Notre Dame Journal of Formal Logic
Volume 42, Number 2, 2001

The Fixed Point Property in Modal Logic

Lorenzo Sacchetti

Abstract This paper deals with the modal logics associated with (possibly nonstandard) provability predicates of Peano Arithmetic. One of our goals is to present some modal systems having the fixed point property and not extending the Gödel-Löb system **GL**. We prove that, for every $n \geq 2$, $K + \Box(\Box^{n-1} p \to p) \to \Box p$ has the explicit fixed point property. Our main result states that every complete modal logic L having the Craig's interpolation property and such that $L \vdash \Delta(\nabla(p) \to p) \to \Delta(p)$, where $\nabla(p)$ and $\Delta(p)$ are suitable modal formulas, has the explicit fixed point property.

1 Introduction

Gödel-Löb logic is the modal logic of the "standard" provability predicates where by standard provability predicate we mean a provability predicate $Pr(\cdot)$ satisfying the following conditions:

1. $\mathrm{PA} \vdash p$ iff $\mathrm{PA} \vdash Pr(\overline{\ulcorner p \urcorner})$,
2. $\mathrm{PA} \vdash Pr(\overline{\ulcorner p \to q \urcorner}) \to (Pr(\overline{\ulcorner p \urcorner}) \to Pr(\overline{\ulcorner q \urcorner}))$,
3. $\mathrm{PA} \vdash Pr(\overline{\ulcorner p \urcorner}) \to Pr(\overline{\ulcorner Pr(\overline{\ulcorner p \urcorner}) \urcorner})$,

where $\overline{\ulcorner p \urcorner}$ denotes the numeral for the Gödel number of p.

In fact, in view of a modal analysis of Gödel's Second Incompleteness Theorem and of arithmetical self-reference, condition 1 can be replaced by the weaker condition

(1′) if $\mathrm{PA} \vdash p$ then $\mathrm{PA} \vdash Pr(\overline{\ulcorner p \urcorner})$.

As is well known, Gödel-Löb logic (this system is often called **GL**, but it is also known as **G**, **L**, **PRL**, and **K4W**) is the modal logic axiomatized by the following schemes:

1. all the propositional tautologies in the modal language,

Received June 14, 2000; accepted May 14, 2001; printed May 22, 2003
2001 Mathematics Subject Classification: Primary, 03B45, 03F40; Secondary, 03F45, 03F30
Keywords: modal logic, provability predicates, fixed points

2. $\Box(A \to B) \to (\Box A \to \Box B)$,
3. $\Box(\Box A \to A) \to \Box A$,

together with the rules of modus ponens (MP), substitution, and necessitation (MN), that is, $A/\Box A$.

In general, if $Pr(\cdot)$ is a (possibly nonstandard) provability predicate, then the modal logic associated to $Pr(\cdot)$ is defined as follows: let $\mathcal{G}$ be the set of the functions f from the set of modal formulas to the set of formulas of PA such that f commutes with the propositional connectives and such that, for every sentence A, $f(\Box A) = Pr(\overline{\ulcorner f(A) \urcorner})$. The modal logic associated to $Pr(\cdot)$ is defined as the set of modal formulas A such that, for every $f \in \mathcal{G}$, $\mathrm{PA} \vdash f(A)$.

One of the most interesting features of the Gödel-Löb logic is that the Diagonalization Lemma is modally expressed in **GL** by the Fixed Point Theorem. This theorem splits into a uniqueness and an existence part and concerns formulas A with a distinguished propositional variable p that occurs in A only in the scope of a $\Box$ (see, e.g., Sambin and Valentini [9], Boolos [2], and Smoryński [11]).

The analysis of incompleteness phenomena, especially of the second incompleteness theorem, showed the dependence of the second Gödel's theorem not only on the extension of the provability predicate, but also on its intension (see Feferman [3] for discussion about intensionally correct provability predicates versus extensionally correct provability predicates): Feferman proved in fact that there are extensionally correct formal arithmetical statements $Pr(\cdot)$ expressing provability in PA such that the corresponding consistency statement $\neg Pr(\overline{\ulcorner \bot \urcorner})$ is provable.

Different provability predicates generate different modal logics. Thus the question arises of what these modal logics have in common. First, let us observe that every provability predicate $Pr(\cdot)$ must satisfy the following conditions.

1. $Pr(\cdot)$ numerates the set of theorems of PA.
2. $Pr(\cdot)$ satisfies the formalization of the closure under modus ponens, that is,

$$\mathrm{PA} \vdash Pr(\overline{\ulcorner p \to q \urcorner}) \to (Pr(\overline{\ulcorner p \urcorner}) \to Pr(\overline{\ulcorner q \urcorner})).$$

When based on the usual classical propositional logic, conditions 1 and 2 give us the basic modal logic **K**, that is, the modal logic having the same schemes and the same inference rules of **GL**, except the Löb scheme $\Box(\Box A \to A) \to \Box A$. Another important feature of a correct provability predicate $Pr(\cdot)$ is that

3. the Diagonalization Lemma holds for $Pr(\cdot)$.

In the seventies many authors (see Smoryński [10] and Boolos [1]) began a modal investigation of the fixed point property. Surprisingly, it turned out that a restricted version of this property, concerning formulas built from sentence variables, connectives, and the standard provability predicate, is provable using only the (purely modal) principles of **GL**. Turning our attention to nonstandard provability predicates, it is interesting to investigate which modal principles are needed to prove the fixed point property on purely modal grounds. In our opinion the investigation of the fixed point property in modal logics has a purely modal interest (i.e., solving fixed point equations). Moreover, even though we have not yet found any example of a nonstandard provability predicate whose provability logic does not extend **GL** but has the fixed point property, we are confident that such provability predicates exist and that the results proved in this paper will be helpful in view of a modal analysis of them.

To this purpose we introduce two kinds of fixed point properties: we say that a modal logic L (i.e., a system including all tautologies and distribution axioms, and closed under modus ponens, substitution, and necessitation) has the *nonexplicit fixed point property* if every formula $A(p)$, in which p is under the scope of a $\Box$, has a fixed point in every model of L, whereas we say that a logic L has the *explicit fixed point property* if for every formula $A(p)$, in which p is under the scope of a $\Box$, there exists a formula H containing only those variables of $A(p)$ other than p such that $L \vdash A(H) \leftrightarrow H$.

We will see that the fixed point property is strictly related to the following semantic properties:

1. the reverse well-foundedness of the accessibility relation,
2. a form of weak transitivity of the accessibility relation.

The outline of this paper is as follows. In Section 2 we shall recall some facts about the logics having the (explicit or nonexplicit) fixed point property; then we shall prove some general properties about them. In Section 3 we shall investigate the modal systems L such that $L \vdash \Delta(\nabla(p) \to p) \to \Delta(p)$, where $\Delta(p)$ and $\nabla(p)$ are modal formulas satisfying suitable conditions. The main result of this section is that, under suitable hypothesis, every logic L such that $L \vdash \Delta(\nabla(p) \to p) \to \Delta(p)$ has the fixed point property. As an example, in Section 4, we prove that, for every $n \geq 2$, the system $K + \Box(\Box^{n-1} p \to p) \to \Box p$ has the explicit fixed point property.

2 Preliminaries

First, we need some basic definitions.

Definition 2.1 Let p and $A(p)$ be given. Say that p is *boxed* in $A(p)$ if every occurrence of p in $A(p)$ lies within the scope of a $\Box$.

Definition 2.2 Let L be a modal logic. Say that L has the *explicit fixed point property* if for each formula $A(p, q_1, \ldots, q_n)$ in which the variable p is boxed, there is a formula $H(q_1, \ldots, q_n)$ such that

1. $H(q_1, \ldots, q_n)$ contains only those variables of $A(p, q_1, \ldots, q_n)$ other than p,
2. $L \vdash A(H(q_1, \ldots, q_n), q_1, \ldots, q_n) \leftrightarrow H(q_1, \ldots, q_n)$.

Any such formula $H(q_1, \ldots, q_n)$ is called a *fixed point* of $A(p, q_1, \ldots, q_n)$.

Definition 2.3 Let L be a logic. Say that L has the *nonexplicit fixed point property* if for every model $M = \langle X, R, \Vdash \rangle$ of L and for each $A(p)$ in which the variable p is boxed, there is a formula H such that $M \models A(H) \leftrightarrow H$.

It is obvious that every logic having the explicit fixed point property has the nonexplicit fixed point property.

Definition 2.4 Let $X_1 = \langle W_1, R_1 \rangle, \ldots, X_n = \langle W_n, R_n \rangle$ be frames. Then $\bigoplus_{i=1}^{n} X_i$ denotes the frame $\langle W, R \rangle$ where

1. $W = \{\bigcup_{i=1}^{n} \{W_i \times \{i\}\}\} \cup \{0\}$,
2. $0R(x, i)$ for all $(x, i) \in W$,
3. $(x, i)R(y, j)$ iff $i = j$ and $x R_i y$.

Definition 2.5 Say that L has the *amalgamation property* if whenever $X_1, \ldots, X_n$ are frames for L, $\bigoplus_{i=1}^{n} X_i$ is a frame for L.

Definition 2.6 Say that a logic L has *Craig's interpolation property* (see Maksimova [6]) if, for any formulas A and B, the condition $L \vdash A \to B$ implies $L \vdash A \to C$ and $L \vdash C \to B$ for some formula C, such that all its variables are in both A and B.

We begin with a brief survey of some facts regarding the fixed point property (see Sacchetti [8]).

1. The logic $K + \Box^n \bot$ has the explicit fixed point property.
2. Let L be a logic having the (explicit or nonexplicit) fixed point property. Thus
   (a) if L has the strong disjunction property then L is not canonical (i.e., the canonical model of L is based on a frame which is not a frame for L);
   (b) if L has the finite model property then L is complete with respect to a class of reverse well-founded frames;
   (c) L can be invalidated in every frame containing a cycle (in particular, in every frame containing a reflexive node), hence every finite frame for L is reverse well-founded;
   (d) every finite distinguishable model of L is reverse well-founded. Since for every finite model there is an equivalent (i.e., validating the same formulas) distinguishable model, every finite model of L is, up to equivalence, reverse well-founded.

We now turn to other results. Throughout this section, we assume L to be a logic having the (explicit or nonexplicit) fixed point property.

Proposition 2.7 *$L \subseteq K + \Box\bot$ and $L \not\subseteq K + \Box p \leftrightarrow p$.*

Proof If $L \subseteq K + \Box p \leftrightarrow p$ then the formula $\neg\Box q$ has no fixed point. Since every modal system is contained either in $K + \Box p \leftrightarrow p$ or in $K + \Box\bot$, the claim follows. □

It follows from Proposition 2.7 that every logic having the fixed point property is *compatible* with **GL** in the sense that if we add the axioms of **GL** to L we obtain a consistent logic. The following corollary states this more generally.

Corollary 2.8 *The union of any family of logics having the (explicit or nonexplicit) fixed point property is consistent.*

Proof The proof is trivial. □

We are going to prove a theorem regarding the connections between the fixed point property and the amalgamation property. Recall that by a *tree* is meant (see [11], p. 102) a frame $\langle X, R\rangle$ in which

1. R is a strict partial ordering, that is, R is transitive and asymmetric,
2. the set of predecessors of any element is finite and linearly ordered by R.

Observe that in this definition roots are not required.

Lemma 2.9 *If L has the amalgamation property then L is valid in every finite tree.*

Proof Finite trees can be defined by induction as follows:

1. for all x $\langle \{x\}, \varnothing\rangle$ is a tree,
2. if $T_1, \dots, T_n$ are trees so is $\bigoplus_{i=1}^{n} T_i$.

Now $\langle \{x\}, \varnothing \rangle$ models $K + \Box\bot$, so it models L. Since L has the amalgamation property if $T_1, \ldots, T_n$ model L, then $\bigoplus_{i=1}^{n} T_i$ models L. Thus every finite tree models L. □ □

Theorem 2.10 *Let L be a logic having the amalgamation property. For every formula A, if $L \vdash A$ then* $\mathbf{GL} \vdash A$.

Proof Let $L \vdash A$. Therefore A is valid in every frame for L, in particular, by Lemma 2.9, L is valid in every finite tree. Since $\mathbf{GL}$ is complete with respect to the class of finite trees, it follows that $\mathbf{GL} \vdash A$. □

3 The Main Theorem

We are interested in the systems having the form $K + \Delta(\nabla(p) \to p) \to \Delta(p)$, where $\Delta(p)$ and $\nabla(p)$ are modal formulas satisfying suitable conditions. The main theorem of this section states that, under some conditions on $\Delta(p)$ and $\nabla(p)$, every complete logic having the Craig's interpolation property and such that $L \vdash \Delta(\nabla(p) \to p) \to \Delta(p)$ has the explicit fixed point property.

3.1 A particular case We begin with the case in which $\Delta \equiv \nabla$.

Definition 3.1 Let $B(p)$ be a formula such that

1. if $L \vdash p$ then $L \vdash B(p)$,
2. $L \vdash B(p \to q) \to (B(p) \to B(q))$,
3. $L \vdash B(p) \to \Box p$,
4. $L \vdash B(B(p) \to p) \to B(p)$.

Proposition 3.2 $L \vdash B(p) \to B(B(p))$.

Proof It is just a repetition of the proof that $\mathbf{GL} \vdash \mathbf{4}$ (see, e.g., Hughes and Cresswell [5], p. 150). □

In the sequel, when there is no danger of confusion, we shall write Bp instead of $B(p)$.

Proposition 3.3 *Let $B_s p \equiv Bp \wedge p$, then*

1. $L \vdash B_s p \to p$,
2. $L \vdash Bp \to \Box B_s p$,
3. $L \vdash B_s p \to \Box B_s p$,
4. $L \vdash Bp \to B B_s p$,
5. *if* $L \vdash Bp \to q$ *then* $L \vdash Bp \to Bq$,
6. *if* $L \vdash B_s p \to q$ *then* $L \vdash Bp \to Bq$.

Proof We have for

1. it is an obvious consequence of the definition;
2. (a) $L \vdash Bp \to \Box p$ from Definition 3.1(3),
   (b) $L \vdash Bp \to BBp$ from Proposition 3.2,
   (c) $L \vdash BBp \to \Box Bp$ from Definition 3.1(3),
   (d) $L \vdash Bp \to \Box Bp$ from (2b) and (2c),
   (e) $L \vdash Bp \to \Box B_s p$ from (2a) and (2d).
3. it is an obvious consequence of (2).

Proofs of (4), (5), and (6) are completely parallel to the proofs of the analogous statements for **GL** (see [11], pp. 66–70, Lemma 1.4, Lemma 1.8, Lemma 1.11). One has just to replace $\Box$ by B. $\Box$

We now prove the First Substitution Lemma.

Lemma 3.4 (FSL) *Let $A(p)$ be given. Then $L \vdash B_s(E \leftrightarrow C) \to (A(E) \leftrightarrow A(C))$.*

Proof The proof is by induction on the complexity of $A(p)$. Atomic cases and propositional cases are trivial. Let $A(p)$ be $\Box D(p)$. We have

1. $L \vdash B_s(E \leftrightarrow C) \to (D(E) \leftrightarrow D(C))$ by induction hypothesis;
2. $L \vdash \Box B_s(E \leftrightarrow C) \to \Box(D(E) \leftrightarrow D(C))$ from (1) by MN and distributivity;
3. $L \vdash B_s(E \leftrightarrow C) \to \Box B_s(E \leftrightarrow C)$ by Proposition 3.3(3);
4. $L \vdash B_s(E \leftrightarrow C) \to \Box(D(E) \leftrightarrow D(C))$ from (2) and (3);
5. $L \vdash B_s(E \leftrightarrow C) \to (\Box D(E) \leftrightarrow \Box D(C))$ from (4) by distributivity;
6. $L \vdash B_s(E \leftrightarrow C) \to (A(E) \leftrightarrow A(C))$ from (5). $\Box$

We can now prove the Second Substitution Lemma.

Lemma 3.5 (SSL) *Let $A(p)$ be given. Then $L \vdash B(F \leftrightarrow C) \to \Box(A(F) \leftrightarrow A(C))$.*

Proof Set $D \equiv F \leftrightarrow C$ and $E \equiv A(F) \leftrightarrow A(C)$. Then

1. $L \vdash B_s(D) \to E$ by FSL;
2. $L \vdash \Box B_s(D) \to \Box E$ from (1) by MN and distributivity;
3. $L \vdash B(D) \to \Box B_s(D)$ by Proposition 3.3(2);
4. $L \vdash B(D) \to \Box E$ from (2) and (3). $\Box$

In the following proposition we prove the uniqueness of fixed points.

Proposition 3.6 (Uniqueness of Fixed Points) *Let p be boxed in $A(p)$ and let q be a new variable. Then $L \vdash B_s(p \leftrightarrow A(p)) \wedge B_s(q \leftrightarrow A(q)) \to (p \leftrightarrow q)$.*

Proof Let $A(p) \equiv P(\Box C_1(p), \ldots, \Box C_n(p))$ be a Boolean combination of $\Box C_1(p), \ldots, \Box C_n(p)$ and of formulas without occurrences of p. We have

1. $L \vdash B(p \leftrightarrow q) \to \Box(C_i(p) \leftrightarrow C_i(q))$ by SSL;
2. $L \vdash B(p \leftrightarrow q) \to (\Box C_i(p) \leftrightarrow \Box C_i(q))$ from (1) by distributivity;
3. $L \vdash B(p \leftrightarrow q) \to B(\Box C_i(p) \leftrightarrow \Box C_i(q))$ from (2) by Proposition 3.3(5);
4. $L \vdash B(p \leftrightarrow q) \to B_s(\Box C_i(p) \leftrightarrow \Box C_i(q))$ from (2) and (3);
5. $L \vdash B(p \leftrightarrow q) \to (A(p) \leftrightarrow A(q))$ from (4) by FSL;
6. $L \vdash [B_s(p \leftrightarrow A(p)) \wedge B_s(q \leftrightarrow A(q)) \wedge B(p \leftrightarrow q)] \to [(p \leftrightarrow A(p)) \wedge (q \leftrightarrow A(q)) \wedge (A(p) \leftrightarrow A(q))]$ from (5) and from the definition of B_s;
7. $L \vdash [B_s(p \leftrightarrow A(p)) \wedge B_s(q \leftrightarrow A(q)) \wedge B(p \leftrightarrow q)] \to (p \leftrightarrow q)$ from (6);
8. $L \vdash B_s(p \leftrightarrow A(p)) \wedge B_s(q \leftrightarrow A(q)) \to (B(p \leftrightarrow q) \to (p \leftrightarrow q))$ from (7);
9. $L \vdash B[(p \leftrightarrow A(p)) \wedge (q \leftrightarrow A(q))] \to B(B(p \leftrightarrow q) \to (p \leftrightarrow q))$ from (8) by Proposition 3.3(6);
10. $L \vdash B_s(p \leftrightarrow A(p)) \wedge B_s(q \leftrightarrow A(q)) \to B(p \leftrightarrow q)$ from (9) and from the Löb axiom for B;
11. $L \vdash B_s(p \leftrightarrow A(p)) \wedge B_s(q \leftrightarrow A(q)) \to (p \leftrightarrow q)$ from (10) and (8). $\Box$

Proposition 3.7 (Beth Definability Theorem) *Let L be a modal logic having the Craig's interpolation property. Let $D(r)$ be a formula that does not contain the atoms p or q. Suppose that $L \vdash D(p) \wedge D(q) \to (p \leftrightarrow q)$. Then there is a formula H such that*

1. *H contains only sentence letters contained in $D(p)$ other than p,*
2. *$L \vdash D(p) \to (p \leftrightarrow H)$.*

Proof See [1], Chapter 14. □

Proposition 3.8 *Let L be a logic having the Craig's interpolation property. Let p be boxed in $A(p)$. Then there exists a formula H containing only sentence letters contained in $A(p)$ other than p such that $L \vdash B(p \leftrightarrow A(p)) \to B(p \leftrightarrow H)$.*

Proof We have

1. $L \vdash B_s(p \leftrightarrow A(p)) \wedge B_s(q \leftrightarrow A(q)) \to (p \leftrightarrow q)$ by Proposition 3.6;
2. $L \vdash B_s(p \leftrightarrow A(p)) \to (p \leftrightarrow H)$ by Proposition 3.7;
3. $L \vdash B(p \leftrightarrow A(p)) \to B(p \leftrightarrow H)$ from (2) by Proposition 3.3(6). □

In the following lemmas we assume L, $A(p)$ and H as in the previous proposition, and we abbreviate $A(p, q_1, \ldots, q_s)$ with $A(p, \vec{q})$.

Lemma 3.9 $L \vdash B(p \leftrightarrow A(p, \vec{q})) \to (A(p, \vec{q}) \leftrightarrow A(H(\vec{q}), \vec{q}))$.

Proof As p is boxed in $A(p, \vec{q})$, there are formulas $A_1(p, \vec{q}), \ldots, A_n(p, \vec{q})$ such that $A(p, \vec{q})$ is a Boolean combination of $\Box A_1(p, \vec{q}), \ldots, \Box A_n(p, \vec{q})$ and of formulas without occurrences of p. Therefore, if $1 \leq i \leq n$, we have

1. $L \vdash B(p \leftrightarrow H(\vec{q})) \to \Box(A_i(p, \vec{q}) \leftrightarrow A_i(H(\vec{q}), \vec{q}))$ by SSL;
2. $L \vdash B(p \leftrightarrow H(\vec{q})) \to (\Box A_i(p, \vec{q}) \leftrightarrow \Box A_i(H(\vec{q}), \vec{q}))$ by distributivity;
3. $L \vdash B(p \leftrightarrow H(\vec{q})) \to (A(p, \vec{q}) \leftrightarrow A(H(\vec{q}), \vec{q}))$ from (2) by the propositional calculus;
4. $L \vdash B(p \leftrightarrow A(p, \vec{q})) \to B(p \leftrightarrow H(\vec{q}))$ by Proposition 3.8;
5. $L \vdash B(p \leftrightarrow A(p, \vec{q})) \to (A(p, \vec{q}) \leftrightarrow A(H(\vec{q}), \vec{q}))$ from (3) and (4). □

Lemma 3.10 $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \to (H(\vec{q}) \leftrightarrow A(H(\vec{q}), \vec{q}))$.

Proof We have

1. $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \to (p \leftrightarrow H(\vec{q}))$ by the proof of Proposition 3.8,
2. $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \to [(p \leftrightarrow H(\vec{q})) \wedge (p \leftrightarrow A(p, \vec{q}))]$ from (1) by definition of B_s,
3. $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \to (H(\vec{q}) \leftrightarrow A(p, \vec{q}))$ from (2),
4. $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \to (A(p, \vec{q}) \leftrightarrow A(H(\vec{q}), \vec{q}))$ from Lemma 3.9,
5. $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \to (H(\vec{q}) \leftrightarrow A(H(\vec{q}), \vec{q}))$ from (3) and (4). □

Lemma 3.11 *If C is a formula in which p does not occur and $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \to C$ then $L \vdash C$.*

Proof For the sake of readability we write $F(p)$ for $p \leftrightarrow A(p, \vec{q})$ (F comes from 'Fixed point'), $A(p)$ for $A(p, \vec{q})$, and $A(H)$ for $A(H(\vec{q}), \vec{q})$. Recall (Lemma 3.9) that

1. $L \vdash B(F(p)) \to (A(p) \leftrightarrow A(H))$. Now suppose $L \vdash B_s(F(p)) \to C$, where p does not occur in C. Then
2. $L \vdash \neg C \to (\neg F(p) \vee \neg B(F(p)))$, hence
3. $L \vdash \neg C \to [(\neg F(p)) \wedge B(F(p)) \vee \neg B(F(p))]$. Now from (1) and from the definition of $F(p)$ we have
4. $L \vdash (\neg F(p) \wedge B(F(p))) \to (\neg(p \leftrightarrow A(p)) \wedge (A(p) \leftrightarrow A(H)))$. Therefore,
5. $L \vdash (\neg F(p) \wedge B(F(p))) \to \neg(p \leftrightarrow A(H))$. Now using Löb's Theorem for B, and letting $\widetilde{B} \equiv \neg B \neg$,
6. $L \vdash \neg B(F(p)) \to \widetilde{B}(\neg F(p) \wedge B(F(p)))$. Using (5) and the monotonicity of $\widetilde{B}$,
7. $L \vdash \neg B(F(p)) \to \widetilde{B}(\neg(p \leftrightarrow A(H))$. Therefore,
8. $L \vdash \neg B(F(p)) \to \neg B(p \leftrightarrow A(H))$. Putting (3), (5), and (8) together, we obtain
9. $L \vdash \neg C \to [\neg(p \leftrightarrow A(H)) \vee \neg B(p \leftrightarrow A(H))]$. Therefore,
10. $L \vdash \neg C \to \neg B_s(p \leftrightarrow A(H))$. Now L is closed under substitution. Replacing p by $A(H)$ in (10) (recall that p does not occur in C) we obtain
11. $L \vdash \neg C \to \neg B_s(\top)$. Thus $L \vdash C$. □

We can now prove the fixed point theorem for L.

Theorem 3.12 (Explicit Definability of Fixed Points) *Let L be a logic having the Craig's interpolation property and such that there is a formula Bp satisfying the following conditions:*

1. *if $L \vdash p$ then $L \vdash Bp$,*
2. *$L \vdash B(p \to q) \to (Bp \to Bq)$,*
3. *$L \vdash Bp \to \Box p$,*
4. *$L \vdash B(Bp \to p) \to Bp$.*

Let p be boxed in $A(p, \vec{q})$. Then there is a formula $H(\vec{q})$ containing only sentence letters contained in $A(p, \vec{q})$ other than p such that

(a) *$L \vdash A(H(\vec{q}), \vec{q}) \leftrightarrow H(\vec{q})$,*
(b) *$L \vdash B(p \leftrightarrow A(p)) \to B(p \leftrightarrow H(\vec{q}))$.*

Proof
For (a) we have

(1) $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \to (H(\vec{q}) \leftrightarrow A(H(\vec{q}), \vec{q}))$ from Lemma 3.10;
(2) $L \vdash H(\vec{q}) \leftrightarrow A(H(\vec{q}), \vec{q})$ from (1) by Lemma 3.11.

For (b) we have

(1) $L \vdash A(H(\vec{q}), \vec{q}) \leftrightarrow H(\vec{q})$ from (a);
(2) $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \wedge B_s(H(\vec{q}) \leftrightarrow A(H(\vec{q}), \vec{q})) \to (p \leftrightarrow H(\vec{q}))$ from Proposition 3.6;
(3) $L \vdash B_s(A(H(\vec{q}), \vec{q}) \leftrightarrow H(\vec{q}))$ from (1);
(4) $L \vdash B_s(p \leftrightarrow A(p, \vec{q})) \to (p \leftrightarrow H(\vec{q}))$ from (2) and (3);
(5) $L \vdash B(p \leftrightarrow A(p)) \to B(p \leftrightarrow H(\vec{q}))$ from (4) by Proposition 3.3(6). □

We close this section with some results regarding modal formulas satisfying a suitable semantic property. We shall use the following results in the subsequent section.

Definition 3.13 Let $\langle X, R, \Vdash\rangle$ be a model; let x be a world in X. Let

1. $\overline{S_1(x)} = \{y | \exists n \geq 0 : xR^n y\}$,
2. $S_1(x) = \{y | \exists n > 0 : xR^n y\}$

and let

1. $\overline{S(x)} = \langle \overline{S_1(x)}, R \restriction \overline{S_1(x)}, \Vdash\restriction \overline{S_1(x)}\rangle$,
2. $S(x) = \langle S_1(x), R \restriction S_1(x), \Vdash\restriction S_1(x)\rangle$,

where $R \restriction S_1(x)$ denotes the restriction of R to $S(x)$ and $\Vdash\restriction S_1(x)$ denotes the restriction of $\Vdash$ to $S(x)$. The model $\overline{S(x)}$ is called the *submodel generated by* x.

Corollary 3.14 *Let L be a logic such that*

1. *L has the Craig's interpolation property;*
2. *there is a formula Bp such that*
   (a) *p is boxed in Bp,*
   (b) *$L \vdash B(Bp \to p) \to Bp$,*
   (c) *for every model $\langle X, R, \Vdash\rangle$ of L, for every $x \in X$ and for every formula φ, $x \Vdash B\varphi$ iff $S(x) \models \varphi$.*

Then L has the explicit fixed point property.

Proof It is easy to verify that if the formula Bp satisfies conditions 2a, 2b, and 2c, then it satisfies also conditions 1, 2, 3, and 4 of Theorem 3.12. □

In [8] we proved that $K + \Box^n\bot$ has the explicit fixed point property. Corollary 3.14 allows us to obtain an alternative proof of this fact.

Corollary 3.15 *For every $n \geq 1$, $K + \Box^n\bot$ has the explicit fixed point property.*

Proof It suffices to verify that $K + \Box^n\bot$ satisfies conditions 1 and 2 of Corollary 3.14. Set $Bp \equiv \bigwedge_{i=1}^{n} \Box^i p$. It is easy to verify that Bp satisfies conditions 2a, 2b, and 2c. It remains to be proven that $K + \Box^n\bot$ has Craig's interpolation property. Observe that K has Craig's interpolation property (see Gabbay [4]), and $K + \Box^n\bot$ is axiomatized over K by formulas without variables (i.e., $K + \Box^n\bot$ is a *constant extension*—see [6] and Rautenberg [7]—of K). In [7] Rautenberg proves that if a logic L has the Craig's interpolation property, then any constant extension of L has Craig's interpolation property. It follows that $K + \Box^n\bot$ possesses Craig's interpolation property. Therefore condition 1 is also satisfied. □

Proposition 3.16 *Let L be a logic having the (explicit or nonexplicit) fixed point property. Assume that there is a formula Bp such that*

1. *p is boxed in Bp,*
2. *for every model $\langle X, R, \Vdash\rangle$ of L, for every $x \in X$, and for every formula φ, $x \Vdash B\varphi$ iff $S(x) \models \varphi$.*

Then every frame for L is reverse well-founded.

Proof Let $\langle X, R\rangle$ be a frame for L. By way of contradiction assume that there is a chain $\alpha_0 R \alpha_1 R \ldots$ of length ω. Set $\Lambda = \{\alpha_0, \alpha_1, \ldots\}$. Let q be a sentence letter. Set $A(p, q) \equiv \neg Bp \vee q$. For every node x in X let $x \Vdash q$ if and only if $x \notin \Lambda$. For any sentence letter r, $r \neq q$, let the forcing be defined arbitrarily. It is easy to prove that in this model $A(p, q)$ does not have any fixed point. □

Proposition 3.17 *Let L be a logic. Assume that there is a formula Bp such that*

1. *p is boxed in Bp,*
2. *for every model $\langle X, R, \Vdash\rangle$ of L, for every $x \in X$, and for every formula φ, $x \Vdash B\varphi$ iff $S(x) \models \varphi$.*

Let $\langle X, R\rangle$ be a frame of L. Then the following statements are equivalent:

1. *R is reverse well-founded;*
2. *$\langle X, R\rangle \models B(Bp \to p) \to Bp$.*

Proof Let $\langle X, R\rangle$ be a frame of L. Let R^* be the transitive closure of R. Then under our assumptions, B is the modal operator associated with R^*. Now R^* is transitive, and it is reverse well-founded if and only if R is such. Moreover, it is well known (see [11] or [1]) that reverse well-foundedness for transitive relations can be modally expressed by the Löb axiom. The claim follows. □

3.2 The general case In this section we shall turn to the general case, that is, to the modal systems of the kind $K + \Delta(\nabla(p) \to p) \to \Delta(p)$, where $\Delta(p)$ and $\nabla(p)$ are modal formulas satisfying suitable conditions.

Definition 3.18 Say that a modal formula $\sigma(p)$ has the property $(*)$ if for every frame $\langle X, R\rangle$ and for every node x in X there is a subset $T_\sigma(x)$ of X such that for every model $M = \langle X, R, \Vdash\rangle$ based on $\langle X, R\rangle$ and for every formula φ we have $x \Vdash \sigma(\varphi)$ if and only if $T_\sigma(x) \models \varphi$ (where $T_\sigma(x) \models \varphi$ means for every $y \in T_\sigma(x)$, $y \Vdash \varphi$).

In the sequel, when there is no danger of confusion, we shall write Δp, ∇p, and σp instead of $\Delta(p)$, $\nabla(p)$, and $\sigma(p)$.

Proposition 3.19 *Let $\sigma_1 p$ and $\sigma_2 p$ be modal formulas having the property $(*)$. Then*

1. *the formula $\sigma_1\sigma_2 p$ has the property $(*)$ with respect to*
$$T_{\sigma_1\sigma_2}(x) = \{y | \exists z \in T_{\sigma_1}(x) : y \in T_{\sigma_2}(z)\};$$
2. *for every frame $\langle X, R\rangle$ the following facts are equivalent:*
   (a) *for every $x \in X$, $T_{\sigma_1\sigma_2}(x) \subseteq T_{\sigma_1}(x)$,*
   (b) *$\langle X, R\rangle \models \sigma_1 p \to \sigma_1\sigma_2 p$;*
3. *for every logic L, if $L \vdash p$, then $L \vdash \sigma_i p$, for i=1,2;*
4. *for every logic L, $L \vdash \sigma_i(p \to q) \to (\sigma_i p \to \sigma_i q)$, for i=1,2;*
5. *for every frame $\langle X, R\rangle$ and for all $x, y, z \in X$, if $y \in T_{\sigma_1}(x)$ and $z \in T_{\sigma_2}(y)$, then $z \in T_{\sigma_1\sigma_2}(x)$.*

Proof It is a routine matter. □

Definition 3.20 Say that a modal formula σp has the property $(**)$ if

1. σp has the property $(*)$,
2. there exists $n \in \mathbb{N}$ such that for every logic L, $L \vdash \sigma p \to \Box^n p$.

From the definition it follows that if σp has the property $(**)$ then there exists $n \in \mathbb{N}$ such that for every frame $\langle X, R\rangle$, for all $x, y \in X$, for every forcing relation $\Vdash$ on X and for every formula φ, if $x R^n y$ and $x \Vdash \sigma\varphi$, then $y \Vdash \varphi$.

Example 3.21 Let $A_1, \ldots, A_n$ be closed formulas, and let $k_1, \ldots, k_n$ be positive integers. Then $\Delta p \equiv \bigwedge_{i=1}^{n} \Box^{k_i}(A_i \to p)$ satisfies the property $(*)$. If, in addition, $K \vdash A_i$ for some A_i, then Δp satisfies the property $(**)$.

In the sequel, we shall assume that $\sigma_1 p$ and $\sigma_2 p$ are formulas such that, for every logic L and for every $i = 1, 2$,

(a) if $L \vdash p$ then $L \vdash \sigma_i p$,
(b) $L \vdash \sigma_i(p \to q) \to (\sigma_i p \to \sigma_i q)$.

Proposition 3.22 *Let L be a logic. If $L \vdash \sigma_1(\sigma_2 p \to p) \to \sigma_1 p$ then $L \vdash \sigma_1 p \to \sigma_1 \sigma_2 p$.*

Proof We have

1. $L \vdash p \to ((\sigma_2 p \wedge \sigma_2\sigma_2 p) \to (p \wedge \sigma_2 p))$, tautology;
2. $L \vdash \sigma_1 p \to \sigma_1((\sigma_2 p \wedge \sigma_2\sigma_2 p) \to (p \wedge \sigma_2 p))$ from (1) by rules (a) and (b);
3. $L \vdash \sigma_1 p \to \sigma_1(\sigma_2(p \wedge \sigma_2 p) \to (p \wedge \sigma_2 p))$ from (2) and by the distributivity of σ_2;
4. $L \vdash \sigma_1(\sigma_2(p \wedge \sigma_2 p) \to (p \wedge \sigma_2 p)) \to \sigma_1(p \wedge \sigma_2 p)$ by hypothesis;
5. $L \vdash \sigma_1 p \to \sigma_1(p \wedge \sigma_2 p)$ from (3) and (4);
6. $L \vdash \sigma_1 p \to \sigma_1\sigma_2 p$ from (5). □

Corollary 3.23 *Assume L to be a logic such that $L \vdash \Delta(\nabla p \to p) \to \Delta p$ where Δp and ∇p are formulas having the property $(*)$. Then $L \vdash \Delta p \to \Delta\nabla p$.*

Proof It follows from Proposition 3.22 and Proposition 3.19. □

Proposition 3.24 *Let L be a logic. If $L \vdash \sigma_1 p \to \sigma_1\sigma_2 p$ then, for every $n \in \mathbb{N}$, $L \vdash \sigma_1 p \to \sigma_1\sigma_2^n p$.*

Proof The proof is easy. □

This allows us to prove the following proposition.

Proposition 3.25 *Let L be a logic. If $L \vdash \sigma_1(\sigma_2 p \to p) \to \sigma_1 p$ then, for every $n \in \mathbb{N}$, $L \vdash \sigma_1 p \to \sigma_1\sigma_2^n p$.*

Proof It follows from Proposition 3.22 and from Proposition 3.24. □

Proposition 3.26 *Let Δp be a formula having the property $(*)$. For every frame $\langle X, R\rangle$ and for every $x \in X$ we have*

1. *$T_\Delta(x) \subseteq \overline{S_1(x)}$,*
2. *if p is boxed in Δp then $T_\Delta(x) \subseteq S_1(x)$.*

Proof The proof is obvious. □

In the sequel, we shall assume Δp and ∇p to be formulas such that

1. Δp and ∇p have the property $(**)$,
2. p is boxed in Δp and in ∇p.

Proposition 3.27 *Let L be a logic. If $L \vdash \Delta p \to \Delta\nabla p$ then there exists a formula Bp such that*

1. *p is boxed in Bp,*

2. *for every model* $\langle X, R, \Vdash \rangle$ *of* L*, for every* $x \in X$*, and for every formula* φ*,* $x \Vdash B\varphi$ *if and only if* $S(x) \models \varphi$*.*

Proof Let $n, m \in \mathbb{N}$ be such that $L \vdash \Delta p \to \Box^n p$ and $L \vdash \nabla p \to \Box^m p$. Now let $Bp \equiv \bigwedge_{i=1}^{n} \Box^i p \wedge \Delta \bigwedge_{j=1}^{m} \Box^j p$. We claim that for all h, $L \vdash Bp \to \Box^h p$. The claim is trivial for $h \leq n$.

Now let $h > n$. We can write h as $h = n + qm + r$, where $r < m$. Then $Bp \to \Delta\Box^r p$. Since, by Proposition 3.24, $L \vdash \Delta p \to \Delta\nabla^q p$, substituting $\Box^r p$ for p we get $L \vdash \Delta\Box^r p \to \Delta\nabla^q\Box^r p$. Therefore $L \vdash \Delta\Box^r p \to \Box^{n+mq+r} p$ and finally $L \vdash Bp \to \Box^{n+mq+r} p$. Thus $L \vdash Bp \to \Box^h p$ for all h. It follows that if $x \Vdash B\varphi$ then $S(x) \models \varphi$.

For the other direction, if $S(x) \models \varphi$ then $x \Vdash \Box^i\varphi$ for $i = 1, \ldots, n$ and $x \Vdash \Delta\Box^i\varphi$ for $i = 1, \ldots, m$, as $T_{\Delta\Box^i}(x) = \{y | \exists z \in T_\Delta(x) : zR^i y\}$ and, since $T_\Delta(x) \subseteq S_1(x)$, $T_{\Delta\Box^i}(x) \subseteq S_1(x)$. It follows that if $S(x) \models \varphi$, then $x \Vdash B\varphi$. □

Corollary 3.28 *Let* L *be a logic. If* L *has the (explicit or nonexplicit) fixed point property and* $L \vdash \Delta p \to \Delta\nabla p$ *then every frame for* L *is reverse well-founded.*

Proof It is an immediate consequence of Propositions 3.16 and 3.27. □

The following proposition proves an interesting result about the modal definability of well-foundedness. It is well known (see, e.g., van Benthem [12]) that well-foundedness is not modally definable alone but is modally definable together with transitivity. The following proposition states that well-foundedness is modally definable not only together with transitivity but also together with weaker conditions.

Proposition 3.29 *For every frame* $\langle X, R \rangle$*, the following facts are equivalent.*

1. R *is reverse well-founded and, for every* $x \in X$*,* $T_{\Delta\nabla}(x) \subseteq T_\Delta(x)$*,*
2. $\langle X, R \rangle \models \Delta(\nabla p \to p) \to \Delta p$*.*

Proof (1 $\Rightarrow$ 2) Suppose (1) holds. Let, by contradiction, $x \in X$ be such that $x \Vdash \Delta(\nabla\varphi \to \varphi)$ and $x \nVdash \Delta\varphi$ for some formula φ and some forcing $\Vdash$. Then there is $x_1 \in T_\Delta(x)$ such that $x_1 \nVdash \varphi$. Since $x \Vdash \Delta(\nabla\varphi \to \varphi)$, $x_1 \Vdash \nabla\varphi \to \varphi$. So $x_1 \nVdash \nabla\varphi$. Thus there is $x_2 \in T_\nabla(x_1)$ such that $x_2 \nVdash \varphi$. Now from $x_1 \in T_\Delta(x)$ and $x_2 \in T_\nabla(x_1)$ we deduce, by Proposition 3.19, $x_2 \in T_{\Delta\nabla}(x)$, and, by (1), $x_2 \in T_\Delta(x)$. Again, $x_2 \nVdash \varphi$ and $x_2 \Vdash \nabla\varphi \to \varphi$, therefore $x_2 \nVdash \nabla\varphi$. Iterating we get $x_1, x_2, \ldots, x_n, \ldots$ such that $x_{i+1} \in T_\nabla(x_i)$. Since, by Proposition 3.26, $T_\nabla(x_i) \subseteq S_1(x_i)$, there are $m_1, m_2, \ldots, m_n, \ldots$ such that $x_1 R^{m_1} x_2 R^{m_2} x_3 \ldots x_n R^{m_n} x_{n+1}$, contradicting the well-foundedness of R.

(2 $\Rightarrow$ 1) From Proposition 3.22 and from Proposition 3.19 it follows that for every $x \in X$, $T_{\Delta\nabla}(x) \subseteq T_\Delta(x)$. It remains to be proven that R is reverse well-founded. By contraposition, assume that $\langle X, R \rangle$ contains a chain $\alpha_0 R \alpha_1 R \ldots$ of length ω. Set $\Lambda = \{\alpha_0, \alpha_1, \ldots\}$. Define $\Vdash$ by $x \Vdash p$ if and only if $x \notin \Lambda$, for any atom p and any node $x \in X$. We claim that every node forces $\nabla p \to p$. The claim is trivial for all nodes x such that $x \notin \Lambda$. Assume that $x \in \Lambda$. Let $x = \alpha_k$. By definition, $x \nVdash p$. Since ∇p has the property $(**)$ there is $m \in \mathbb{N}$ such that for every $y \in X$, if $xR^m y$ and $x \Vdash \nabla p$ then $y \Vdash p$. Since $\alpha_k R^m \alpha_{m+k}$ and $\alpha_{m+k} \nVdash p$, it follows that $\alpha_k \nVdash \nabla p$. Hence $x \Vdash \nabla p \to p$. Therefore, for every node $x \in X$, $x \Vdash \nabla p \to p$. In particular, for every node $y \in T_\Delta(\alpha_0)$, $y \Vdash \nabla p \to p$. Therefore $\alpha_0 \Vdash \Delta(\nabla p \to p)$. Since Δp has the property $(**)$, there is $n \in \mathbb{N}$ such that for every $y \in X$, if $\alpha_0 R^n y$

and $\alpha_0 \Vdash \Delta p$ then $y \Vdash p$. Since $\alpha_0 R^n \alpha_n$ and $\alpha_n \nVdash p$, it follows that $\alpha_0 \nVdash \Delta p$. Therefore $\alpha_0 \nVdash \Delta(\nabla p \to p) \to \Delta p$. $\square$

Proposition 3.30 *Let L be a logic. If L is complete and $L \vdash \Delta(\nabla p \to p) \to \Delta p$ then there is a modal formula Bp such that*

1. *p is boxed in Bp,*
2. *for every model $\langle X, R, \Vdash\rangle$ of L, for every $x \in X$, and for every formula φ, $x \Vdash B\varphi$ iff $S(x) \models \varphi$,*
3. *$L \vdash B(Bp \to p) \to Bp$.*

Proof If $L \vdash \Delta(\nabla p \to p) \to \Delta p$ then $L \vdash \Delta p \to \Delta\nabla p$. Therefore, by Proposition 3.27 there exists a modal formula Bp satisfying conditions 1 and 2. From Proposition 3.29 it follows that every frame for L is reverse well-founded. From Proposition 3.17 it follows that $B(Bp \to p) \to Bp$ is valid in every frame for L, hence, by completeness, $L \vdash B(Bp \to p) \to Bp$. $\square$

We can now prove the main theorem of this section.

Theorem 3.31 *Let L be a logic such that*

1. *L has Craig's interpolation property,*
2. *L is complete,*
3. *$L \vdash \Delta(\nabla p \to p) \to \Delta p$, where Δp and ∇p are formulas having the property $(**)$ and such that p is boxed in Δp and ∇p.*

Then L has the explicit fixed point property.

Proof From Proposition 3.30 it follows that $L \vdash B(Bp \to p) \to Bp$ where Bp is a modal formula such that

1. p is boxed in Bp,
2. for every model $\langle X, R, \Vdash\rangle$ of L, for every $x \in X$, and for every formula φ, we have $x \Vdash B\varphi$ iff $S(x) \models \varphi$.

Since L has the Craig's interpolation property, from Corollary 3.14 it follows that L has the explicit fixed point property. $\square$

4 An Example: Weak Transitivity

Since any complete modal logic whose frames are transitive and reverse well-founded is an extension of **GL**, in order to find new modal logics having the explicit fixed point property, we will relax the transitivity condition, and we shall replace it by a weaker condition.

We begin with the semantical notion of a weak transitive relation. Before turning to the definition, we state the idea. Informally, just as transitivity can be thought of as the condition imposing that if a node x sees a node y in two steps then x can see y in one step, we can think of weak transitivity as the condition imposing that if a node x sees a node y in n steps (where $n \in \mathbb{N}$, $n \geq 2$), then x can see y in one step.

Definition 4.1 Let $\langle X, R\rangle$ be a frame. Say that R is *weakly transitive* if there is $n \in \mathbb{N}$, $n \geq 2$, such that for all $x, y \in X$, we have if $xR^n y$ then xRy.

Every transitive frame is obviously weakly transitive. Therefore any frame for **GL** is weakly transitive. Furthermore, it is easy to see that every frame for $K + \Box^n\bot$ is weakly transitive. Therefore all the logics having the fixed point property known so far are valid in weakly transitive frames.

Consider now the following formulas $\Box p$ and $\Box^{n-1} p$. It is easy to prove that each of them has the property $(**)$, and that, for every frame $\langle X, R\rangle$ and each node $x \in X$, we have $T_{\Box}(x) = \{y|xRy\}$ and $T_{\Box^{n-1}}(x) = \{y|xR^{n-1}y\}$.

Proposition 4.2 *Let $n \in \mathbb{N}$, $n \geq 2$, be given. Then*

1. *for every frame $\langle X, R\rangle$ the following statements are equivalent:*
   (a) $\langle X, R\rangle \models \Box p \to \Box^n p$,
   (b) *for all $x, y \in X$, if $xR^n y$ then xRy;*
2. $K + \Box(\Box^{n-1} p \to p) \to \Box p \vdash \Box p \to \Box^n p$;
3. *for every frame $\langle X, R\rangle$ the following statements are equivalent:*
   (a) *R is reverse well-founded and for all $x, y \in X$, if $xR^n y$ then xRy,*
   (b) $\langle X, R\rangle \models \Box(\Box^{n-1} p \to p) \to \Box p$.

Proof (1) follows from Proposition 3.19, (2) follows from Proposition 3.22, and (3) follows from Proposition 3.29, but it is a routine matter to prove them directly. $\Box$

Our goal is to reach the fixed point theorem for $K + \Box(\Box^{n-1} p \to p) \to \Box p$. It suffices to verify that $K + \Box(\Box^{n-1} p \to p) \to \Box p$ satisfies the hypothesis of Theorem 3.31. As we have already remarked, the formulas $\Box p$ and $\Box^{n-1} p$ have the property $(**)$. It remains to be proven that $K + \Box(\Box^{n-1} p \to p) \to \Box p$ is complete and has the Craig's interpolation property.

4.1 Completeness Our first task is to prove that $K + \Box(\Box^{n-1} p \to p) \to \Box p$ has the finite model property. Let L be the logic $K + \Box(\Box^{n-1} p \to p) \to \Box p$.

Definition 4.3 Let A be a modal formula such that $L \nvdash A$. Let

$$\begin{aligned} S &= \{\varphi|\varphi \text{ is a subformula of A}\}, \\ S_1 &= \{\varphi, \neg\varphi|\varphi \in S\} \cup \{\Box^i\varphi, \neg\Box^i\varphi|1 \leq i \leq 2n, \varphi \in S\}. \end{aligned}$$

Definition 4.4 Let X be the family of maximal consistent subsets of S_1. For all $x, y \in X$ we define xRy if and only if

1. for all $\varphi \in S$, if $\Box\varphi \in x$ then $\varphi \in y$;
2. for all $\varphi \in S$ and for all $i < n$, if $\Box^{i+1}\varphi \in x$ then $\Box^i\varphi \in y$;
3. there is some $\varphi \in S$ such that $\Box^{n-1}\varphi \in y$ and $\Box\varphi \notin x$.

Obviously, X is finite.

Lemma 4.5 *R is weakly transitive, irreflexive, and acyclic.*

Proof We prove that R is weakly transitive. Suppose that $x_0 R^n x_n$. We must show that $x_0 R x_n$. We prove that x_0 and x_n satisfy condition 1. Let $\varphi \in S$ such that $\Box\varphi \in x_0$. Then $\Box^n\varphi \in S_1$ and, since $L \vdash \Box\varphi \to \Box^n\varphi$, $\Box^n\varphi \in x_0$. Therefore $\Box^{n-1}\varphi \in x_1, \Box^{n-2}\varphi \in x_2, \ldots, \varphi \in x_n$. We now prove that x_0 and x_n satisfy condition 2. Let $\varphi \in S$ such that $\Box^{i+1}\varphi \in x_0$. Then $\Box^{n+i}\varphi \in S_1$ so, since $L \vdash \Box^{i+1}\varphi \to \Box^{n+i}\varphi$, $\Box^{n+i}\varphi \in x_0$. It follows that $\Box^{n+i-1}\varphi \in x_1, \ldots,$ $\Box^{n+i-n}\varphi \in x_n$. Then $\Box^i\varphi \in x_n$. We finally prove that x_0 and x_n satisfy condition 3. From $x_0 R x_1$ it follows that there is some $\varphi \in S$ such that $\Box^{n-1}\varphi \in x_1$ and $\Box\varphi \notin x_0$. Therefore from $\Box^{2n-2}\varphi \in S_1$ and from $L \vdash \Box^{n-1}\varphi \to \Box^{2n-2}\varphi$ it follows that $\Box^{2n-2}\varphi \in x_1$. Thus $\Box^{2n-3}\varphi \in x_2, \ldots, \Box^{2n-1-n}\varphi \in x_n$, that is, $\Box^{n-1}\varphi \in x_n$.

We prove that R is irreflexive. Suppose, by way of contradiction, that there is some $x \in X$ such that xRx. From (3) it follows that there is $\varphi \in S$ such that $\Box^{n-1}\varphi \in x$ and $\Box\varphi \notin x$. Thus, by (2), $\Box^{n-2}\varphi \in x, \dots, \Box\varphi \in x$, a contradiction.

We finally prove that there are no loops. Suppose that $x_0 R^k x_k R x_0$. Since R is weakly transitive we can assume, without loss of generality, that $k < n$. Moreover, if $k = n$ then we would deduce $x_0 R x_0$, against the irreflexiveness of R. So $k < n - 1$ and $k + 1 < n$. Let $x_0 R x_1 R \dots R x_k R x_0$ be a loop of minimal length. Let $n = q(k+1) + r$, with $r < k + 1$. Therefore $x_0 R^{q(k+1)+r} x_r$. Since R is weakly transitive, it follows that $x_0 R x_r$. Since by hypothesis the loop has minimal length, we cannot have $r \neq 1$, otherwise there would be a shorter loop. Therefore $r = 1$ and $n \equiv 1 \mod(k+1)$. From $x_0 R x_1$ it follows that there is some $\varphi \in S$ such that $\Box^{n-1}\varphi \in x_1$ and $\Box\varphi \notin x_0$. Therefore $\Box^{n-2}\varphi \in x_2, \dots, \Box^{n-k}\varphi \in x_k, \Box^{n-k-1}\varphi \in x_0, \Box^{n-k-2}\varphi \in x_1, \dots, \Box^{n-2k-1}\varphi \in x_k$, $\Box^{n-2k-2}\varphi \in x_0$. By an iteration argument we obtain $\Box^{n-q(k+1)}\varphi \in x_0$, that is, $\Box^r\varphi \in x_0$. Since $r = 1$ we have $\Box\varphi \in x_0$, a contradiction. $\Box$

Definition 4.6 Let $x \in X$. For any atom p, $x \Vdash p$ if and only if $p \in x$.

Lemma 4.7 *For all $\varphi \in S$ and for all $x \in X$, $x \Vdash \varphi$ if and only if $\varphi \in x$.*

Proof The proof is by induction on the complexity of φ. If φ is atomic the result holds by definition of $\Vdash$. Boolean cases are almost trivial.

Let $\varphi \equiv \Box\psi$. Let $\Box\psi \in x$ and xRy. Then $\psi \in y$. By induction hypothesis, $y \Vdash \psi$. Therefore if $\Box\psi \in x$ and xRy then $y \Vdash \psi$. It follows that $x \Vdash \Box\psi$. Vice versa, assume that $\Box\psi \notin x$. Define

$$\overline{y} = \{\gamma | \gamma \in S, \Box\gamma \in x\} \cup \bigcup_{i=1}^{n-1} \{\Box^i\delta | \delta \in S, \Box^{i+1}\delta \in x\} \cup \{\neg\psi, \Box^{n-1}\psi\}.$$

We claim that $\overline{y}$ is consistent. For otherwise

$$\{\gamma | \gamma \in S, \Box\gamma \in x\} \cup \bigcup_{i=1}^{n-1} \{\Box^i\delta | \delta \in S, \Box^{i+1}\delta \in x\} \vdash \Box^{n-1}\psi \to \psi.$$

From $\Gamma \vdash A$ it follows that $\Box\Gamma \vdash \Box A$. Therefore

$$\{\Box\gamma | \gamma \in S, \Box\gamma \in x\} \cup \bigcup_{i=1}^{n-1} \{\Box^{i+1}\delta | \delta \in S, \Box^{i+1}\delta \in x\} \vdash \Box(\Box^{n-1}\psi \to \psi).$$

Thus $\{\Box\gamma | \gamma \in S, \Box\gamma \in x\} \cup \bigcup_{i=1}^{n-1} \{\Box^{i+1}\delta | \delta \in S, \Box^{i+1}\delta \in x\} \vdash \Box\psi$. Since

$$\{\Box\gamma | \gamma \in S, \Box\gamma \in x\} \cup \bigcup_{i=1}^{n-1} \{\Box^{i+1}\delta | \delta \in S, \Box^{i+1}\delta \in x\} \subseteq x,$$

we have $x \vdash \Box\psi$. By the maximality of x, and since $\Box\psi \in S_1$, we obtain $\Box\psi \in x$, a contradiction.

Therefore $\overline{y}$ is consistent so that $\overline{y}$ can be extended to a maximal consistent set y. We must prove that xRy. Let $\Box\gamma \in x$, $\gamma \in S$. Then, by definition, $\gamma \in \overline{y}$, hence $\gamma \in y$. This proves condition 1. Let $\Box^{i+1}\gamma \in x$, $\gamma \in S$, then, by definition, $\Box^i\gamma \in \overline{y}$, then $\Box^i\gamma \in y$. This proves condition 2. Finally, $\Box\psi \notin x$ and $\Box^{n-1}\psi \in \overline{y} \subseteq y$. Then condition 3 is also satisfied.

Therefore xRy. Since $\neg\psi \in \overline{y} \subseteq y$, $\psi \notin y$. By the induction hypothesis $y \not\Vdash \psi$. Thus there is some $y \in X$ such that xRy and $y \not\Vdash \psi$. This yields that $x \not\Vdash \Box\psi$. $\Box$

Theorem 4.8 *For all $n \geq 2$, $K + \Box(\Box^{n-1}p \to p) \to \Box p$ has the finite model property.*

Proof From the lemmas it follows that if $K + \Box(\Box^{n-1}p \to p) \to \Box p \nvdash A$ then there is a finite model in which the accessibility relation is weakly transitive and reverse well-founded, invalidating A. $\Box$

4.2 Interpolation We now must prove that $K + \Box(\Box^{n-1}p \to p) \to \Box p$ (which in the sequel we abbreviate by L) has the Craig's interpolation property. Fix A, B and suppose no interpolant exists. We must show that $L \nvdash A \to B$. Let $\mathcal{L}(A)$ be the set consisting of all formulas possessing only atoms occurring in A and let $\mathcal{L}(B)$ be the set consisting of all formulas possessing only atoms occurring in B.

Definition 4.9 Two finite sets $X \subseteq \mathcal{L}(A)$, $Y \subseteq \mathcal{L}(B)$ are said to be *separable* if there is a formula $C \in \mathcal{L}(A) \cap \mathcal{L}(B)$ such that

1. $L \vdash \bigwedge\!\!\bigwedge X \to C$,
2. $L \vdash \bigwedge\!\!\bigwedge Y \to \neg C$.

Say that X and Y are *inseparable* if they are not separable.

With this terminology, we can restate our hypothesis that there is no interpolant between A and B by saying that the sets $\{A\}$ and $\{\neg B\}$ are inseparable.

Remark 4.10 If X and Y are inseparable, then $L + \bigwedge\!\!\bigwedge X$ and $L + \bigwedge\!\!\bigwedge Y$ are consistent.

Proof Suppose that $L + \bigwedge\!\!\bigwedge X$ is not consistent. Therefore we have

$$L \vdash \bigwedge\!\!\bigwedge X \to \bot \text{ and } L \vdash \bigwedge\!\!\bigwedge Y \to \neg\bot,$$

contrary to the inseparability of X and Y. The case in which $L + \bigwedge\!\!\bigwedge Y$ is not consistent is treated in a similar manner. $\Box$

Definition 4.11 For any formula D let

$$\begin{aligned} S(D) &= \{\varphi | \varphi \text{ is a subformula of } D\}, \\ S_1(D) &= \{\varphi, \neg\varphi | \varphi \in S(D)\} \cup \{\Box^i\varphi, \neg\Box^i\varphi | 1 \leq i \leq 2n, \varphi \in S(D)\}. \end{aligned}$$

Definition 4.12 A pair $X \subseteq S_1(A)$ and $Y \subseteq S_1(\neg B)$ of sets is called S_1-*complete* if

1. X and Y are inseparable,
2. for all $D \in S_1(A)$, either $D \in X$ or $\neg D \in X$,
3. for all $D \in S_1(\neg B)$, either $D \in Y$ or $\neg D \in Y$.

Thus if X,Y form an S_1-complete pair, then X and Y are maximal consistent with respect to $S_1(A)$ and $S_1(\neg B)$, respectively. Moreover, the three conditions imply for any $D \in S_1(A) \cap S_1(\neg B)$, $D \in X$ if and only if $D \in Y$ (if, e.g., there is $D \in S_1(A) \cap S_1(\neg B)$ such that $D \in X$ and $D \notin Y$ then $D \in X$ and $\neg D \in Y$against the inseparability of X, Y).

Lemma 4.13 *Let $X_0 \subseteq S_1(A)$, $Y_0 \subseteq S_1(\neg B)$ be inseparable. Then there is an S_1-complete pair X, Y such that $X_0 \subseteq X$, $Y_0 \subseteq Y$.*

Proof Let $A_0, \ldots, A_m$ enumerate $S_1(A)$ and let $B_0, \ldots, B_n$ enumerate $S_1(\neg B)$. Define sequences X_k, Y_k inductively, beginning with X_0, Y_0, as follows. Suppose we have defined X_k, Y_k.

If $k \leq m$, set

$$X_{k+1} = \begin{cases} X_k \cup \{A_k\} & \text{if } X_k \cup \{A_k\} \text{ and } Y_k \text{ are inseparable,} \\ X_k \cup \{\neg A_k\} & \text{otherwise.} \end{cases}$$

If $k > m$, set $X_{k+1} = X_k$.

If $k \leq n$, set

$$Y_{k+1} = \begin{cases} Y_k \cup \{B_k\} & \text{if } Y_k \cup \{B_k\} \text{ and } X_{k+1} \text{ are inseparable,} \\ Y_k \cup \{\neg B_k\} & \text{otherwise.} \end{cases}$$

If $k > n$, set $Y_{k+1} = Y_k$.

Finally, define $X = \bigcup_k X_k = X_{m+1}, Y = \bigcup_k Y_k = Y_{n+1}$. We claim that X, Y is an S_1-complete pair. The completeness clauses 2 and 3 of the definition of S_1-completeness (4.12) hold trivially. To establish inseparability, we prove the inseparability of X_k, Y_k by induction on k. Suppose X_k, Y_k to be inseparable, but X_{k+1}, Y_k not to be inseparable. This means that $X_k \cup \{A_k\}$ and Y_k and also $X_k \cup \{\neg A_k\}$ and Y_k are separable. Thus, for some $C_1, C_2 \in \mathcal{L}(A) \cap \mathcal{L}(B)$, we have

$$L \vdash \bigwedge\!\!\!\bigwedge X_k \wedge A_k \rightarrow C_1,$$
$$L \vdash \bigwedge\!\!\!\bigwedge Y_k \rightarrow \neg C_1,$$

and

$$L \vdash \bigwedge\!\!\!\bigwedge X_k \wedge \neg A_k \rightarrow C_2,$$
$$L \vdash \bigwedge\!\!\!\bigwedge Y_k \rightarrow \neg C_2,$$

hence

$$L \vdash \bigwedge\!\!\!\bigwedge X_k \rightarrow (C_1 \vee C_2),$$
$$L \vdash \bigwedge\!\!\!\bigwedge Y_k \rightarrow \neg(C_1 \vee C_2),$$

contradicting the inseparability of X_k, Y_k. The inseparability of X_{k+1}, Y_{k+1} follows from that of X_{k+1}, Y_k in a similar manner. $\Box$

Now we construct a countermodel to $A \rightarrow B$ from S_1-complete pairs. Greek letters $\alpha_0, \alpha_1, \ldots$ will denote such pairs, with $X_{\alpha_i}, Y_{\alpha_i}$ denoting the components of a pair $\alpha_i = (X_{\alpha_i}, Y_{\alpha_i})$. Fix $\underline{\alpha} = (X_{\underline{\alpha}}, Y_{\underline{\alpha}})$ to be a particular S_1-complete extension of the inseparable pair, $\{A\}, \{\neg B\}$.

Definition 4.14 Let $\langle W, R \rangle$ be the frame defined as follows: $\alpha_i R \alpha_j$ if and only if

1. for all $\varphi \in S(A) \cup S(\neg B)$, if $\Box\varphi \in X_{\alpha_i} \cup Y_{\alpha_i}$, then $\varphi \in X_{\alpha_j} \cup Y_{\alpha_j}$;
2. for all $\varphi \in S(A) \cup S(\neg B)$ and for all $i < n$, if $\Box^{i+1}\varphi \in X_{\alpha_i} \cup Y_{\alpha_i}$, then $\Box^i\varphi \in X_{\alpha_j} \cup Y_{\alpha_j}$;
3. there is $\varphi \in S(A) \cup S(\neg B)$ such that $\Box^{n-1}\varphi \in X_{\alpha_j} \cup Y_{\alpha_j}$ and $\Box\varphi \notin X_{\alpha_i} \cup Y_{\alpha_i}$.

Let $W = \{\alpha_i \mid \text{there is } n \in \mathbb{N} \text{ with } \underline{\alpha} R^n \alpha_i\}$.

The set W is finite because W consists of subsets of a fixed finite set.

Lemma 4.15 *The accessibility relation R is weakly transitive, irreflexive, and acyclic.*

Proof Observe that if φ is a formula such that $\varphi \in S(A)$ and $\Box^k\varphi \in X_{\alpha_i} \cup Y_{\alpha_i}$ (for some k with $1 \leq k \leq 2n$), then $\Box^k\varphi \in X_{\alpha_i}$. For suppose that $\Box^k\varphi \notin X_{\alpha_i}$. Since $\varphi \in S(A)$ we have that $\Box^k\varphi \in S_1(A)$, hence, by the S_1-completeness of X_{α_i}, $\neg\Box^k\varphi \in X_{\alpha_i}$. From $\Box^k\varphi \in X_{\alpha_i} \cup Y_{\alpha_i}$ and $\Box^k\varphi \notin X_{\alpha_i}$ it follows that $\Box^k\varphi \in Y_{\alpha_i}$. From $\varphi \in S(A)$ it follows $\Box^k\varphi \in \mathcal{L}(A)$ and from $\Box^k\varphi \in Y_{\alpha_i}$ it follows $\Box^k\varphi \in \mathcal{L}(B)$. From $\neg\Box^k\varphi \in X_{\alpha_i}$ and $\Box^k\varphi \in Y_{\alpha_i}$ it follows that $L \vdash \bigwedge\!\!\!\bigwedge X_{\alpha_i} \to \neg\Box^k\varphi$ and $L \vdash \bigwedge\!\!\!\bigwedge Y_{\alpha_i} \to \Box^k\varphi$. Since $\Box^k\varphi \in \mathcal{L}(A) \cap \mathcal{L}(B)$ the last two implications yield the separability of X_{α_i} and Y_{α_i}, a contradiction. In a similar manner we prove that if φ is a formula such that $\varphi \in S(\neg B)$ and $\Box^k\varphi \in X_{\alpha_i} \cup Y_{\alpha_i}$ (for some k with $1 \leq k \leq 2n$), then $\Box^k\varphi \in Y_{\alpha_i}$.

We now prove that R is weakly transitive. Suppose that $\alpha_0 R^n \alpha_n$. We must prove that $\alpha_0 R \alpha_n$. We first prove condition 1. Let $\varphi \in S(A) \cup S(\neg B)$ be such that $\Box\varphi \in \alpha_0$. Therefore $\Box\varphi \in X_{\alpha_0} \cup Y_{\alpha_0}$. Suppose that $\varphi \in S(A)$ (if $\varphi \in S(\neg B)$ the argument is similar). Then $\Box\varphi \in X_{\alpha_0}$, $\Box^n\varphi \in S_1(A)$ and, since $L \vdash \Box\varphi \to \Box^n\varphi$ and $L + \bigwedge\!\!\!\bigwedge X_{\alpha_0}$ is consistent, $\Box^n\varphi \in X_{\alpha_0}$ (as X_{α_0} is complete). Thus $\Box^n\varphi \in X_{\alpha_0} \cup Y_{\alpha_0}, \Box^{n-1}\varphi \in X_{\alpha_1} \cup Y_{\alpha_1}, \ldots, \varphi \in X_{\alpha_n} \cup Y_{\alpha_n}$, that is, $\varphi \in \alpha_n$.

We now prove condition 2. Let $\varphi \in S(A) \cup S(\neg B)$ such that $\Box^{i+1}\varphi \in X_{\alpha_0} \cup Y_{\alpha_0}$. Suppose that $\varphi \in S(A)$ (if $\varphi \in S(\neg B)$ the argument is similar). Therefore $\Box^{i+1}\varphi \in X_{\alpha_0}$ and $\Box^{n+i}\varphi \in S_1(A)$. Since $L \vdash \Box^{i+1}\varphi \to \Box^{n+i}\varphi$, $\Box^{n+i}\varphi \in X_{\alpha_0}$. It follows that $\Box^{n+i}\varphi \in X_{\alpha_0} \cup Y_{\alpha_0}$, $\Box^{n+i-1}\varphi \in X_{\alpha_1} \cup Y_{\alpha_1}, \ldots, \Box^{n+i-n}\varphi \in X_{\alpha_n} \cup Y_{\alpha_n}$, that is, $\Box^i\varphi \in \alpha_n$.

We finally prove condition 3. From $\alpha_0 R \alpha_1$ it follows that there is $\varphi \in S(A) \cup S(\neg B)$ such that $\Box^{n-1}\varphi \in X_{\alpha_1} \cup Y_{\alpha_1}$ and $\Box\varphi \notin X_{\alpha_0} \cup Y_{\alpha_0}$. Suppose that $\varphi \in S(A)$ (if $\varphi \in S(\neg B)$ the argument is similar). Therefore $\Box^{n-1}\varphi \in X_{\alpha_1}$. Thus, from $\Box^{2n-2}\varphi \in S_1(A)$ and from $L \vdash \Box^{n-1}\varphi \to \Box^{2n-2}\varphi$ it follows that $\Box^{2n-2}\varphi \in X_{\alpha_1}$. Thus, $\Box^{2n-2}\varphi \in X_{\alpha_1} \cup Y_{\alpha_1}, \Box^{2n-3}\varphi \in X_{\alpha_2} \cup Y_{\alpha_2}, \ldots, \Box^{2n-1-n}\varphi \in X_{\alpha_n} \cup Y_{\alpha_n}$, that is, $\Box^{n-1}\varphi \in X_{\alpha_n} \cup Y_{\alpha_n}$.

The proof that there are no reflexive nodes and no cycles is just a repetition of the proof given in Lemma 4.5. $\Box$

Definition 4.16 Define a model $\langle W, R, \Vdash\rangle$ based on $\langle W, R\rangle$ as follows: $\alpha_i \Vdash p$ if and only if $p \in X_{\alpha_i} \cup Y_{\alpha_i}$, for all α_i and for every atom p.

Lemma 4.17 *For $\langle W, R, \Vdash\rangle$ we have for all $\alpha \in W$ and for all $D \in S(A) \cup S(\neg B)$,*

$$(*) \qquad \alpha \Vdash D \text{ iff } D \in X_\alpha \cup Y_\alpha.$$

Proof We prove $(*)$ by induction on the complexity of D. If D is an atom then the result holds by definition of $\Vdash$.

Boolean cases are routine. For example we treat the case $D \equiv E \wedge F \in S(A) \cup S(\neg B)$. Without loss of generality, we assume $D \in S(A)$ (hence $E, F \in S(A)$). Suppose that $\alpha \Vdash D$, we want to prove that $D \in X_\alpha \cup Y_\alpha$. From $\alpha \Vdash D$ it follows that $\alpha \Vdash E$ and $\alpha \Vdash F$ and then, by the induction hypothesis, $E \in X_\alpha \cup Y_\alpha$ and $F \in X_\alpha \cup Y_\alpha$. We show that $E \in X_\alpha$. For suppose that $E \notin X_\alpha$, thus, since $E \in S(A)$, $\neg E \in X_\alpha \subseteq \mathcal{L}(A)$. From $E \in X_\alpha \cup Y_\alpha$ and $E \notin X_\alpha$ it follows that $E \in Y_\alpha \subseteq \mathcal{L}(B)$. Therefore we have $L \vdash \bigwedge\!\!\!\bigwedge X_\alpha \to \neg E$ and $L \vdash \bigwedge\!\!\!\bigwedge Y_\alpha \to E$, contradicting the inseparability of X_α and Y_α.

For the converse, suppose that $D \in X_\alpha \cup Y_\alpha$, and let us prove $\alpha \Vdash D$. Suppose that $E \wedge F \in S(A)$. Then $E \in S(A)$ and $F \in S(A)$. We show that $D \in X_\alpha$. If $D \notin X_\alpha$ then $\neg D \in X_\alpha$. Moreover $D \in Y_\alpha \subseteq \mathcal{L}(B)$. Thus $L \vdash \bigwedge\!\!\!\bigwedge X_\alpha \to \neg D$ and $L \vdash \bigwedge\!\!\!\bigwedge Y_\alpha \to D$, contradicting the inseparability of X_α and Y_α. Thus $E, F \in X_\alpha$. By the induction hypothesis, $\alpha \Vdash E$, $\alpha \Vdash F$ and, finally, $\alpha \Vdash E \wedge F$. The other Boolean cases are similar, therefore they are left to the reader.

Consider the case $D \equiv \Box E$. Let $\Box E \in S(A) \cup S(\neg B)$. Suppose that $\Box E \in X_\alpha \cup Y_\alpha$ and let us prove that $\alpha \Vdash \Box E$. From $\Box E \in X_\alpha \cup Y_\alpha$ it follows, by definition of R, that for every β such that $\alpha R \beta$, $E \in X_\beta \cup Y_\beta$. Therefore, by the induction hypothesis, for every β such that $\alpha R \beta$ we have $\beta \Vdash E$. Thus $\alpha \Vdash \Box E$.

For the converse, suppose by contradiction that $\alpha \Vdash \Box E$ and $\Box E \notin X_\alpha \cup Y_\alpha$. Define

$$
\begin{aligned}
\overline{X}_\beta &= \{\psi | \psi \in S(A), \Box\psi \in X_\alpha\} \cup \\
&\quad \{\Box^i \psi | \psi \in S(A), i < n, \Box^{i+1}\psi \in X_\alpha\} \text{ if } E \notin \mathcal{L}(A); \\
\overline{X}_\beta &= \{\psi | \psi \in S(A), \Box\psi \in X_\alpha\} \cup \{\Box^i \psi | \psi \in S(A), i < n, \Box^{i+1}\psi \in X_\alpha\} \cup \\
&\quad \{\neg E, \Box^{n-1} E\} \text{ if } E \in \mathcal{L}(A); \\
\overline{Y}_\beta &= \{\psi | \psi \in S(\neg B), \Box\psi \in Y_\alpha\} \cup \\
&\quad \{\Box^i \psi | \psi \in S(\neg B), i < n, \Box^{i+1}\psi \in Y_\alpha\} \text{ if } E \notin \mathcal{L}(B); \\
\overline{Y}_\beta &= \{\psi | \psi \in S(\neg B), \Box\psi \in Y_\alpha\} \cup \{\Box^i \psi | \psi \in S(\neg B), i < n, \Box^{i+1}\psi \in Y_\alpha\} \cup \\
&\quad \{\neg E, \Box^{n-1} E\} \text{ if } E \in \mathcal{L}(B).
\end{aligned}
$$

We prove that $\overline{X}_\beta$ and $\overline{Y}_\beta$ are inseparable. Suppose, by way of contradiction, that there is some $C \in \mathcal{L}(A) \cap \mathcal{L}(B)$ such that

1. $L \vdash \bigwedge\!\!\!\bigwedge \overline{X}_\beta \to C$,
2. $L \vdash \bigwedge\!\!\!\bigwedge \overline{Y}\beta \to \neg C$.

Case 1 Let $E \in \mathcal{L}(A) \setminus \mathcal{L}(B)$. Then from (1) it follows that

$$
\begin{aligned}
L \vdash\ & \textstyle\bigwedge_{\psi \in S(A), \Box\psi \in X_\alpha} \psi \wedge \\
& \textstyle\bigwedge_{\psi \in S(A), i<n, \Box^{i+1}\psi \in X_\alpha} \Box^i \psi \to (\neg C \to (\Box^{n-1} E \to E)).
\end{aligned}
$$

By MN we obtain

$$
\begin{aligned}
L \vdash\ & \textstyle\bigwedge_{\psi \in S(A), \Box\psi \in X_\alpha} \Box\psi \wedge \\
& \textstyle\bigwedge_{\psi \in S(A), i<n, \Box^{i+1}\psi \in X_\alpha} \Box^{i+1} \psi \to (\Box\neg C \to \Box(\Box^{n-1} E \to E)).
\end{aligned}
$$

Since $L \vdash \Box(\Box^{n-1} E \to E) \to \Box E$, we have

$$
\begin{aligned}
L \vdash\ & \textstyle\bigwedge_{\psi \in S(A), \Box\psi \in X_\alpha} \Box\psi \wedge \\
& \textstyle\bigwedge_{\psi \in S(A), i<n, \Box^{i+1}\psi \in X_\alpha} \Box^{i+1} \psi \to (\Box\neg C \to \Box E).
\end{aligned}
$$

It follows that

$$L \vdash \bigwedge\!\!\!\bigwedge X_\alpha \to (\Box\neg C \to \Box E).$$

On the other hand, since $\Box E \notin \mathcal{L}(B)$, $\Box E \in \mathcal{L}(A)$, $\Box E \in S(A)$, $\Box E \notin X_\alpha \cup Y_\alpha$, and X_α is S_1-complete, we have that $\neg\Box E \in X_\alpha$. Thus

$$L \vdash \bigwedge\!\!\!\bigwedge X_\alpha \to \neg\Box\neg C.$$

Since $\Box E \notin \mathcal{L}(B)$, from (2) it follows that

$$L \vdash \bigwedge_{\psi \in S(\neg B), \Box\psi \in Y_\alpha} \psi \wedge \bigwedge_{\psi \in S(\neg B), i<n, \Box^{i+1}\psi \in Y_\alpha} \Box^i \psi \to \neg C.$$

By MN we obtain

$$L \vdash \bigwedge_{\psi \in S(\neg B), \Box\psi \in Y_\alpha} \Box\psi \wedge \bigwedge_{\psi \in S(\neg B), i<n, \Box^{i+1}\psi \in Y_\alpha} \Box^{i+1} \psi \to \Box\neg C.$$

Thus
$$L \vdash \bigwedge\!\!\!\bigwedge Y_\alpha \to \Box\neg C.$$

Therefore
$$L \vdash \bigwedge\!\!\!\bigwedge X_\alpha \to \neg\Box\neg C,$$
$$L \vdash \bigwedge\!\!\!\bigwedge Y_\alpha \to \Box\neg C,$$

contradicting the inseparability of X_α, Y_α.

Case 2 $E \in \mathcal{L}(B) \setminus \mathcal{L}(A)$. This case is treated similarly.

Case 3 Let $E \in \mathcal{L}(A) \cap \mathcal{L}(B)$. From (1) and (2) it follows that

$$\begin{aligned} L \ \vdash \ & \bigwedge_{\psi \in S(A), \Box\psi \in X_\alpha} \psi \wedge \\ & \bigwedge_{\psi \in S(A), i<n, \Box^{i+1}\psi \in X_\alpha} \Box^i \psi \to ((\neg E \wedge \Box^{n-1} E) \to C); \end{aligned}$$

$$\begin{aligned} L \ \vdash \ & \bigwedge_{\psi \in S(\neg B), \Box\psi \in Y_\alpha} \psi \wedge \\ & \bigwedge_{\psi \in S(\neg B), i<n, \Box^{i+1}\psi \in Y_\alpha} \Box^i \psi \to ((\neg E \wedge \Box^{n-1} E) \to \neg C). \end{aligned}$$

Hence, by MN,

$$\begin{aligned} L \ \vdash \ & \bigwedge_{\psi \in S(A), \Box\psi \in X_\alpha} \Box\psi \wedge \\ & \bigwedge_{\psi \in S(A), i<n, \Box^{i+1}\psi \in X_\alpha} \Box^{i+1} \psi \to \Box((\neg E \wedge \Box^{n-1} E) \to C); \end{aligned}$$

$$\begin{aligned} L \ \vdash \ & \bigwedge_{\psi \in S(\neg B), \Box\psi \in Y_\alpha} \Box\psi \wedge \\ & \bigwedge_{\psi \in S(\neg B), i<n, \Box^{i+1}\psi \in Y_\alpha} \Box^{i+1} \psi \to \Box((\neg E \wedge \Box^{n-1} E) \to \neg C). \end{aligned}$$

Thus

$$L \ \vdash \ \bigwedge\!\!\!\bigwedge X_\alpha \to \Box((\neg E \wedge \Box^{n-1} E) \to C);$$

$$L \ \vdash \ \bigwedge\!\!\!\bigwedge Y_\alpha \to \Box((\neg E \wedge \Box^{n-1} E) \to \neg C).$$

By hypothesis $\Box E \in S(A) \cup S(\neg B)$. Without loss of generality we can suppose $\Box E \in S(A)$. Since $\Box E \notin X_\alpha \cup Y_\alpha$, $\Box E \notin X_\alpha$. Therefore $\neg\Box E \in X_\alpha$.

We prove that $L \vdash \bigwedge\!\!\!\bigwedge X_\alpha \to \neg\Box((\neg E \wedge \Box^{n-1} E) \to \neg C)$. Observe

$$L \vdash ((\Box^{n-1} E \wedge \neg E. \to \neg C) \wedge (\Box^{n-1} E \wedge \neg E. \to C)) \to (\Box^{n-1} E \to E).$$

Thus, by MN,

$$L \vdash (\Box(\Box^{n-1} E \wedge \neg E. \to \neg C) \wedge \Box(\Box^{n-1} E \wedge \neg E. \to C)) \to \Box(\Box^{n-1} E \to E).$$

Therefore, since $L \vdash \Box(\Box^{n-1} E \to E) \to \Box E$,

$$L \vdash (\Box(\Box^{n-1} E \wedge \neg E. \to \neg C) \wedge \Box(\Box^{n-1} E \wedge \neg E. \to C)) \to \Box E.$$

On the other hand, we have

$$L \ \vdash \ \bigwedge\!\!\!\bigwedge X_\alpha \to \Box(\neg E \wedge \Box^{n-1} E. \to C),$$

$$L \ \vdash \ \bigwedge\!\!\!\bigwedge X_\alpha \to \neg\Box E.$$

Hence

$$L \vdash \bigwedge\!\!\!\bigwedge X_\alpha \to \neg\Box((\neg E \wedge \Box^{n-1} E) \to \neg C).$$

We have thus established that

$$L \vdash \bigwedge\!\!\!\bigwedge X_\alpha \to \neg\Box((\neg E \wedge \Box^{n-1} E) \to \neg C),$$

$$L \vdash \bigwedge\!\!\!\bigwedge Y_\alpha \to \Box((\neg E \wedge \Box^{n-1} E) \to \neg C).$$

Since $E \in \mathcal{L}(A) \cap \mathcal{L}(B)$ and $C \in \mathcal{L}(A) \cap \mathcal{L}(B)$,

$$\Box((\neg E \wedge \Box^{n-1} E) \to \neg C) \in \mathcal{L}(A) \cap \mathcal{L}(B),$$

contrary the inseparability of X_α e Y_α.

In a similar manner we treat the case $\Box E \in S(\neg B)$. We have thus proved that $\overline{X}_\beta$ and $\overline{Y}_\beta$ are inseparable. From Lemma 4.13 it follows that there exist X_β, Y_β S_1-complete such that $\overline{X}_\beta \subseteq X_\beta$ and $\overline{Y}_\beta \subseteq Y_\beta$. We show that $\alpha R \beta$. Let $\psi \in S(A) \cup S(\neg B)$ be such that $\Box\psi \in X_\alpha \cup Y_\alpha$. If $\Box\psi \in X_\alpha$ then, by definition, $\psi \in \overline{X}_\beta \subseteq X_\beta$. If $\Box\psi \in Y_\alpha$ then, by definition, $\psi \in \overline{Y}_\beta \subseteq Y_\beta$. At any rate $\psi \in X_\beta \cup Y_\beta$. This proves condition 1. Now, let $\psi \in S(A) \cup S(\neg B)$ with $i < n$ be such that $\Box^{i+1}\psi \in X_\alpha \cup Y_\alpha$. If $\Box^{i+1}\psi \in X_\alpha$ then, by definition, $\Box^i\psi \in \overline{X}_\beta \subseteq X_\beta$. If $\Box^{i+1}\psi \in Y_\alpha$ then $\Box^i\psi \in \overline{Y}_\beta \subseteq Y_\beta$. At any rate $\Box^i\psi \in X_\beta \cup Y_\beta$. This proves condition 2. Finally, we prove condition 3. By hypothesis $\Box E \in S(A) \cup S(\neg B)$ (hence $E \in S(A) \cup S(\neg B)$) and $\Box E \notin X_\alpha \cup Y_\alpha$. From $\Box E \in S(A) \cup S(\neg B)$ it follows that $E \in \mathcal{L}(A)$ or $E \in \mathcal{L}(B)$. If $E \in \mathcal{L}(A)$ then, by definition, $\Box^{n-1}E \in \overline{X}_\beta \subseteq X_\beta$. If $E \in \mathcal{L}(B)$ then $\Box^{n-1}E \in \overline{Y}_\beta \subseteq Y_\beta$. At any rate, $\Box^{n-1}E \in X_\beta \cup Y_\beta$. We have thus shown that if $\Box E \notin X_\alpha \cup Y_\alpha$ then there is some β, $\alpha R \beta$, such that $E \notin X_\beta \cup Y_\beta$. If $E \notin X_\beta \cup Y_\beta$ then, by the induction hypothesis, $\beta \nVdash E$ and therefore, since $\alpha R \beta$, $\alpha \nVdash \Box E$. $\Box$

We can now prove Craig's interpolation property for L.

Theorem 4.18 *If $L \vdash A \to B$ then there is a formula C possessing only atoms common to A and B and such that $L \vdash A \to C$ and $L \vdash C \to B$.*

Proof Suppose no interpolant exists between A and B. Then $\langle W, R, \Vdash\rangle$, as just defined, is a model for L, and there is $\underline{\alpha} \in W$ such that $\underline{\alpha} \nVdash A \to B$, a contradiction. $\Box$

We conclude this section with two open problems: (1) Is there a provability predicate $Pr(\cdot)$ for PA whose provability logic is $K + \Box(\Box^{n-1}p \to p) \to \Box p$? (2) Is there a constructive proof of the fixed point theorem for $K + \Box(\Box^{n-1}p \to p) \to \Box p$?

References

[1] Boolos, G., *The Unprovability of Consistency*, Cambridge University Press, Cambridge, 1979. Zbl 0409.03009. MR 81c:03013. 66, 71, 74

[2] Boolos, G., *The Logic of Provability*, Cambridge University Press, Cambridge, 1993. Zbl 0891.03004. MR 95c:03038. 66

[3] Feferman, S., "The arithmetization of metamathematics in a general setting," *Fundamentum Mathematicae*, vol. 49 (1960), pp. 35–92. 66

[4] Gabbay, D. M., "Craig's interpolation theorem for modal logics," pp. 111–27 in *Conference in Mathematical Logic—London '70*, edited by W. Hodges, Springer, Berlin, 1972. Zbl 0233.02009. MR 49:8815. 73

[5] Hughes, G. E., and M. J. Cresswell, *A New Introduction to Modal Logic*, Routledge, London, 1996. Zbl 0855.03002. MR 99j:03011. 69

[6] Maksimova, L., "Amalgamation and interpolation in normal modal logic," *Studia Logica*, vol. 50 (1991), pp. 457–71. Zbl 0754.03013. MR 93e:03030. 68, 73

[7] Rautenberg, W., "Modal tableau calculi and interpolation," *Journal of Philosophical Logic*, vol. 12 (1983), pp. 403–23. Zbl 0547.03015. MR 85b:03029. 73

[8] Sacchetti, L., "Modal logics with the fixed-point property," *Bollettino della Unione Matematica Italiana*, vol. 2 (1999), pp. 279–90. MR 2000f:03060. 68, 73

[9] Sambin, G., and S. Valentini, "The modal logic of provability. The sequential approach," *Journal of Philosophical Logic*, vol. 11 (1982), pp. 311–42. Zbl 0523.03014. MR 84b:03033. 66

[10] Smoryński, C., "Beth's theorem and self-referential sentences," pp. 253–61 in *Logic Colloquium '77*, edited by A. Macintyre, L. Pacholski, and J. Paris, North-Holland, Amsterdam, 1978. Zbl 0453.03018. MR 80c:03023. 66

[11] Smoryński, C., *Self-reference and Modal Logic*, Springer-Verlag, New York, 1985. Zbl 0596.03001. MR 88d:03001. 66, 68, 70, 74

[12] van Benthem, J., *Modal Logic and Classical Logic*, Bibliopolis, Naples, 1985. Zbl 0639.03014. MR 88k:03029. 76

Acknowledgments

I am indebted to Professor Franco Montagna for his cooperation and for a number of helpful suggestions.

Dipartimento di Scienze dell'Informazione
via Comelico 39–41
20135 Milano
ITALY
sacchetti@dsi.unimi.it