while there is a family of models $\boldsymbol{N}^k_d$ – one for each d in $\boldsymbol{M}^k$ lacking P – together with partial maps f^k_d that preserve all first-order formulas α of the language $L(P^+, \boldsymbol{Q})$ w.r.t. all object tuples $\boldsymbol{e}$ all of whose objects occur in the domain of f^k_d:

($) if $\boldsymbol{M}^k \models \alpha\, [\boldsymbol{e}]$, then $\boldsymbol{N}^k_d \models \alpha\, [f^k_d(\boldsymbol{e})]$

So, the current maps may be partial, and non-surjective. We now give a three-step procedure for extending these models and maps to larger domains, while restoring the properties that we start with here – in particular, the crucial invariant *($)*.

Step A We find an elementary extension for each model $\boldsymbol{N}^k_d$, and we simultaneously extend the given f^k_d from $\boldsymbol{M}^k$ into it to a new map that is total on $\boldsymbol{M}^k$, while still preserving all true $L(P^+, \boldsymbol{Q})$-formulas.

By way of preparation, we add new individual constants $\underline{e}$ denoting objects e in $\boldsymbol{M}^k$ that are in the domain of f_d, and interpret these in $\boldsymbol{N}^k_d$ via their f_d-images. Thus, both models get expanded. By *($)*, every $L(P^+, \boldsymbol{Q})$-sentence true in $\boldsymbol{M}^k$ is also true in $\boldsymbol{N}^k_d$. Next, we add new individual constants for all objects in $\boldsymbol{M}^k$, and expand the latter model once more. Our first-order language now contains all new constants. Then, finding the extended model and mapping uses the following fact:

(A#) The following set $\boldsymbol{\Sigma}$ is finitely satisfiable:
(a) all P-positive $P, \boldsymbol{Q}$–sentences α true in the twice-expanded $\boldsymbol{M}^k$,
plus (b) the complete first-order theory of the expanded model $\boldsymbol{N}^k_d$.

Proof In fact, $\boldsymbol{\Sigma}$ is finitely satisfiable in the expanded model $\boldsymbol{N}^k_d$. Consider any finite subset $\boldsymbol{\Sigma}_0$. It may have some (b)-type formulas that are true in the expanded $\boldsymbol{N}^k_d$ as they stand. As for the formulas of the (a)-type, we can take any finite conjunction of these in the extended language, and existentially quantify over the new constants naming objects different from those already in the domain of f_d. Using the fact that (a) P-positive formulas are closed under conjunctions and existential quantifiers, and

(b) the old map f_d preserved $L(P^+, \boldsymbol{Q})$-formulas, we see that this existential formula was already true in the expanded $\boldsymbol{N}^k{}_d$.This provides the required model for Σ_0. ♣

Next, by Compactness, we find a model for the whole formula set Σ, which yields the required elementary extension, as well as the extended map as in earlier proofs. In particular, the new map extends the old. Let the object e in the domain of f_d be named by the constant $\underline{e}$. Let $\underline{f}$ be the old constant naming e as an element of the domain of f_d. Then the atom $\underline{e}=\underline{f}$ is true in the expansion of $\boldsymbol{M}^k$, and hence it was preserved into the new model.

To summarize the result of Step *A*, we note that:

$\boldsymbol{M}^{k+1, A} = \boldsymbol{M}^k$
$\boldsymbol{N}^{k+1, A}{}_d$ *is an elementary extension of* $\boldsymbol{N}^k{}_d$
$f^{k+1, A}{}_d \supseteq f^k{}_d$ *is total on* $\boldsymbol{M}^{k+1, A}$*, but not necessarily surjective*

Step B We find an elementary extension $\boldsymbol{M}^{k+1, B}$ of $\boldsymbol{M}^{k+1, A}$, as well as an extension of each map $f^{k+1, A}{}_d$ to a surjection onto the model $\boldsymbol{N}^{k+1, A}{}_d$ that still preserves all $L(P^+, \boldsymbol{Q})$-formulas true with parameters in $\boldsymbol{M}^{k+1, B}$.

First, given $\boldsymbol{M}^{k+1, A}$ and any model $\boldsymbol{N}^{k+1, A}{}_d$, we can extend $\boldsymbol{M}^{k+1, A}$ and the existing map $f^{k+1, A}{}_d$ so that the extended map still preserves all *P*-positive *P*, $\boldsymbol{Q}$-formulas, while having all of $\boldsymbol{N}^{k+1, A}{}_d$ inside its image. The argument is similar to that in Step *A*. We add new individual constants as before for all objects in $\boldsymbol{N}^{k+1, A}{}_d$, and show

(B#) The following set Σ in the extended language is finitely satisfiable:
(a) the complete first-order theory of $\boldsymbol{M}^{k+1, A}$, plus
(b) the set of all negations $\neg\alpha$ of *P*-positive *P*, $\boldsymbol{Q}$–formulas α that are true in the expanded model $(\boldsymbol{N}^{k+1, A}{}_d, N^{k+1, A}{}_d)$.

Proof The set is finitely satisfiable in $\boldsymbol{M}^{k+1, A}$. If not, then $\boldsymbol{M}^{k+1, A}$ would satisfy some formula $\forall \boldsymbol{x}(\alpha_1 \vee ... \vee \alpha_k)$ with all α_i *P*-positive – with the universal quantifier $\forall \boldsymbol{x}$ running over all new object names used from $(\boldsymbol{N}^{k+1, A}{}_d, N^{k+1, A}{}_d)$. By closure under disjunctions and universal quantifiers, this formula is still in our class of *P*–positive $L(P^+, \boldsymbol{Q})$-formulas – and so, by *($)*, it would have been true in $\boldsymbol{N}_d$: *quod non*. ♣

It follows that the whole set of formulas Σ is satisfiable, and any model for it will be the required *P*, $\boldsymbol{Q}$–elementary extension of $\boldsymbol{M}^{k+1, A}$, while also yielding the right extension $f^{k+1, B}{}_d$ for the map $f^{k+1, A}{}_d$ in an obvious way. But this is not enough! We must achieve this for *all* models $\boldsymbol{N}^{k+1, A}{}_d$ that existed at the end of Step *A*.

For that purpose, we now arrange all these models in some countable enumeration, and repeat the preceding construction through all finite ordinals. In particular, in each of these steps, the current descendant of the initial model $\boldsymbol{M}^{k+1,\ A}$ changes to some $L(P, \boldsymbol{Q})$-elementary extension. But this does not affect the crucial preservation property *($)* for our partial maps, as truth values for all relevant formulas do not change between an elementary extension and the original model $\boldsymbol{M}^{k+1,\ A}$. Finally, taking the union of the resulting elementary chain of $\boldsymbol{M}$-models is the required model $\boldsymbol{M}^{k+1, B}$, while the maps $f^{\,k+1,\ B}{}_d$ constructed during the stages are the required surjections still satisfying *($)*. Note that these maps are also injective, as the invariance condition *($)* implies preservation of negated identity atoms.

To summarize the result of Step *B*, we write:

$\boldsymbol{M}^{k+1, B}$ *is an elementary extension of* $\boldsymbol{M}^{k+1, A}$
$\boldsymbol{N}^{\,k+1,\,B}{}_d = \boldsymbol{N}^{\,k+1,\,A}{}_d$
$f^{\,k+1, B}{}_d \supseteq f^{\,k+1, A}{}_d$ *is surjective on* $\boldsymbol{N}^{\,k+1, B}{}_d$*, but not necessarily total*

Step C In taking the union $\boldsymbol{M}^{k+1, B}$ of an elementary chain in Step *B*, the domain of this model may have acquired many new objects *d* that lack *P*, though they are not in the domain of any map onto a matching model $\boldsymbol{N}_d$. Finally,

> Create a family of such models, plus embedding maps satisfying *($)*, exactly as in the argument setting up our first stage. This does not change $\boldsymbol{M}^{k+1, B}$, or any of the other models and maps existing by the end of Step *B*.

The result of Steps *A, B, C* executed successively is

(i) a model $\boldsymbol{M}^{k+1}$ which is an elementary extension of $\boldsymbol{M}^{k}$,
(ii) a family of models $\boldsymbol{N}^{k+1}{}_d$ elementarily extending the models $\boldsymbol{N}^{k}{}_d$ existing at the end of Stage *k*,
(iii) a family of partial maps $f^{k+1}{}_d$ from $\boldsymbol{M}^{k+1}$ onto $\boldsymbol{N}^{k+1}{}_d$ satisfying the preservation condition *($)* for $L(P^{+}, \boldsymbol{Q})$-formulas, whose domain includes $\boldsymbol{M}^{k}$ and whose range includes $\boldsymbol{N}^{k}{}_d$ – and
(iv) new models $\boldsymbol{N}_d$ witnessing all objects in $\boldsymbol{M}^{k+1}$ that lack the property *P*, with maps as in (iii) – not necessarily surjective.

In particular, the initial situation has been restored.

Iteration to an Elementary Chain To conclude the proof of the Set-Up Lemma, we iterate the inductive step described here through all finite ordinals. The result is an elementary chain of models $\boldsymbol{M}^1$, $\boldsymbol{M}^2$, ... whose union is the required model $\boldsymbol{M}^*$. Moreover, the iterative procedure guarantees that each of its elements lacking *P* has started an elementary chain of models $\boldsymbol{N}^k{}_d$, $\boldsymbol{N}^{k+1}{}_d$, ... from some stage *k* onward, whose union is the required model $\boldsymbol{N}_d$ of the Set-Up Lemma. Finally, the union of all partial maps $f^n{}_d$ between $\boldsymbol{M}^n$ and $\boldsymbol{N}^n{}_d$ constructed at stages *n* of this process is the required $L(\boldsymbol{Q})$-isomorphism and *P*-homomorphism f_d from $\boldsymbol{M}^*$ to $\boldsymbol{N}_d$. In particular, the map has become a bijection because of the back-and-forth domain extension steps in Step *A* and Step *B*, while the preservation condition *($)* for $L(P^+, \boldsymbol{Q})$-formulas with finitely many parameters $\boldsymbol{d}$ still holds because $\boldsymbol{M}^*$ is elementarily equivalent to the model $\boldsymbol{M}^k$ where all objects $\boldsymbol{d}$ first appeared together. ♣

Now we are ready to clinch our argument.

Lemma 6 (Transfer Lemma) $\boldsymbol{M} \models \phi$.

Proof Consider the situation in the Set-Up Lemma. Each model $\boldsymbol{N_d}$ satisfies ϕ, and moreover, it is *P*, $\boldsymbol{Q}$-isomorphic to the model $(\boldsymbol{M}^*, P_d)$ which is like $\boldsymbol{M}^*$, but with the interpretation of the predicate letter *P* replaced by one copied from that of $\boldsymbol{N_d}$ via the map f_d. This makes f_d into a complete *P*, $\boldsymbol{Q}$-isomorphism, and hence

$$(\boldsymbol{M}^*, P_d) \models \phi$$

Also, the *P*-homomorphism condition ensures that

the copied predicate P_d contains $P^{\boldsymbol{M}^*}$

Finally, note that

the object *d* in $\boldsymbol{M}^*$ does not satisfy P_d

Now we use the given Intersection Property of ϕ. The model $\boldsymbol{M}^\#$ with the intersection of all predicates P_d interpreting the predicate letter *P* must also satisfy ϕ. But by the preceding observations, that intersection is just $P^{\boldsymbol{M}^*}$, and so $\boldsymbol{M}^\#$ is in fact just the model $\boldsymbol{M}^*$. It follows that $\boldsymbol{M}^* \models \phi$. But then also $\boldsymbol{M} \models \phi$, since $\boldsymbol{M}^*$ is an elementary extension of the original model $\boldsymbol{M}$ for *PIA-Cons(*ϕ*).* ♣

There are several variations on Theorem 2; some much simpler to prove. The universal Horn clauses of Theorem 1 were one example. Another special case lets *P* occur in consequent position only. Van Benthem 1996 shows this is equivalent to strengthening the Intersection Property to an equivalence – or more perspicuously, to adding a separate semantic requirement to *IP* that $\phi(P, \boldsymbol{Q})$ be *monotone* w.r.t. *P*.

3 Predicate-minimizing and fixed-point logics

Predicate minimization can also be added as a general device to first-order logic. The result is the following formalism.

Definition 4 The language of *first-order logic with predicate minimization (MIN(FO))* has all the recursive formation rules of standard first-order logic plus a new formation rule for formulas

$MIN\ P\bullet\ \phi(P, \mathbf{Q})$ where $\phi(P, \mathbf{Q})$ is an *extended PIA-condition*.

The latter still have the syntactic shape of Definition 2, but *MIN(FO)* syntax allows any *P*-positive antecedents $\psi(P, \mathbf{Q}, \mathbf{x})$ *from MIN(FO)*. Here, positive occurrences of atoms in $\alpha(R, \mathbf{Q})$ not involving *R* are also positive in $MIN\ R\bullet\ \alpha(R, \mathbf{Q})$. ♣

MIN(FO) is closely related to the more standard language *LFP(FO)* extending first-order logic with a recursive formation rule for fixed-point operators.

Definition 5 *LFP(FO)* extends the usual inductive formation rules for first-order syntax with an operator defining smallest fixed-points

$\mu P, \mathbf{x}\bullet\ \phi(P, \mathbf{Q}, \mathbf{x})$

where *P* may occur only positively in $\phi(P, \mathbf{Q}, \mathbf{x})$, and $\mathbf{x}$ is a tuple of variables of the right arity for *P*. The relevant fixed-points are those of the following monotone set operation on predicates in any given model $\boldsymbol{M}$:

$F^{\boldsymbol{M}}_{\phi} = \lambda P\bullet \{\mathbf{d}\ in\ M \mid (\boldsymbol{M}, P), \mathbf{d} \models \phi(P, \mathbf{Q})\}$

By the Tarski-Knaster Theorem, the denotation of $\mu P, \mathbf{x}\bullet\ \phi(P, \mathbf{Q}, \mathbf{x})$ may be defined correctly as the intersection of all predicates *P* on $\boldsymbol{M}$ with $F^{\boldsymbol{M}}_{\phi}(P) \subseteq P$ – which is also the smallest subset *X* of $\boldsymbol{M}$ such that $F^{\boldsymbol{M}}_{\phi}(X) = X$. ♣

In this definition, the syntactic condition of positive occurrence for *P* in ϕ guarantees the monotonicity of the map $F^{\boldsymbol{M}}_{\phi}$. This condition is backed up by a well-known model-theoretic result. A simple variant of Lyndon's preservation theorem for homomorphisms states that a first-order formula $\phi(P, \mathbf{Q})$ defines a monotone set operation F^M_ϕ iff $\phi(P, \mathbf{Q})$ is definable by a formula with only positive occurrences of the predicate *P*. We can look at our Theorem 2 as doing the same for *MIN(FO).*

A modal variant of *LFP(FO)* is the modal *μ–calculus*, where all predicates are unary, and the $\phi(P, \mathbf{Q})$ are modal formulas. These modal and first-order languages can also define *greatest fixed-points* by dualization, just as we can look at predicate *maximization* instead of minimization – but we do not pursue this angle here.

Despite the phrasing of 'minimality', the smallest fixed-point for $\phi(P, \mathbf{Q})$ is usually not a predicate *P* for which $\phi(P, \mathbf{Q})$ holds. For instance, in Example 2 of Section 1, the minimal predicate *P* satisfying a modal formula *[]p* (i.e., $\forall y\ (Rxy \rightarrow Py)$) at some point *x* was just $\{s|Rxs\}$. But the smallest fixed-point for $\forall y\ (Rxy \rightarrow Py)$) – written μp• *[]p* in the μ–calculus – is much more complicated: it defines the well-founded part of the given relation *R*, which occurred in Example 4 of Section 1. Nevertheless, there is an intimate connection between the two formalisms.

Proposition 2 *MIN(FO)* and *LFP(FO)* have equal expressive power.

Proof (a) From *LFP(FO)* to *MIN(FO)*. The smallest fixed-point for the operation *F* as described above is also a smallest 'pre-fixed point', which can be represented as follows, writing $\mathbf{x}$ for the tuple of the relevant free variables:

$$\mu P, \mathbf{x}\bullet\ \phi(P, \mathbf{Q}, \mathbf{x}) = MIN\ P\bullet\ \forall \mathbf{x}\ (\phi(P, \mathbf{Q}, \mathbf{x}) \rightarrow P\mathbf{x})$$

Here we can assume inductively that the *LFP(FO)*-antecedent $\phi(P, \mathbf{Q}, \mathbf{x})$ already has a *MIN(FO)*-equivalent. (b) From *MIN(FO)* to *LFP(FO)*. Minimization just occurs over *PIA*-conditions $\forall \mathbf{x}\ (\phi(P, \mathbf{Q}, \mathbf{x}) \rightarrow P\mathbf{x})$, with *P* occurring only positively in $\phi(P, \mathbf{Q}, \mathbf{x})$. But the same predicate can be described as $\mu P, \mathbf{x}\bullet\ \phi(P, \mathbf{Q}, \mathbf{x})$. ♣

A choice between the languages *LFP(FO)* and *MIN(FO)* seems largely a matter of practical convenience. More theoretically, our preservation results in Section 2 are the counterpart of the above 'Lyndon justification' for imposing the constraint of positive occurrence in *LFP(FO).* We have tried to find some more direct reduction of our preservation results in Section 2 to a Lyndon-style one, but without success.

Remark: an open preservation problem In this connection, there is a natural model-theoretic question that seems open for both formalisms. E.g., is it still true in the full language *LFP(FO)* that the formulas $\phi(P)$ defining monotone set operations are just those definable by formulas having only positive occurrences of *P*? No Lyndontype theorem is known for *LFP(FO),* because the usual compactness-based model-theoretic techniques for first-order logic fail – including those of Section 2. And their substitutes for infinitary languages like $L_{\infty\omega}$ (cf. Barwise & van Benthem 1999) fail, too – as *LFP(FO)* defines well-foundedness of binary orders, which is beyond these. Exactly the same question is open for our semantic Intersection Property and the extended *PIA*-format used in the definition of *MIN(FO)* . ♣

Another aspect of the comparison concerns *fine-structure*. Smallest fixed-point denotations for *LFP(FO)* can be computed in ordinal stages, following a well-known bottom-up approximation procedure starting from the empty predicate. In particular, some fixed-points are uniformly computable in any model by stage ω: e.g., predicate denotations in minimal Herbrand models for logic programs. Other fixed-points require growing stages up to the cardinality of the model, with the well-founded part of a given binary ordering as a key example. One can predict some of this behaviour from the shape of the formulas $\mu P, \mathbf{x}\bullet\ \phi(P, \mathbf{Q}, \mathbf{x})$. E.g., van Benthem 1996 analyzes stabilization by stage ω in terms of *finite continuity* in the predicate *P*: '$\phi(P, \mathbf{Q})$ holds iff $\phi(P_0, \mathbf{Q})$ holds for some finite subpredicate P_0 of *P*'. Its syntactic counterpart turns out to be some form of *positive-existential* occurrence of *P* in ϕ, without the universal quantification that makes the well-founded part case so complex. It may be of interest to find similar fine-structure inside *MIN(FO).*

4 Minimization and fixed-points in modal correspondence theory

In this final section, we explore some new uses of predicate minimization and *PIA* syntax. Minimal predicates are used extensively in modal logic, when computing so-called 'frame correspondents' for modal formulas. Here are the basic notions – for more details concerning modal logic we refer to the standard literature (e.g., van Benthem 1983, Blackburn, de Rijke & Venema 2001). A modal formula $\phi(p_1, ..., p_k)$ is called *true in a frame* $\mathbf{F} = (W, R)$ if, for each valuation for its proposition letters $p_1,..., p_k$, ϕ holds in every world of that frame. This notion treats modal formulas ϕ as monadic second-order closures of their standard first-order translation $ST(\phi)$ on relational models, viz. as monadic Π^1_1–formulas

$$\forall P_1 \dots P_k\ ST(\phi)$$

But in many cases, better equivalent properties exist, indeed first-order ones, which can be computed from the form of the modal axioms. The case of transitivity and the modal axiom $[]p \rightarrow [][]p$ in Section 1 was a key example. How far does this go?

4.1 Computing first-order frame correspondents

First-order frame correspondences are often proven ad-hoc. In Section 1, we already mentioned that the *K4*-axiom *[]p → [][]p* corresponds to first-order *transitivity*: $\forall xy\, (Rxy \rightarrow \forall z\, (Ryz \rightarrow Rxz))$. Another well-known case is the following

Example 5 The *.2-axiom <a>[b]p → [b]<a>p* corresponds to *confluence*:

$\forall xy\, (R_a xy \rightarrow \forall z\, (R_b xz \rightarrow \exists u\, (R_a zu \wedge R_b yu))$ ♣

Such results can be computed more uniformly using a well-known *substitution algorithm* (van Benthem 1983, Blackburn, de Rijke & Venema 2001). It turns modal axioms of suitable syntactic shapes ('Sahlqvist forms') into equivalent first-order conditions on accessibility relations on frames.

Theorem 3 Modal formulas $\alpha \rightarrow \beta$ of the following form have first-order frame correspondents. Antecedents α must be constructed from atoms *p, q, ...* possibly prefixed by universal modalities, conjunctions, disjunction, and existential modalities, while consequents β can be any modal formula positive in all its proposition letters. Also, the first-order correspondents can be computed uniformly and effectively from the given modal axioms.

Proof sketch The proof of this result is widely available in the modal literature. Here is the effective procedure. The substitution algorithm computing the frame equivalents works as follows for modal axioms $\alpha \rightarrow \beta$ of the given syntactic form:

(a) Translate the modal axiom into its canonical first-order form, prefixed with monadic set quantifiers for proposition letters:
$\forall x\!: \forall \mathbf{P}\!: translation(\alpha \rightarrow \beta)(\mathbf{P}, x)$,
(b) Pull existential modalities in the antecedent forward to become bounded universal quantifiers in the prefix,
(c) Compute first-order *minimal values* for the proposition letters making the remaining portion of the antecedent true,
(d) Substitute these definable values for the proposition letters occurring in the body of the consequent – and if convenient,
(e) Perform some simplifications modulo logical equivalence.

Example 6 For the modal transitivity formula *[]p → [][]p,*

(a) yields $\forall x\!: \forall P\!: \forall y\, (Rxy \rightarrow Py) \rightarrow \forall z\, (Rxz \rightarrow \forall u\, (Rzu \rightarrow Pu))$,
(b) is vacuous – as there are no existential modalities in *[]p*, while
(c) yields the minimal value $Ps := Rxs$ – and then

(d) substitution gives $\forall x$: $\forall y\ (Rxy \rightarrow Rxy) \rightarrow \forall z\ (Rxz \rightarrow \forall u\ (Rzu \rightarrow Rxu))$.
(e) the latter simplifies to the usual form $\forall x$: $\forall z\ (Rxz \rightarrow \forall u\ (Rzu \rightarrow Rxu))$. ♣

Example 7 For the modal confluence formula $<a>[b]p \rightarrow [b]<a>p$,

(a) yields $\forall x$: $\forall P$: $\exists y\ (R_a xy \wedge \forall z\ (R_b yz \rightarrow Pz))$
$\rightarrow \forall u\ (R_b xu \rightarrow \exists v\ (R_a uv \wedge Pv))$,
(b) yields $\forall x$: $\forall P$: $\forall y\ (R_a xy \rightarrow$
$(\forall z\ (R_b yz \rightarrow Pz) \rightarrow \forall u\ (R_b xu \rightarrow \exists v\ (R_a uv \wedge Pv)))$,
(c) yields the minimal value $Ps := R_b ys$
(d) substitution gives $\forall x$: $\forall y\ (R_a xy \rightarrow$
$(\forall z\ (R_b yz \rightarrow R_b yz) \rightarrow \forall u\ (R_b xu \rightarrow \exists v\ (R_a uv \wedge R_b yv)))$,
(e) the latter simplifies to the usual form
$\forall x\ \forall y\ (R_a xy \rightarrow \forall u\ (R_b xu \rightarrow \exists v (R_a uv \wedge R_b yv))$. ♣

For the correctness of the substitution algorithm, we refer to the cited literature – since it is not our main concern here. The main idea is this. Clearly, the formulas of step (a) imply their special substitution instance in step (d). Vice versa, assume that the latter is true in a modal frame ***F***. If an antecedent α in Sahlqvist form is true in ***F*** at a point *x*, for any valuation *V(p)* for its proposition letters *p*, then it will also be true for the minimal values computed in step (c), which are contained in the sets *V(p)*. Therefore, the substitution instance in (d) says that the consequent β holds at *x* for those minimal values. But then it also holds for the original *V(p)*-values, by the semantic monotonicity induced by its positive syntactic form. This shows that the second-order formula in (a) expressing frame truth of $\alpha \rightarrow \beta$ is true at *x* in ***F***. ♣

In our perspective, what happens here is this. In dealing with Sahlqvist antecedents α, step (c) of the above algorithm uses predicate minimizations *MIN P*• $\phi(P, \mathbf{Q})$, where all conditions ϕ are *PIA*. This follows from the syntactic form of the translated α, after existential modalities have become universal prefix quantifiers, so that only iterations of the form *[]...[]p* remain. The corresponding *PIA* conditions are even very special, as the predicate *P* to be minimized does not occur in their antecedents, which only refer to relational successor chains. This is the special case mentioned at the end of Section 3, which explains the first-order definability. Nevertheless, minimization would also work with other types of modal antecedent – and this suggests a surprising extension of the above theorem in Section 4.2.

Remark Not all first-order frame-definable modal formulas have equivalents by first-order substitution (van Benthem 1983). A counter-example is the conjunction of the *K4* transitivity axiom and the *McKinsey Axiom* $[]<>p \rightarrow <>[]p$. ♣

4.2 Generalized frame correspondents in fixed-point logic

The preceding substitution algorithm typically runs into difficulties with modal implications $\alpha \rightarrow \beta$ whose antecedents are of more complex forms.

Example 8 *Löb's Axiom* $[]([]p \rightarrow p) \rightarrow []p$ defines the conjunction of the following two frame conditions: (a) transitivity of *R*, (b) upward well-foundedness of *R*. This property of binary relations is evidently not first-order definable. ♣

The failure of the earlier substitution algorithm here can be understood as follows. We do not get a minimal value for an antecedent $[]([]p \rightarrow p)$ which is first-order definable in terms of *R*, = only. But still, we may observe that

the Löb antecedent has the *PIA*-form $\forall y\ ((Rxy \wedge \forall z\ (Ryz \rightarrow Pz)) \rightarrow Py)$

Therefore, this antecedent supports a minimal value: not in the first-order language of *R* and =, but in *MIN(FO)*, or equivalently, the fixed-point language *LFP(FO)*.

Example 8, continued Computing the minimal valuation for Löb's Axiom.
Analyzing $[]([]p \rightarrow p)$ a bit more closely, the minimal predicate satisfying the antecedent of Löb's Axiom at a world *x* describes the following set of worlds:

$\{y \mid \forall z\ (Ryz \rightarrow Rxz)$ & no infinite sequence of *R*-successors starts from $y\}$.

Then, if we plug this description into the Löb consequent $[]p$, precisely the usual, earlier-mentioned conjunctive frame condition will result automatically. ♣

Here is the general upshot of these observations:

Proposition 3 Modal implicational axioms with positive consequents and antecedents that are *PIA* modulo extracting existential outer quantifiers have effectively computable frame correspondents in *LFP(FO)*.

We could also define these frame correspondents in *MIN(FO)*. Either way, many modal axioms beyond Sahlqvist forms have correspondents in fixed-point logics. By itself, this observation is not new. *LFP(FO)* has also been used explicitly in Nonnengart & Salas 1999, as part of their 'SCAN-algorithm' for analyzing second-order frame properties and turning them into more manageable logical forms.

Here is one more illustration of the fine-structure of useful fixed-point equivalents.

Example 9 'Cyclic Return'.
The modal axiom $(<>p \wedge [](p \rightarrow []p)) \rightarrow p$ expresses the frame property that "every point *x* with an *R*-successor *y* can be reached from *y* by a finite sequence of successive *R*-steps" (van Benthem 1983, Benton 2002). The antecedent of Cyclic Return becomes *PIA* after first pulling out a prefixed universal quantifier $\forall y\ (Rxy \rightarrow$ for the existential modality $<>p$. The resulting minimal predicate is 'being reachable from *y* in finitely many *R*-steps'. After substitution in the consequent, the eventual *LFP(FO)*-equivalent is exactly the mentioned frame condition. ♣

Cyclic Return involves only transitive closure of *R*, and hence a simple fixed-point suffices, reached at the first infinite approximation stage ω. The reason for this simplicity is a syntactic one, related to our observations in Sections 1, 2. The *PIA* antecedent computed from the antecedent $<>(p \wedge [](p \rightarrow []p))$ is a universal Horn clause. By contrast, in the minimal value computed for the non-Horn *PIA* antecedent $[]([]p \rightarrow p)$ of Löb's Axiom, the fixed-point may take any infinite ordinal stage before it is reached, as it computes the well-founded part of a binary relation.

4.3 A hierarchy up to non fixed-point definability

There are limits to minimization. Consider the earlier McKinsey Axiom

$$[]<>p \rightarrow <>[]p,$$

another well-known modal principle without a first-order equivalent. Our fixed-point analysis does not apply here, as the modal antecedent $[]<>p$ has a typically non-*PIA* first-order quantifier pattern $\forall x\ (Rxy \rightarrow \exists y\ (Ryz \wedge \ldots$ There are other higher-order correspondence algorithms which can deal with this case (Gabbay & Ohlbach 1992, Nonnengart & Salas 1999), but these do not deliver an *LFP(FO)*-condition.

Indeed, we have a hierarchy here. Perhaps the simplest example of a non-fixed-point definable frame condition comes from basic temporal logic with modalities *F* for future and *P* for past. Consider the well-known *Dedekind Axiom*

$$(Fp \wedge FG\neg p) \rightarrow F(G\neg p \wedge H(p \vee Fp))$$

On strict linear orders $(T, <)$, this expresses Dedekind Completeness: every subset of *T* with a lower bound has a greatest lower bound. Again, the antecedent is not *PIA*. And indeed, this monadic Π^1_1–property is not definable in the above style.

Proposition 4 Dedekind Completeness is not definable in *LFP(FO)*.

Proof Dedekind continuity holds in the reals *R*, and fails in the rationals *Q*. But there exists a well-known *potential isomorphism* between these frames. As potential isomorphisms preserve all formulas of *LFP(FO)*, non-definability follows. ♣

Thus, we find a new hierarchy among frame correspondents in temporal logic:

first-order, fixed-point definable, essentially higher-order.

A similar hierarchy exists in modal logic:

Theorem 4 There are modal formulas which are not definable in *LFP(FO)*.

Proof Le Bars 2002 presents a modal formula whose truth on finite frames does not satisfy the Zero-One Law for the probability of truth with increasing domain size. But all formulas definable in *LFP(FO)* do satisfy this Zero-One Law (Ebbinghaus & Flum 1995). More precisely, Le Bars looks at finite frames satisfying a simple first-order condition saying that a frame has relational width *2*, which is known to hold with probability *1* in the limit on finite models. Then he considers a further modal formula which may be written as follows:

$$(q \wedge []<>p) \rightarrow (p \vee <><>((p \vee q) \wedge <>(p \vee q)))$$

This formula does not obey a Zero-One Law on the finite frames of width *2*. Hence it is not even *LFP(FO)*-definable in that special case. Again, our method would fail, because the antecedent $q \wedge []<>p$ is typically non-*PIA*. ♣

We conjecture that already the McKinsey Axiom is not definable in *LFP(FO)*.

Remark Several points similar to those made in Sections 5.2, 5.3 are found in Goranko & Vakarelov 2003, which we learnt about after writing this paper. In particular, their 'regular formulas' can be shown to be equivalent to Sahlqvist implications with *PIA*-antecedents, restricted to a modal language with only unary base predicates *P*. The authors point out that such formulas have frame conditions definable in *LFP(FO)*. Moreover, they announce further work on correspondence and completeness in modal fixed-point formalisms extending the above language.

4.4 Some possible extensions

Multiple minimization. The substitution algorithm of Section 4.1 works smoothly on formulas α with several proposition letters p_1, p_2, …. Here, antecedents in the *PIA*-based definitions of minimal values do not involve any predicates P_1, P_2, … But more generally, one can simultaneously minimize a bunch of predicates with respect to a conjunction ϕ of *PIA*-conditions, in the syntactic format *MIN* $P_1, P_2, \ldots \bullet \phi(P_1, P_2, \ldots, \mathbf{Q})$. The proviso is that, in the antecedent of the implicational condition for *P*, all other predicates being minimized occur only positively. We omit details.

Richer modal languages. Minimal substitution analysis applies far beyond the basic modal language of *[], <>*. Suitable antecedents include temporal *Until* modalities *U(p, q)* of quantifier form $\exists\forall$ – and consequents are acceptable in any language (first-order, higher-order) as long as they are monotonic in the proposition letters. In particular, to get expressive harmony between the language of the modal axioms and that of their natural frame correspondents, it would make sense to do frame correspondence theory on modal fixed-point languages like the μ–calculus.

Higher-order model theory. Modal correspondence theory is a pilot study for the model theory of simple fragments of higher-order logic (van Benthem 1983). In particular, it would be of interest to extend its basic definability results. E.g., are modal formulas *LFP(FO)*-frame-definable iff, like all sentences in the latter fixed-point language, they are invariant for potential isomorphisms between frames? Or, can we generalize the Goldblatt-Thomason Theorem characterizing the modally definable elementary frame classes to frame classes that are *LFP(FO)*-definable? To do our model-theoretic analysis of syntactic formats then would require preservation theorems for fixed-point languages. But, as already observed in connection with monotonicity in Section 3, positive results of this sort are scarce, as the typical first-order routines used in the above proofs are no longer available.

5 Conclusion and further directions

We have analysed predicate minimization as a logical device, determining the circumstances when it is appropriate in both semantic and syntactic terms. Our main result is a syntactic characterization of all first-order formulas satisfying a semantic property of predicate intersection underlying many uses of minimization. When the latter device is added in full generality to first-order logic, the resulting formalism *MIN(FO)* provides an alternative to fixed-point languages like *LFP(FO).* Moreover, it sheds new light on old issues of frame definability in modal logic, leading to a new hierarchy of frame conditions with a natural level of fixed-point definability in between first-order and general higher-order. Eventually, this connection suggests a more thorough-going use of fixed-points, matching up stronger modal fixed-point languages like the μ–calculus with first-order fixed-point frame conditions.

6 Acknowledgment

Valentin Goranko, Martin Otto, and Andrzej Salas provided very helpful comments, and so did the anonymous *JSL* referee.

7 References

P. Aczel, 1977, 'An Introduction to Inductive Definitions', in J. Barwise, ed., *Handbook of Mathematical Logic*, North-Holland, 739–782, Amsterdam.

J. Barwise & J. van Benthem, 1999, 'Interpolation, Preservation, and Pebble Games', *Journal of Symbolic Logic* 64:2, 881–903.

J. van Benthem, 1983, *Modal Logic and Classical Logic*, Bibliopolis, Napoli.

J. van Benthem, 1985, *Logic Programming,* lecture notes, Philosophical Institute, Rijksuniversiteit Groningen.

J. van Benthem, 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford University.

R. Benton, 2002, 'A Simple Incomplete Extension of T', *Journal of Philosophical Logic* 31:6, 527–541.

P. Blackburn, M. de Rijke & Y. Venema, 2001, *Modal Logic*, Cambridge University Press, Cambridge.

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

H.C. Doets, 1994, *From Logic to Logic Programming*, The MIT Press, Cambridge (Mass.).

H-D Ebbinghaus & J. Flum, 1995, *Finite Model Theory*, Springer, Berlin.

D. Gabbay & H-J Ohlbach, 1992, 'Quantifier Elimination in Second-Order Predicate Logic', *South African Computer Journal* 7, 35–43.

V. Goranko & D. Vakarelov, 2003, 'Elementary Canonical Formulas I. Extending Sahlqvist's Theorem', Department of Mathematics, Rand Afrikaans University, Johannesburg & Faculty of Mathematics and Computer Science, Kliment Ohridski University, Sofia.

J-M Le Bars, 2002, 'The 0-1 Law fails for Frame Satisfiability of Propositional Modal Logic', *Proceedings LICS (Logic in Computer Science).*

B. Mahr & J. Makovsky, 1983, 'Characterizing Specification Languages which Admit Initial Semantics', *Proceedings 8th CAAP*, Springer, Berlin.

A.I. Malcev, 1971, *The Metamathematics of Algebraic Systems*, North-Holland, Amsterdam.

J. McCarthy, 1980, 'Circumscription – A Form of Nonmonotonic Reasoning', *Artificial Intelligence* 13, 27–39.

Y. N. Moschovakis, 1974, *Elementary Induction on Abstract Structures*, North-Holland, Amsterdam.

A. Nonnengart & A. Salas, 1999, 'A Fixed-Point Approach to Second-Order Quantifier Elimination with Applications to Modal Correspondence Theory', in E. Orlowska, ed., *Logic at Work*, Physica-Verlag, Heidelberg, 89–108.

M-A. Papalaskari & S. Weinstein, 1990, 'Minimal Consequence in Sentential Logic', *Journal of Logic Programming* 9, 19–31.

C. Stirling, 1999, 'Bisimulation, Modal Logic and Model Checking Games', *Logic Journal of the IGPL* 7:1, 103–124.

J. M. Weinstein, 1965, *First-Order Properties Preserved by Direct Products,* Ph. D. Thesis, University of Wisconsin, Madison.