Rosalie Iemhoff
Dick de Jongh
Chunlai Zhou

Properties of Intuitionistic Provability and Preservativity Logics

Abstract. We study the modal properties of intuitionistic modal logics that belong to the provability logic or the preservativity logic of Heyting Arithmetic. We describe the $\Box$-fragment of some preservativity logics and we present fixed point theorems for the logics iL and iPL, and show that they imply the Beth property. These results imply that the fixed point theorem and the Beth property hold for both the provability and preservativity logic of Heyting Arithmetic. We present a frame correspondence result for the preservativity principle Wp that is related to an extension of Löb's principle.

Keywords: Intuitionistic modal logic, provability logic, preservativity logic, Heyting Arithmetic, Beth definability, fixed points.

1. Introduction

In this paper[1] we study some intuitionistic modal logics that arise from a specific mathematical interpretation of the modal operations. The modalities we consider are $\Box$ and $\rhd$, and their interpretation is given by

$$\begin{array}{ll} \Box\varphi & \text{“}\varphi \text{ is provable in } \mathsf{HA}\text{”, i.e. } \mathsf{HA} \vdash \varphi \\ \varphi \rhd \psi & \text{“for all } \sigma \in \Sigma_1\text{: } \mathsf{HA} \vdash \sigma \to \varphi \text{ implies } \mathsf{HA} \vdash \sigma \to \psi\text{”}, \end{array}$$

where HA is Heyting Arithmetic, the constructive counterpart of PA, i.e. it is a theory in intuitionistic predicate logic IQC that has as axioms the non-logical axioms of PA, and Σ_1 is the first level of the arithmetical hierarchy. All the logics we consider are part of the provability or preservativity logic of HA. This means that all these logics consist of propositional schemes that HA proves about the provability predicate $\Box_{\mathsf{HA}}$ or the preservativity predicate $\rhd_{\mathsf{HA}}$ of HA. In particular, the theorems of these logics are (constructively) valid schemes. Note that provability logic is part of preservativity logic, as

$$\mathsf{HA} \vdash \Box_{\mathsf{HA}}\varphi \equiv \top \rhd_{\mathsf{HA}} \varphi.$$

Preservativity logic was introduced by Visser[2002] as a constructive alternative for interpretability logic, to which it is equivalent for many classical

[1] A shorter version of which was published as Iemhoff et al.[2004].

theories, in particular for PA. No axiomatization is known for the preservativity logic of HA, but over the last few years at least part of the logic has been axiomatized[2]. It is a logic in the language of preservativity logic, $L_{\rhd}$, i.e. the language of propositional logic extended with one binary modal operator $\rhd$. $L_{\Box}$ is the language of provability logic, i.e. the language of propositional logic extended with one modal operator $\Box$. As mentioned above, in preservativity logic we can define $\Box A$ as $\top \rhd A$. In this paper we consider the following principles of the preservativity logic of HA ($\rhd$ and $\Box$ bind stronger than $\wedge, \vee$, that bind stronger than $\to$).

IPC	intuitionistic propositional logic		
$P1$	$A \rhd B \wedge B \rhd C \to A \rhd C$		
$P2$	$A \rhd B \wedge A \rhd C \to A \rhd (B \wedge C)$		
Dp	$A \rhd B \to (A \vee C) \rhd (B \vee C)$		
Mp	$A \rhd B \to (\Box C \to A) \rhd (\Box C \to B)$		
Wp	$A \wedge \Box B \rhd B \to A \rhd B$		
		K	$\Box(A \to B) \to (\Box A \to \Box B)$
$4p$	$A \rhd \Box A$	4	$\Box A \to \Box\Box A$
Lp	$(\Box A \to A) \rhd A$	L	$\Box(\Box A \to A) \to \Box A$
		Le	$\Box(A \vee B) \to \Box(A \vee \Box B)$
Rules:			
$Pres$	$A \to B \ / \ A \rhd B$	Nec	$A \ / \ \Box A$
MP	$A \ \ (A \to B) \ / \ B$		

iP^- denotes the logic given by IPC, the principles $P1, P2$, and the rules $Pres$ and MP. iP is the logic iP^- extended by Dp and is called the *basic preservativity logic* for reasons explained in the next section. By $iP4$ we denote the logic iP extended by the principle $4p$. Similarly for the other preservativity principles Lp, Mp, Wp. iK denotes the logic given by IPC, K, and the rules Nec and MP. The logic iK extended by the principle 4 is denoted by $iK4$. Similarly for L, Le. Conform tradition, iKL is denoted by iL. $iLLe$ denotes iL extended by Le. iPX denotes an arbitrary extension of iP. Lemma 1.1 below shows that all provability principles can be derived from the preservativity principles.

Readers familiar with provability logic will note that at the right side the non-logical axioms $K, 4, L$ of the provability logic GL of PA are listed. This in contrast to Dp and Le, that do not belong to the preservativity logic

[2] Visser[2002] Iemhoff[2003]

of PA. Since it is difficult to think of a natural classical interpretation of $\Box$ or $\rhd$ for which these principles would hold, they are not likely to appear in classical modal logic. On the classical side, the modal study of the provability logic of PA has been a useful tool, as it was shown in Solovay [1976] that for a formula $A(p_1, \ldots, p_n)$ not derivable in GL one can construct, on the basis of a countermodel of A, arithmetical formulas φ_i such that $\mathsf{PA} \nvdash \mathsf{A}(\varphi_1, \ldots, \varphi_n)$, (where $\Box$ in A are interpreted as the provability predicate of PA). An analogous result holds for the interpretability (preservativity) logic of PA. Although it is open whether similar results hold for HA, the modal study of the principles above is interesting for two main reasons. First, these principles express principles of HA. Therefore, knowledge about them is likely to provide insights in HA, and might help in the search for a complete axiomatization of the provability and preservativity logic of HA. In fact, modal results have lead to new principles of these logics before. Second, as mentioned above, some of these principles do not belong to the logics regularly studied in intuitionistic modal logic. We think that therefore the modal study of these principles might be a valuable addition to the field.

In Iemhoff[2003] modal completeness results were presented for all logics given by some or all of these principles, except for iPW and iPL. In this paper we continue the modal study of these logics by investigating the relation between the preservativity and provability logics (Section 3), and by presenting fixed point theorems for both iPL and iL (Section 4). From the latter it follows that both the fixed point theorem and the Beth property hold for any extension of these logics in the appropriate language. In particular, it follows that they hold for the provability and preservativity logic of HA. The proof of the fixed point theorem for iPL also provides another proof of the fixed point theorem for the interpretability logic IL. Furthermore, we present a correspondence theorem for iPW and explain the connection of this principle to iPL. It is not difficult to show that Wp is derivable in $iPLM$, so in view of the preservativity logic of HA it does not add anything new. However, this principle came up in Zhou [2003] in relation to the open problem of frame completeness for iPL, where it plays an interesting role. This will be explained in more detail in Section 3.4 on Wp.

$\Box$-fragments

The first part (Section 3) on the relation between preservativity and provability logics, needs a little more explanation. The $\Box$-fragment of a preservativity logic iPX in $L_{\rhd}$ is defined to be

$$iPX_{\Box} := \{A \text{ in } L_{\Box} \mid iPX \vdash A\}.$$

Here we ask ourselves what the $\Box$-fragment of a given preservativity logic is. An obvious relation between $\Box$ and $\rhd$ is given by the following lemma[3].

LEMMA 1.1. *$iP^- \vdash \Box(A \to B) \to A \rhd B$ and $iP^- \vdash A \rhd B \to (\Box A \to \Box B)$.*

Now the guiding idea behind the description of the $\Box$-fragments is the translation $^\circ$ on formulas that inductively replaces all occurrences of $A \rhd B$ by $\Box A \to \Box B$. All preservativity principles except Dp, Mp and Wp are derivable in iL under this translation[4]. For Wp, it is explained in Section 3.4. why its translation under $^\circ$ does not belong to the provability logic of HA. For Dp and Mp, the translation of which under $^\circ$ does not belong to the provability logic of HA either[5], it turns out that there are rules that somehow cover the effect of Dp and Mp on the $\Box$-fragment of the preservativity logics that contain them. These are the rules

$$\begin{array}{ll} DR & \Box A \to \Box B \ / \ \Box(A \vee C) \to \Box(B \vee C) \\ MoR & \Box A \to \Box B \ / \ \Box(\Box C \to A) \to \Box(\Box C \to B). \end{array}$$

We show that for all preservativity logics considered in this paper, these rules determine the $\Box$-fragment of a preservativity logic in the following way.

THEOREM 1.2. (Numbers indicate the sections where the equality is proved.)

$$\begin{array}{lllll} iP_{\Box} & \overset{3.3}{=} & iK & \overset{3.3}{=} & iK + DR \\ iP4_{\Box} & \overset{3.1}{=} & iLe & \overset{3.1.1}{=} & iK4 + DR \\ iPL_{\Box} & \overset{3.2}{=} & iLLe & \overset{3.2.1}{=} & iL + DR \\ iPM_{\Box} & \overset{3.3}{=} & iK & \overset{3.3}{=} & iK + DR + MoR \\ iPW_{\Box} & \overset{3.4}{=} & iLLe & \overset{3.4}{=} & iL + DR \end{array}$$

In particular, if X is one of $4p, Lp$ or empty, then $iPX_{\Box} = iKX^\circ + DR$. For $X = Mp$, $iPX_{\Box} = iK + DR + MoR = iK$. Wp is an exception of this regularity, as $iPW_{\Box} \neq iKW^\circ + DR$ (Section 3.4).

[3] Iemhoff[2003]
[4] Iemhoff[2003]
[5] Iemhoff[2001]

In all these cases the general method to prove these equalities is similar. As an example, we explain the way in which the equalities on the second line $iP4_\Box = iLe = iK4 + DR$ are proved. First it is shown that $iP4_\Box = iLe$, essentially by proving completeness of iLe and $iP4$ with respect to the same class of frames. As $iLe \subseteq iK4 + DR$, this gives

$$iP4_\Box = iLe \subseteq iK4 + DR.$$

It remains to prove that

$$iK4 + DR \subseteq iP4_\Box.$$

Since by Lemma 1.1 it follows that $iP4_\Box \vdash 4$, it suffices to show that DR is admissible for $iP4_\Box$. This is shown via proving that the rule BP (Box Pres)

$$\Box A \to \Box B \ / \ A \rhd B,$$

is admissible for iPX, i.e. $iPX + BP = iPX$, and then applying the following lemma.

LEMMA 1.3. *If the rule BP is admissible for iPX, then DR is admissible for iPX, and whence for $iPX_\Box$. If in addition $iPX \vdash Mp$, then both DR and MoR are admissible for iPX, and whence for $iPX_\Box$.*

PROOF. We show the first part of the lemma, as the second part is similar. Suppose $iPX \vdash \Box A \to \Box B$. By the admissibility of BP this implies that $iPX \vdash A \rhd B$. Since iP derives Dp, so does iPX. Whence $iPX \vdash (A \vee B) \rhd (B \vee C)$. Lemma 1.1 gives $iPX \vdash \Box(A \vee C) \to \Box(B \vee C)$. ∎

All the other equalities are proved in a similar manner. However, the applied techniques differ with the logic. In some cases modal completeness results are used, but for iPL, for which such a result is not available, other methods had to be found.

2. Preliminaries and Tools

2.1. Semantics for Preservativity Logic

DEFINITION 2.1. *A frame F* (for preservativity logic) is a triple $\langle W, R, \leq \rangle$, where W is a nonempty set of possible worlds, points or nodes, $\leq$ is a partial order and R is a binary relation satisfying

$$(\leq \circ R) \subseteq R.$$

A model M is a quadruple $\langle W, R, \leq, \Vdash \rangle$ where $\langle W, R, \leq \rangle$ is a frame and $\Vdash$ is a forcing relation between points in W and propositional letters which satisfies the following condition:

- *(persistence)* If $x \Vdash p$ and $x \leq y$, then $y \Vdash p$.

Next we model the whole language by extending the forcing relation $\Vdash$ to relate points to complex formulae by interpreting the connectives in IPC in the usual manner:

$M, w \Vdash A \wedge B \equiv_{def} M, w \Vdash A$ and $M, w \Vdash B$;
$M, w \Vdash A \vee B \equiv_{def} M, w \Vdash A$ or $M, w \Vdash B$;
$M, w \Vdash A \to B \equiv_{def} \forall v \geq w (M, v \Vdash A$ implies $M, v \Vdash B)$;
$M, w \Vdash \top$ for any w;
$M, w \nVdash \bot$ for any w.

Define $\neg A$ as $A \to \bot$. From the above third and fifth clauses and the definition of $\neg A$ as $A \to \bot$, it is easy to deduce that $M, w \Vdash \neg A$ iff $\forall v \geq w (M, v \nVdash A)$. The most important and characteristic clause is the following one for $\rhd$ formulas:

$$M, w \Vdash A \rhd B \equiv_{def} \text{ for any } v \text{ such that } wRv, \text{ if } M, v \Vdash A, \text{ then } M, v \Vdash B.$$

It follows immediately that $M, w \Vdash \Box A$ iff for any v such that wRv, $M, v \Vdash A$. Also it is easy to check that

- (persistence for all formulas) for any formula A in $L_{\rhd}$, if $M, w \Vdash A$ and $w \leq v$, then $M, v \Vdash A$.

As a matter of fact, given the persistence for propositional letters, the condition that $(\leq \circ R) \subseteq R$ is a *necessary and sufficient* condition to guarantee persistence for all formulas[6], which is different from the condition $(\leq \circ R) \subseteq (R \circ \leq)$ for intuitionistic modal logic (sometimes we write $R \circ \leq$ as $\bar{R}$).

LEMMA 2.2. *Let $\langle W, R, \leq \rangle$ be a frame such that W is a non-empty set, $\leq$ is a partial order and $(\leq \circ R) \nsubseteq R$. Then there is a formula A of $L_{\rhd}$ and a forcing relation $\Vdash$ such that in $\langle W, R, \leq, \Vdash \rangle$ for some $x, y \in W$, $x \leq y$ and $x \Vdash A$ but $y \nVdash A$.*

[6] Zhou[2003]

PROOF. Since not $(\leq \circ R) \subseteq R$, there are two worlds $x, y \in W$ such that $x(\leq \circ R)y$ but not xRy. That is to say, there is a world $u \in W$ such that $x \leq uRy$. Define $z \Vdash p$ iff $y \leq z$; $z \Vdash q$ iff $z \not\leq y$; $z \nVdash r$ for any other propositional letter r.
It is easy to check that $\Vdash$ is a forcing relation, i.e. it satisfies that for any propositional letter t, if $x \leq y$ and $x \Vdash t$, then $y \Vdash t$. On one hand, $u \nVdash p \rhd q$. For $uRy, y \Vdash p$, $y \nVdash q$ and hence $u \nVdash p \rhd q$. On the other hand, $x \Vdash p \rhd q$. To see this we shall show that if xRz and $z \Vdash p$, then $z \Vdash q$. Since xRz, $z \neq y$. Moreover, $y \leq z$, for $z \Vdash p$. Therefore $y < z$. This implies that $z \not\leq y$ and hence $z \Vdash q$. So we get that $x \leq u$, $x \Vdash p \rhd q$ but $u \nVdash p \rhd q$.

■

A is *valid in a model* M if for any $w \in W$ $M, w \Vdash A$. A is *valid in a frame* F if A is valid on any model $M = \langle F, \Vdash \rangle$ on the frame.

THEOREM 2.3. $iP \vdash A$ iff A is valid on all frames iff A is valid on all finite frames. [7]

As we can easily see, iP stands in the same position in preservativity logic as iK does in normal modal logics. In this sense, it is the *basic preservativity logic*.

2.2. Semantics for Intuitionistic Provability Logic

The semantics for $L_\Box$ should be part of the semantics for $L_\rhd$ because we define $\Box A$ to be $\top \rhd A$. Although, given the persistence for propositional letters, the condition that $(\leq \circ R) \subseteq (R \circ \leq)$ is a sufficient and necessary condition that guarantee the persistence for all formulas in $L_\Box$ (in fact $(\leq \circ R) \subseteq R$ implies $(\leq \circ R) \subseteq (R \circ \leq)$), it is well justified[8] to define frames for intuitionistic provability logics in the following simpler way, which is the same as that of frames for preservativity logics.

DEFINITION 2.4. *A frame* F (for intuitionistic provability logic) is a triple $\langle W, R, \leq \rangle$ where W is a nonempty set of possible worlds, points or nodes, $\leq$ is a partial order and R is a binary relation satisfying

$$(\leq \circ R) \subseteq R.$$

[7] Proposition 4.1.1 in Iemhoff [2001].

[8] For details, see Bozic and Došen [1983] or Zhou [2003].

However, for $L_\Box$ we can impose extra conditions on frames that we cannot require of frames for preservativity logic. All intuitionistic modal logics iT that we will consider below are complete with respect to some class of frames satisfying additionally:

- (*brilliancy*) $(R\circ \leq) \subseteq R$.

In particular, iK is complete w.r.t the class of finite brilliant frames. Moreover, all the notions and propositions above can be adapted into intuitionistic modal logic automatically. We will not go into details about that.

2.3. Some Useful Facts

In the following we achieve some basic propositions in preservativity logics that will be very useful to other sections in this paper. First we establish the connection between the natural rule for preservativity logic: preservation rule and the more-often-used rule: necessitation rule.

THEOREM 2.5. In any preservativity logic iT containing all theorems in iP^-, the preservation rule and the necessitation rule are equivalent.[9]

PROOF. Assume that $A \to B/A \rhd B$ is admissible in iT and $iT \vdash A$. Then $iT \vdash \top \to A$ because $IPC \vdash A \to (\top \to A)$. By the preservation rule, we get that $iT \vdash \top \rhd A$, i.e $iT \vdash \Box A$. Now for the other direction. Assume that the necessitation rule is admissible in iT and $iT \vdash A \to B$. By applying the necessitation rule, we get that $iT \vdash \Box(A \to B)$. It follows from Lemma 1.1 that $iT \vdash A \rhd B$.

∎

By Lemma 1.1, we immediately get the following two corollaries:

COROLLARY 2.6. *The following two forms of the Mp principle are equivalent over iP^-:*

1. $A \rhd B \to (\Box C \to A) \rhd (\Box C \to B)$
2. $(A \wedge \Box C) \rhd B \to A \rhd (\Box C \to B)$.

COROLLARY 2.7. *The following two forms of W are equivalent over iP^-:*

1. $A \wedge \Box B \rhd B \to A \rhd B$

[9] This theorem is, in fact, implicit in Lemma 3.1.2 in Iemhoff [2001].

2. $A \rhd B \to (\Box B \to A) \rhd B$

We will use the following substitution lemmas in our section on the fixed point theorems.

LEMMA 2.8. (a) $T \vdash \boxdot(A \leftrightarrow B) \to (F[A/p] \leftrightarrow F[B/p])$, for $T = iP4$ or $T = iK4$,
(b) If p occurs only *modalized* in F, i.e. only under $\Box$ or $\rhd$, then
$T \vdash \Box(A \leftrightarrow B) \to (F[A/p] \leftrightarrow F[B/p])$, for $T = iP4$ or $T = iK4$.

PROOF. We can prove (a) directly by induction on the complexity of F, and (b) by induction from (a).

■

The following lemma, the proof of which we leave to the reader, shows that the principle 4 is a very basic principle in intuitionistic provability logic.

LEMMA 2.9. *For $T = iL$ or $T = iLe$, $T \vdash \Box A \to \Box\Box A$ for any formula A in $L_\Box$.*

3. Conservation Results

As you will see, the rule $\Box A \to \Box B / A \rhd B$ plays a dominant role in the following sections. This rule is discussed in Section 5.2 of Iemhoff [2001] where a short proof sketch is given for the admissibility of the rule for iPH. We will give detailed proofs of the admissibility of this rule in many other logics, which may impose an impression that we repeat the same proofs. This is not the case though. Every time we show the admissibility for a different logic, you will find some additional new ideas in the proof.
We will divide the presentation of this section into several parts according to the results which we have previously mentioned.

3.1. Conservation of $iP4$ over iLe

A notational convention: Given a frame M, $[z) := \{w|$ there is a sequence of $w_0 S_0 w_1 \cdots w_n = w$ for some worlds $w_0, w_1, \cdots, w_n$ in M where $S_i \in \{R, \leq\}\}$. Thus $[z)$ stands for the subframe generated by z. The same notation applies to models.

THEOREM 3.1. [10]

[10] Propostions 4.2.1, 4.2.2, 4.4.1 in Iemhoff [2001].

1. In $L_\Box$, 4 corresponds to semi-transitivity: $(R \circ R) \subseteq (R \circ \leq)$.
2. In $L_\Box$, $\vdash_{iK4} A$ iff A is valid on all finite transitive frames.
3. The principle $4p$ corresponds to gatheringness: if $wRvRu$, then $v \leq u$.
4. $\vdash_{iP4} A$ iff A is valid on all finite gathering frames.
5. On finite frames Le corresponds to the Le-property: $\forall wv(wRv \to \exists x(wRx \leq v \wedge \forall u(vRu \to x \leq u)))$.
6. $\vdash_{iLe} A$ iff A is valid on all finite brilliant Le-frames.
7. In $L_\Box$, $\vdash_{iLLe} A$ iff A is valid on all finite transitive conversely well-founded brilliant Le-frames.

For the sake of completeness, we will repeat the conservation of $iP4$ over iLe in Iemhoff [2001] to give a more transparent presentation. Besides, we need again the procedure (used in Lemma 3.3) transforming Le-frames to gathering frames again in the proofs of Lemmas 3.11 and 3.34.

LEMMA 3.2. *Let $M := \langle W, R, \leq, \Vdash \rangle$ and $N := \langle W, R', \leq, \Vdash \rangle$ be two finite models. If $R' \subseteq R \subseteq (R' \circ \leq)$, then $M, w \Vdash B$ iff $N, w \Vdash B$ for any formula B in $L_\Box$ and any world $w \in W$.*

LEMMA 3.3. *Let $M := \langle W, R, \leq, \Vdash \rangle$ be a finite Le brilliant model. Then there is a finite gathering model $N = \langle W, R', \leq, \Vdash \rangle$ such that $R' \subseteq R \subseteq (R' \circ \leq)$.*

PROOF. Assume that $M := \langle W, R, \leq, \Vdash \rangle$ is a finite Le brilliant model. Define:

$$wR'v \equiv_{def} wRv \text{ and } \forall u(vRu \to v \leq u) \text{ and } N := \langle W, R', \leq, \Vdash \rangle.$$

$S(x)$ denotes the property: $\forall u(xRu \to x \leq u)$. Assume that wRv. We need to find an x such that $wR'x \leq v$. That is to say, $wRx \leq v$ and $S(x)$. By the Le-property, there is a successor x_1 of w which is below both v and all its own successors. If $x_1 = v$, then we have found such an x. If $x_1 \neq v$, then there is another successor x_2 of w which is below both x_1 and all its own successors. If $x_2 = x_1$, then we have found such an x. If not, we will repeat the same argument as above. Thus, we will get a sequence $x_1 x_2 \cdots$. Since the frame is finite, there are two nodes $x_{n-1} = x_n$ for some n; x_n is the x that we are looking for.

■

LEMMA 3.4. $\vdash_{iP4} \Box(A \vee B) \to \Box(A \vee \Box B)$.

PROOF. Observe that, by $4p$ and Dp, $\vdash_{iP4} (A \vee B) \rhd (A \vee \Box B)$, and apply Lemma 1.1.

■

THEOREM 3.5. $\vdash_{iLe} A$ iff A is valid on all finite gathering frames.

PROOF. The right-to-left direction follows from the fact that Le is derivable in $iP4$ (Lemma 3.4). We just need to show the other direction. Suppose that $\nvdash_{iLe} A$. Then by the completeness of iLe, we know that there is a world b in some finite brilliant Le model $M = \langle W, R, \leq, \Vdash \rangle$ such that $M, b \nVdash A$. According to Lemma 3.3, there is another new finite gathering model $N = \langle W, R', \leq, \Vdash \rangle$ such that $R' \subseteq R \subseteq (R' \circ \leq)$. From Lemma 3.2, it follows that $N, b \nVdash A$.

■

COROLLARY 3.6. (*Conservation*) $\vdash_{iP4} A$ *iff* $\vdash_{iLe} A$*, for all* A *in* $L_{\Box}$*.*

3.1.1. iLe is equivalent to the logic $iK4$ with DR

LEMMA 3.7. *Let* M *be a model on a gathering frame and* x, y *be two worlds in this model such that* xRy*. If* $y \Vdash A$*, then, for any* $z \in [y)$*,* $z \Vdash \boxdot A$*.*

PROOF. First one observation: for any $z \in [y)$, $y \leq z$. This follows immediately from the fact that M is on a gathering frame. So $z \Vdash A$. Take any successor w of z, $w \Vdash A$ because $w \in [y)$. So $z \Vdash \Box A$ and hence $z \Vdash \boxdot A$.

■

LEMMA 3.8. $\vdash_{iP4} A \rhd B$ *iff* $\vdash_{iP4} (\Box A \to \Box B)$*.*

PROOF. The direction from left to right follows from Lemma 1.1. We prove the other direction by contraposition. Suppose that $iP4 \nvdash A \rhd B$. From the completeness of $iP4$, it follows that $A \rhd B$ is false at a point w of some finite gathering model M. Then there is a point v such that wRv, $v \Vdash A$ and $v \nVdash B$. Now we define from the original one a new model M', which is, in fact, a submodel of the old one. Take $W' := \{w\} \cup [v)$, $R' = R\restriction_{W'}$, $\leq' = \leq\restriction_{W'}$, and $x \Vdash p$ iff $x \Vdash' p$ for any propositional variable p, for all $x \in W'$. Observe that M' has a gathering frame. Note that, for any $x \in [v)$ and for any formula B in $L_{\rhd}$, $M', x \Vdash B$ iff $M, x \Vdash B$.

It is clear that $M', w \nVdash \Box B$ because $wR'v$ and $M', v \nVdash B$. By the above lemma, we get that $M', w \Vdash \Box A$ because $R'[w] \subseteq [v)$ and for any $x \in [v), x \Vdash A$. So $M', w \Vdash \Box A$ but $M', w \nVdash \Box B$, which implies that $M', w \nVdash \Box A \to \Box B$. Therefore $\nvdash_{iP4} \Box A \to \Box B$.

■

THEOREM 3.9. *iLe is equivalent to the logic $iK4$ with the extra rule DR. Whence $iP4_\Box = iLe = iK4 + DR$.*

PROOF. First we prove that iLe is contained in $iK4 + DR$. We only need to show that Le is derivable in the latter logic. Since $iK4 + DR \vdash \Box A \to \Box\Box A$, we can get Le immediately by just applying DR. For the other direction, recall[11] that the principle 4 is derivable in iLe. Whence it remains to show that DR is admissible for iLe, which is the same as showing that it is admissible for $iP4_\Box$, by Corollary 3.6. That DR is admissible for $iP4_\Box$ follows from the previous lemma, by applying Lemma 1.3. ■

3.2. Conservation of iPL over $iLLe$

LEMMA 3.10. *The principle Lp corresponds to gatheringness plus converse well-foundedness of the modal relation. Similarly, L corresponds to semi-transitivity plus well-foundedness.*[12]

LEMMA 3.11. *$iLLe \vdash A$ iff A is valid on all finite gathering conversely well-founded frames.*

PROOF. First the easier left-to-right direction. It suffices to show that both Le and L are valid on all finite gathering conversely well-founded frames. Firstly, Le is valid on all finite gathering frames and hence on all finite gathering conversely well-founded frames. Secondly, L is valid on all finite gathering conversely well-founded frames. For L corresponds to semi-transitivity plus converse well-foundedness, and gatheringness implies semi-transitivity.

Next we show the more difficult direction. Suppose that $iLLe \nvdash A$ where A is a formula in $L_\Box$. According to Theorem 3.1, there is a point b in some model $M = \langle W, \leq, R, \Vdash\rangle$ which is finite transitive conversely well-founded brilliant Le-model such that $M, b \nVdash A$. Define $wR'v$ iff wRv and $\forall u(vRu \to v \leq u)$. By the same argument as that in Lemma 3.3, we get that $R' \subseteq R \subseteq (R' \circ \leq)$. Set $M' = \langle W, R', \leq, \Vdash\rangle$. It is easy to see that M' is on a finite gathering frame, as we impose this property through the definition of R'.

Finally, M' is conversely well-founded. Suppose not. Then there is a loop: $w_0 R' w_1 R' \cdots R' w_n R' w_0$. According to the definition of R', $w_0 R w_1 R \cdots R w_n R w_0$, which is impossible because R is conversely well-founded. So M' is on a finite gathering well-founded frame. It follows immediately from

[11] Lemma 2.9

[12] Iemhoff [2001].

the above Lemma 3.2 that $M', b \not\Vdash A$. Since M' is on a finite gathering conversely well-founded frame, A is not valid on all finite gathering conversely well-founded frames.

■

THEOREM 3.12. (Conservation) $iLLe$ is the $L_{\Box}$-fragment of iPL.

PROOF. Suppose that $iLLe \nvdash A$. Then A is not valid on all gathering conversely well founded frames. It follows from the above correspondence result that $iPL \nvdash A$. On the other hand, it is easy to see that $iLLe$ is contained in iPL. For both L and Le are derivable in iPL.

■

3.2.1. $iLLe$ is Equivalent to iL with the Extra Rule DR

In the following paragraphs we are mainly concerned with the proof of

$$iPL \vdash \Box A \to \Box B \Leftrightarrow iPL \vdash A \rhd B,$$

which immediately implies that $iLLe$ is equivalent to iL with the extra rule DR.

LEMMA 3.13. $iP4 \vdash \Box((\Box C \to C) \rhd C) \to (\Box C \to C) \rhd C$.

PROOF. First note that $iP \vdash \Box(A \rhd B) \to (\Box A \rhd \Box B)$ by Lemma 1.1. Reason inside $iP4$:

$\Box((\Box C \to C) \rhd C) \to \Box(\Box C \to C) \rhd \Box C$
$\Box((\Box C \to C) \rhd C) \to (\Box C \to C) \rhd \Box C$
$\Box((\Box C \to C) \rhd C) \to (\Box C \to C) \rhd (\Box C \wedge (\Box C \to C))$
$\Box((\Box C \to C) \rhd C) \to (\Box C \to C) \rhd C.$

■

COROLLARY 3.14. $iP4 \vdash \Box \boxdot L \leftrightarrow \boxdot L \leftrightarrow \Box L$ *where* L *is* $(\Box C \to C) \rhd C$.

LEMMA 3.15. (*Detour Lemma*) $iPL \vdash A$ *iff there exist* $C_1, C_2, \cdots, C_n$ *such that* $iP4 \vdash \Box((\Box C_1 \to C_1) \rhd C_1) \wedge \cdots \wedge \Box((\Box C_n \to C_n) \rhd C_n) \to A$.

PROOF. Let C range over expressions $((\Box C_1 \to C_1) \rhd C_1) \wedge \cdots \wedge (\Box C_n \to C_n) \rhd C_n))$. Then the above proposition can be put in the following simpler way:

$$iPL \vdash A \text{ iff there exists } C \text{ such that } iP4 \vdash \Box C \to A.$$

The direction from right to left is obvious. We just show the other direction. Assume that $iPL \vdash A$. Let $s_1 s_2 \cdots s_n$ be a proof of A in iPL, i.e. $s_n = A$ and for all $i \leq n$, s_i is an axiom of iPL, or there are $j, h < i$ such that $s_j\ s_h\ /\ s_i$ is an instance of Modus Ponens, or $s_j\ /\ s_h$ is an instance of the rule $Pres$. Let $(\Box C_1 \to C_1) \rhd C_1), \cdots, ((\Box C_n \to C_n) \rhd C_n)$ be the instances of the Löb principle occurring in the sequence and C denote their conjunction. Define $s'_i := \Box C \to s_i$. With induction to i we show that for all i, $iP4 \vdash s'_i$. This proves that $iP4 \vdash \Box C \to A$.

1. If s_i is an instance of axiom of iP, then s'_i is a theorem of iP and hence of $iP4$ because, in fact, $s_i \to (\Box C \to s_i)$ is a tautology;
2. If s_i is an instance of Lp, it is easy to see that s'_i is a theorem of $iP4$ by of the following reasoning:

   $iP4 \vdash \Box C \to \boxdot C$

   $\Rightarrow iP4 \vdash \boxdot C \to s_i$

   $\Rightarrow iP4 \vdash \Box C \to s_i$

   $\Rightarrow iP4 \vdash s'_i$;
3. If there are s_j and $s_k (j, k < i)$ such that $s_j \equiv s_k \to s_i$, then $s'_j \equiv \Box C \to (s_k \to s_i) \equiv ((\Box C \to s_k) \to (\Box C \to s_i)) \equiv s'_k \to s'_i$, and as $iP4 \vdash s'_j \wedge s'_k$ by IH, $iP4 \vdash s'_i$ follows;
4. If there are B, D and $s_j (j < i)$ such that $s_j \equiv B \to D$ and $s_i \equiv B \rhd D$, then $s'_j \equiv \Box C \to (B \to D)$ and $s'_i \equiv \Box C \to (B \rhd D)$. Now we show that we can get s'_i in $iP4$. The following is the argument:

   $iP4 \vdash s'_j$ by IH

   $\Rightarrow iP4 \vdash \Box C \to (B \to D)$

   $\Rightarrow iP4 \vdash \Box\Box C \to \Box(B \to D)$

   $\Rightarrow iP4 \vdash \Box C \to \Box(B \to D)$

   $\Rightarrow iP4 \vdash \Box C \to (B \rhd D)$

■

LEMMA 3.16. *If* $iP4 \vdash \Box C \to (\Box A \to \Box B)$, *then* $iP4 \vdash \Box C \to (A \rhd B)$, *for all formulas* C.

PROOF. The proof is similar to the one of Lemma 3.8.

■

THEOREM 3.17. $iPL \vdash \Box A \to \Box B$ iff $iPL \vdash A \rhd B$

PROOF. Assume that $iPL \vdash \Box A \to \Box B$. Then, according to Lemma 3.15, there are some instances of L: $(\Box C_1 \to C_1) \rhd C_1, \cdots, (\Box C_n \to C_n) \rhd C_n$ (C denotes their conjunction) such that $iP4 \vdash \Box C \to (\Box A \to \Box B)$. By Lemma 3.16, we get that $iP4 \vdash \Box C \to (A \rhd B)$. This implies, according to Lemma 3.15, that $iPL \vdash A \rhd B$.

■

COROLLARY 3.18. *$iLLe$ is equivalent to the logic iL with the extra rule DR. Whence $iPL_{\Box} = iLLe = iL + DR$.*

PROOF. The previous theorem, Theorem 3.12 and Lemma 1.3. ■

In fact, we can show the admissibility of DR in $iLLe$ without that of $\Box A \to \Box B / A \rhd B$ in iPL. The proof strategy here is similar to the above though. First we give a similar detour lemma:

LEMMA 3.19. *For any formula A in $L_{\Box}$, $\vdash_{iLLe} A$ iff $\vdash_{iLe} \Box C \to A$ where C is the conjunction of some instances of Löb's principle L.*

PROOF. Here we only mention that, for any instance C of Löb's provability principle, $\vdash_{iLLe} \Box C \leftrightarrow \boxdot C$.

■

THEOREM 3.20. DR is admissible in $iLLe$.

PROOF. $\vdash_{iLLe} \Box A \to \Box B$

$\Rightarrow \vdash_{iLe} \Box C \to (\Box A \to \Box B)$ for some conjunction C of instances of L.

$\Rightarrow \vdash_{iLe} \Box(C \wedge A) \to \Box B$

$\Rightarrow \vdash_{iLe} \Box((C \wedge A) \vee D) \to \Box(B \vee D)$ (by the admissibility of DR in iLe)

$\Rightarrow \vdash_{iLe} \Box(C \vee D) \to (\Box(A \vee D) \to \Box(B \vee D))$

$\Rightarrow \vdash_{iLe} \Box C \to (\Box(A \vee D) \to \Box(B \vee D))$

$\Rightarrow \vdash_{iLLe} \Box(A \vee D) \to \Box(B \vee D)$.

■

3.3. Conservation of iPM over iK

Before proving that iK is the $L_{\Box}$-fragment of iPM (Theorem 3.29), we show the admissibility of $\Box A \to \Box B / A \rhd B$. Recall that $\bar{R}$ is short for $R\circ \leq$.

LEMMA 3.21. [13] *(i) The principle Mp corresponds to the Mp property:*
$\forall wvu(wRv \leq u \rightarrow \exists x(wRx \wedge v \leq x \leq u \wedge x\bar{R} \subseteq u\bar{R})$.
(ii) $\vdash_{iPM} A$ *iff* A *is valid on all finite Mp frames.*

It is not easy to give a precise proof of the admissibility of rules in iPM, although the intuitive idea is not difficult. Probably the reason for that lies in the fact that the Mp property is a property that states the existence of certain nodes. Hence, in contrast to the situation for e.g. gathering models, the Mp property is not inherited by submodels. In order to make certain submodels into models that have the Mp property, we have to define some notions to help with our formalization of the proof.

DEFINITION 3.22. A triple (w, v, u) in a frame is called a *problem* if it satisfies $wRv \leq u$. It is called an *unsolved problem* if it additionally satisfies:

$$\text{there is no } x \text{ such that } wRx, v \leq x \leq u \text{ and } x\bar{R} \subseteq u\bar{R}.$$

Such an x is called a *solution* to the above problem. If such an x exists, then the problem is called *a solved problem.* Let (w, v, u) and (w, v', u) be two problems. If $v \leq v'$, then we denote $(w, v, u) \preceq (w, v', u)$ and say (w, v, u) is *below* (w, v', u). If $v < v'$, then we denote $(w, v, u) \prec (w, v', u)$. A problem (w, v, u) is called *a dispensable problem* if there is another different problem such that $(w, v, u) \prec (w, v', u)$. A problem is called *indispensable* if it is not dispensable. $<[x] := \{z|x < z\}$. A point w is called a *minimal point* in X if there is no point v such that $v < w$. $min[x]$ denotes the set of minimal points in the set $<[x]$.

LEMMA 3.23. *The following propositions follows immediately from the above definition:*

1. *If* $x \leq y$, *then* $y\bar{R} \subseteq x\bar{R}$. *Therefore, if* $x \leq y$ *and* $x\bar{R} \subseteq y\bar{R}$, *then* $y\bar{R} = x\bar{R}$.
2. *Let* (w, v, u) *and* (w, v', u) *be two problems. If* $(w, v, u) \preceq (w, v', u)$ *and* x *is a solution to the problem* (w, v', u), *then* x *is also a solution to the problem* (w, v, u).
3. *Let* $v \leq v'$. *For all* x, y, *if both* (x, v, y) *and* (x, v', y) *are problems, then each solution to* (x, v', y) *is also a solution to* (x, v, y).
4. *Any problem* (w, v, v) *is a solved problem.*

[13](i) Iemhoff [2001] (ii) Zhou [2003] .

5. *Let F be a frame. If there is no unsolved problem, then F satisfies the Mp-property.*

THEOREM 3.24. $iPM \vdash A \rhd B$ iff $iPM \vdash (\Box A \to \Box B)$.

PROOF. One direction follows from Lemma 1.1. We show the other direction. Assume that $iPM \nvdash A \rhd B$. Then, according to the completeness of iPM, $A \rhd B$ is falsified at some point w_0 of some model M on some finite frame satisfying the Mp-property. This implies that there is a v_0 such that $w_0 R v_0, M, v_0 \Vdash A$ but $M, v_0 \nVdash B$. In the following we will construct a new Mp-model M' such that $M' \nVdash (\Box A \to \Box B)$. Note that, if a problem (x, y, z) is inside $[v_0)$, i.e $x, y, z \in [v_0)$, then it is easy to see that there is a solution to this problem in $[v_0)$. This means that we don't need to consider the problems in $[v_0)$ because all of them are already solved in $[v_0)$.

First we define $W_0 := [v_0) \cup \{w\}, R_0 := R \restriction_{[v_0)} \cup \{(w, v_0)\}; \leq_0 := \leq \restriction_{[v_0)}$ where w is a new world. Then enumerate all the elements in $min[v_0]$: $u_0, u_1, \cdots, u_n$. The purpose of choosing minimal points is to just to make the construction more efficient. At every stage we will add solution to some indispensable unsolved problems. These solutions will be denoted by x_σ, where σ is a sequence of nodes in W_0, used to keep track of the problem to which x_σ is a solution. Let $*$ denote concatenation of sequences, and let σ_l denote the last element of σ, i.e. $\langle y_0, \ldots, y_m \rangle_l = y_m$, and $\tau \preceq \sigma$ denote that τ is an initial segment of σ. Let $TC(S)$ denote the reflexive transitive closure of a relation S. At stage 0, we add for every unsolved problem (w, v_0, u_i) (which is indispensable because the u_i are minimal elements), a new world $x_{\langle u_i \rangle}$ to W_0. We define the new frame $F_1 = \langle W_1, R_1, \leq_1 \rangle$ via $W_1 := W_0 \cup \{x_{\langle u_i \rangle} \mid i \leq n\}; R_1 := R_0 \cup \{(w, x_{\langle u_i \rangle}), (x_{\langle u_i \rangle}, z) \mid i \leq n, (u_i, z) \in R_0\}; \leq_1 := TC(\leq_0 \cup \{(v_0, x_{\langle u_i \rangle}), (x_{\langle u_i \rangle}, u_i) | i \leq n\})$. It is easy to see that $x_{\langle u_i \rangle} \bar{R} = u_i \bar{R}$ and that we can extend the forcing relation to the new nodes by defining $x_{\langle u_i \rangle} \Vdash p$ iff $u_i \Vdash p$ for all propositional letters. Note that in this way persistency is satisfied.

Observe that in F_1 all indispensable unsolved problems are of the form $(w, x_{\langle u_i \rangle}, z)$, for some $z \in< [u_i]$. Namely, all the problems (w, v_0, u_i) and $(w, x_{\langle u_i \rangle}, u_i)$ have been solved by $x_{\langle u_i \rangle}$, and all problems (w, v_0, z) for $z \in< [u_i]$ have become dispensable in F_1 through $x_{\langle u_i \rangle}$. Therefore, we have only to consider problems $(w, x_{\langle u_i \rangle}, z)$, for some $z \in< [u_i]$. At stage 1, we add for every unsolved problem $(w, x_{\langle u_i \rangle}, u)$ with $u \in min[u_i]$, a solution $x_{\langle u_i, u \rangle}$ to W_1, and proceed in the same way as before.

In general, at stage $i+1$, we consider the nodes x_σ in F_i that are newly added at stage i. For every such x_σ and every problem (w, x_σ, u) for $u \in min[\sigma_l]$ we add a solution $x_{\sigma * u}$ to W_i. We define the new frame $F_{i+1} = \langle W_{i+1}, R_{i+1}, \leq_{i+1}\rangle$ via
$W_{i+1} := W_i \cup \{x_{\langle \sigma * u\rangle} \mid x_\sigma \in W_i \backslash W_{i-1}, u \in min[\sigma_l]\}$;
$R_{i+1} := R_i \cup \{(w, x_{\sigma * u}), (x_{\sigma * u}, z) \mid x_\sigma \in W_i \backslash W_{i-1}, u \in min[\sigma_l], (u, z) \in R_0\}$;
$\leq_{i+1} := TC(\leq_i \cup \{(x_\sigma, x_{\sigma * u}), (x_{\sigma * u}, u) | x_\sigma \in W_i \backslash W_{i-1}, u \in min[\sigma_l]\})$. Again, note that $x_{\sigma * u}\bar{R} = u\bar{R}$ and that by extending the forcing relation to the new nodes via $x_{\sigma * u} \Vdash p$ iff $u \Vdash p$, persistency is satisfied.
To see that the procedure terminates, observe that for every x_σ that is added at some stage in the construction, σ is of the form $\langle u_i, y_1, \ldots, y_m\rangle$, for some nodes $u_i < y_1 < \ldots < y_m$ in W_0. Since W_0 is finite, termination follows. Let $i+1$ be a stage at which there are no more unsolved indispensable problems. Consider M_i. As explained in the previous lemma, M_i satisfies the Mp-property. It remains to show that $M_i, w \Vdash \Box A$ and $M_i, w \nVdash \Box B$, as according to the soundness of iPM, this gives $\nvdash_{iPM} \Box A \to \Box B$.
It is not difficult to prove with induction to i that for all formulas C,

1. for all $x_{\sigma * u} \in W_i$, $M_i, u \Vdash C$ iff $M_i, x_{\sigma * u} \Vdash C$.
2. for all nodes $y \in [v_0)$, $M_i, y \Vdash C$ iff $M, y \Vdash C$.

We leave the proof to the reader. For the first part, use the fact that $x_{\sigma * u}\bar{R} = u\bar{R}$. For the last part, use the first part and the facts that no new nodes are added above the u_i, and that above v_0 all new nodes are of the form x_σ. As observed above, for $x_{\sigma * u}$, $u \in W_0$ and $v_0 \leq u$. Thus $u \Vdash A$, and whence $M_i, x_{\sigma * u} \Vdash A$. Whence $x_\sigma \Vdash A$, for all $x_\sigma \in W_i$. As for w, wR_iy implies $y = v_0$ or y is a new node x_σ, it follows that $M_i, w \Vdash \Box A$. $M_i, w \nVdash \Box B$ follows from the fact that $M_i, v_0 \nVdash B$. ■

We can extract a theorem from the above proof, which is very handy when we deal with the admissibility of many other rules in iPM.

THEOREM 3.25. Let $M = \langle W, R, \leq, V\rangle$ be a finite Mp model and $[v)$ be any generated submodel by v. Then there is a new finite Mp model $N = \langle W', R', \leq', V'\rangle$ in which

1. $[v) \subseteq W'$;
2. for any world $x \in [v)$ and any formula E, $M, x \Vdash E$ iff $N, x \Vdash E$;
3. there is a world $w \in W'$ such that $wR'v$ and, if $wR'y$ and $M, v \Vdash A$, then $N, y \Vdash A$.

4. if $M, v \Vdash A$ and $M, v \nVdash B$, then there is a world $w' \in W'$ such that $w'R'v$ and $N, w' \nVdash \Box A \to \Box B$.

By appealing to this theorem, we can easily show that iPM is also closed under the inference rules: $\Box A/A$ and $\Box A \to \Box B / \boxdot A \to B$.

THEOREM 3.26. The logic $iK + DR + MoR$ is contained in $iPM_{\Box}$.

PROOF. By Lemma 1.3 and Theorem 3.24.

■

There also is an interesting syntactic proof of $iK + DR + MoR = iPM_{\Box}$ that uses the following translation on formulas which is related to the translation $^\circ$ given in the introduction.

DEFINITION 3.27. The translation * from formulas in $L_{\rhd}$ to those in $L_{\Box}$ is inductively defined as follows:

- For p, $\top$ and $\bot$, $p^* = p$, $\top^* = \top$ and $\bot^* = \bot$.
- For $\circ \in \{\vee, \wedge, \to\}$, $(A \circ B)^* = A^* \circ B^*$.
- $(\neg A)^* = \neg A^*$.
- $(A \rhd B)^* = \Box(A^* \to B^*)$.

LEMMA 3.28. *If* $iK \vdash X^*$*, then* $iPX_{\Box} = iK$*, where* X *is in* $L_{\rhd}$*.*

PROOF. Clearly, $iK \subseteq iPX_{\Box}$. Thus it remains to show that $iPX_{\Box} \subseteq iK$. Assume that $iPX_{\Box} \vdash A$. Of course we can consider A as a formula in $L_{\rhd}$ according to the fact that $\Box A \equiv (\top \rhd A)$ in iP. It suffices to show that

$$\text{if } iPX \vdash A, \text{ then } iK \vdash A^* \ (*)$$

because, for any formula B in $L_{\Box}$, $B^* = B$.
Since $iPX \vdash A$, there is a finite sequence $s_1 s_2 \cdots s_n (= A)$ of formulas in $L_{\rhd}$ in which, for any $s_i (1 \leq i \leq n)$,

1. either s_i is in the forms of P_1, P_2, Dp or X,
2. or there are some $A_1, A_2, s_j \in L_{\Box}$ $(j < i)$ such that $s_i = A_1 \rhd A_2$ and $s_j = A_1 \to A_2$,
3. or there are some $s_j, s_k (j, k < i)$ such that $s_k = s_j \to s_i$.

The sequence $s_1^* s_2^* \cdots s_n^* (= A^*)$ of formulas in $L_\Box$ is a proof of A^* in iK. We treat the first case and leave the others to the reader. If s_i is an instance of P_1, P_2, Dp or X, then it is easy to see that s_i^* is a theorem of iK for the first three, and it follows by assumption for X.

■

THEOREM 3.29. $iPM_\Box = iK = iK + DR + MoR$.

THEOREM 3.30. $iP_\Box = iK$.

3.4. Conservation of iPW Over $iLLe$

The usual method to prove completeness for L, like the proof method in the proof of completeness for iL, breaks down for iPL. One of the problems is that it is not possible in iPL to infer $A \rhd B$ from $A \wedge \Box B \rhd B$ [14]. This is how the principle Wp emerged. Trivially, Lp is derivable in IPW. We do not know whether Wp is complete, but in the following we will give a correspondence result for Wp and show that the $\Box$-fragment of iPW is $iLLe$. Thus although iPW and iPL are distinct, their $\Box$-fragments are equal.

THEOREM 3.31. [15] Let F be a finite frame. $F \Vdash (A \wedge \Box B) \rhd B \to A \rhd B$ iff F satisfies the following property:

$$\forall wvu(wRvRu \to \exists x(wRx \wedge v < x \leq u)).$$

DEFINITION 3.32. Let $\langle W, R\rangle$ be a finite, transitive, gathering and conversely well-founded (hence irreflexive) frame. An end point w is a world without a w' such that wRw' or $w \leq w'$. It is easy to see that for any $w \in W$, there is a finite sequence s of $w_n, w_{n-1}, \cdots, w_0$ such that $w = w_n S_n w_{n-1} \cdots S_1 w_0$ where s_i is either R or $\leq$ and w_0 is an end point. We define the *grade* $g_s(w)$ of w in this sequence inductively as follows:

1. $g_s(w_0) := 0$
2. If $g_s(w_{i-1}) = k$ and S_i is $\leq$, then $g_s(w_i) := k$; If $g_s(w_{i-1}) = k$ and S_i is R, then $g_s(w_i) := k + 1$.

Of course, $g_s(w_i) \leq n$ for any $i \leq n$. For each $w \in W$, we define the *rank* $r(w)$ of w as the greatest such $g_s(w)$ (we omit the subscript $\langle W, R\rangle$ here). Note that, if wRv, then $r(w) > r(v)$.

[14] Page 68 in Iemhoff [2001].

[15] Lemma 3.5.1 in Zhou [2003].

THEOREM 3.33. $iPW \vdash A$ implies that A is valid on all finite gathering, transitive and conversely well-founded frames.

PROOF. It suffices to show that Wp is valid on all finite gathering, transitive, conversely well-founded frames. Given a model M on such a frame $\langle W, R\rangle$ and any $w', w, v \in W$ such that $w' \leq wRv$, assume that $M, w \Vdash (A \wedge \Box B) \rhd B$ and $M, v \Vdash A$. We need to show that $M, v \Vdash B$. It suffices to show that $M, v \Vdash \Box B$. Suppose that this is not the case: $M, v \nVdash \Box B$.

Now consider the v-generated submodel M'. Obviously, $M', v \nVdash \Box B$, $M', v \Vdash A$ and M' is on a finite transitive, gathering and conversely well-founded frame. Then there is a world $v' \in W'$ of least rank such that $M', v' \nVdash \Box B$. This implies that, for any $v'' \in W'$ such that $v'Rv''$, $M', v'' \Vdash \Box B$ and hence $M, v'' \Vdash \Box B$. Such a v'' can always be found because every end point makes all boxed formulas true. It is easy to check that wRv'' by transitivity and that $M', v'' \Vdash A$ (and hence $M, v'' \Vdash A$) according to the fact that $M', v \Vdash A$ and $v'' \in [v)$. Since $M, w \Vdash (A \wedge \Box B) \rhd B$, $M, v'' \Vdash B$ and hence $M', v'' \Vdash B$. So $M', v' \Vdash \Box B$. We have arrived at a contradiction. So $M, v \Vdash \Box B$ and hence $(A \wedge \Box B) \rhd B$ is valid on any finite gathering, transitive and conversely-well-founded frame.

■

The converse of this lemma is not true. On the one hand, it is easy to check that $(A \rhd B) \rightarrow \Box(A \rhd B)$ is valid on all transitive frames. On the other hand, it is well-known that this formula is not arithmetically valid in HA. Suppose that the converse were true. Then, according to the converse proposition, $iPW \vdash (A \rhd B) \rightarrow \Box(A \rhd B)$, which is impossible because Wp is a valid principle in HA whereas $(A \rhd B) \rightarrow \Box(A \rhd B)$ is not.

LEMMA 3.34. $\vdash_{iLLe} A$ *iff* A *is valid on all finite gathering* transitive *and conversely well-founded frames.*

PROOF. In fact the lemma is not new, just an extension of Lemma 3.11. We only need to show that transitivity is preserved in the new model $N = \langle W, R', \leq, V\rangle$. Assume that $w, v, u \in W$ and $wR'vR'u$. We need to show that $wR'u$. Since $wR'vR'u$, $wRvRu$ and hence wRu because R is transitive. It remains to show that, for any z such that $uR'z$, $u \leq z$. This immediately follows from the assumption that $vR'u$. So $wR'u$.

■

THEOREM 3.35. $iPW_{\Box} = iLLe = iL + DR$.

Proof. Since both L and Le are derivable in iPL, and iPW is a proper extension of iPL, $iLLe \subseteq iPW_\Box$ is clear. For $iPW_\Box \subseteq iLLe$, suppose that $\nvdash_{iLLe} A$ for some A in $L_\Box$. By the completeness of $iLLe$, we know that A is not valid on some finite transitive gathering and conversely well-founded frame. It follows from Lemma 3.33 that $\nvdash_{iPW} A$. This shows that $iPW_\Box = iLLe$. That $iLLe = iL + DR$ follows from Corollary 3.18.

■

Note that in contrast to the principles treated before, $iPW_\Box \neq iKW^\circ + DR$, as $(Wp)^\circ = (\Box(A \wedge \Box B) \to \Box B) \to (\Box A \to \Box B)$. Using the completeneness result for $iLLe$ one can show that $(Wp)^\circ$ does not belong to $iLLe$, but one can also show directly that Wp does not belong to the provability logic of HA, and whence cannot be derivable from $iLLe$. For if so, $(\Box\Box B \to \Box B) \to \Box B$ would belong to the provability logic as well, because it is derivable from Wp. But this principle is not even true, neither classically nor constructively, as it constructively implies $\neg\neg(\Box\Box B \vee \Box B)$.

4. Fixed Points and Beth Definability

In this section we will prove the fixed point theorems for iL and iPL and point out connections with Beth's Definability Theorem. Let us remind the reader that fixed point theorems are of the form: for each formula $A(p)$ in which p occurs only modalized, there exists a unique B not containing p such that B and $A(B)$ are provably equivalent. The proof of the existence of fixed points in iL is an adaptation of the well-known proof of that property for GL; the proof of the existence of fixed points in iPL derives from the one for IL, the basic interpretability logic (de Jongh-Visser [1991]). We will give the main steps of the proof but not all the details where these are sufficiently similar to the classical proofs. In the last subsection, we will discuss the interderivability between fixed points and Beth definability (Definition 4.21) in both intuitionistic provability and preservativity logics. This extends the work of Areces et al.[2000] (see also Hoogland[2001], Ch. 5).

A notational convention: AB is the result of substitution of B for p in the formula Ap.

THEOREM 4.1. (Uniqueness Theorem) Suppose that p occurs modalized in A, then $\vdash_L (\boxdot(p \leftrightarrow Ap) \wedge \boxdot(q \leftrightarrow Aq)) \to (p \leftrightarrow q)$ where $L \in \{iL, iPL\}$.[16]

[16] See Smoryński[1985]. The proof there is intuitionistically acceptable. This is also the case for Lemma 4.3, Corollary 4.4 and Theorem 4.5 below.

Proofs of the existence of fixed points for a system usually consist of proving the existence of fixed points for the basic formulas and proving an inductive step. For the inductive step for iPL, we can borrow the following theorem[17], since its proof did not use classical logic. This means that for iL and iPL we can confine ourselves to proving the basic cases.

THEOREM 4.2. Let U be any extension of iL or iPL satisfying:

FIX: Every formula Ap of the form $\Box Bp$ or $Bp \rhd Cp$ has a fixed point.

Then, for every formula Ap with p modalized, there is a formula J such that p does not occur in J and $\vdash_U J \leftrightarrow AJ$.

4.1. Fixed Point Theorem for iL

LEMMA 4.3. *$iL \vdash \Box A\top \leftrightarrow \Box A\Box A\top$ for all formulas A.*

COROLLARY 4.4. *Let $Ap := B\Box Cp$. Then $iL \vdash AB\top \leftrightarrow AAB\top$.*

Now an application of Theorem 4.2 suffices.

THEOREM 4.5. If in C the propositional letter p occurs exclusively under $\Box$, then there is a formula D not containing p such that $iL \vdash D \leftrightarrow CD$.

4.2. Fixed Point Theorem for iPL

The following proof is similar to the one for interpretability logic in de Jongh-Visser [1991]. To put it more precisely, the fixed point for the formula $A(p) \rhd B(p)$ in iPL is a kind of mirror image of that for the formula $A(p) \rhd_i B(p)$ in IL. This is not surprising since classically $A(p) \rhd B(p)$ is equivalent to $\neg B(p) \rhd_i \neg A(p)$ in IL. Since the latter formula contains negations however the details of the intuitionistic proof ought not to be skipped. The crucial point is Theorem 4.8 which reflects E2 of de Jongh-Visser [1991].

Define: $A \equiv B :\Leftrightarrow \vdash_{iPL} (A \rhd B) \wedge (B \rhd A)$.

LEMMA 4.6. *$A \equiv A \wedge \Box A \equiv \Box A \rightarrow A$.*

LEMMA 4.7. *If $\vdash \Box B\top \rightarrow C$, then $\vdash B\top \wedge \Box B\top \leftrightarrow BC \wedge \Box BC$.*

[17] Theorem 2.4 in de Jongh-Visser[1991].

PROOF. For the left-to-right direction, reason in iPL:

$\Box B\top \to C,$
$\Box B\top \to (\top \leftrightarrow C)$
$\Box\Box B\top \to \Box(\top \leftrightarrow C)$
$\Box B\top \to \boxdot(\top \leftrightarrow C)$
$\Box B\top \to (\boxdot B\top \leftrightarrow \boxdot BC)$
$\boxdot B\top \to \boxdot BC.$

Now the other direction. Again reason in iPL:

$\Box B\top \to \boxdot(\top \leftrightarrow C)$ $(*)$
$\Box(\Box B\top \to \boxdot(\top \leftrightarrow C))$
$BC \wedge \Box BC \wedge \Box(\Box B\top \to \boxdot(\top \leftrightarrow C)) \to (BC \wedge \Box(\Box B\top \to B\top))$
$BC \wedge \Box BC \wedge \Box(\Box B\top \to \boxdot(\top \leftrightarrow C)) \to (BC \wedge \Box B\top))$
$BC \wedge \Box BC \to (BC \wedge \Box B\top))$
$BC \wedge \Box BC \to \boxdot(C \leftrightarrow \top)$ (by $(*)$)
$(BC \wedge \Box BC) \to (B\top \wedge \Box B\top)$

■

THEOREM 4.8. If $\vdash \Box B\top \to C$, then $\vdash B\top \equiv BC$.

PROOF. Follows immediately from Lemmas 4.6 and 4.7. ■

COROLLARY 4.9. $\vdash B\top \equiv B(A\Box B\top \rhd B\top)$

PROOF. Since $\vdash A\Box B\top \rhd \top, \vdash \Box B\top \to (A\Box B\top \rhd B\top)$. It follows from the above theorem that $\vdash B\top \equiv B(A\Box B\top \rhd B\top)$. ■

LEMMA 4.10. $\vdash \Box A\Box B\top \to (A\Box B\top \rhd B\top \leftrightarrow \Box B\top)$

THEOREM 4.11. $\vdash \Box A\Box B\top \to \boxdot(A\Box B\top \rhd B\top \leftrightarrow \Box B\top)$

THEOREM 4.12. $\vdash A\Box B\top \wedge \Box A\Box B\top \leftrightarrow A(A\Box B\top \rhd B\top) \wedge \Box A(A\Box B\top \rhd B\top)$

PROOF. For the left to right direction, reason in iPL as follows:

$\Box A\Box B\top \to \boxdot(A\Box B\top \rhd B\top \leftrightarrow \Box B\top)$
$\boxdot(A\Box B\top \rhd B\top \leftrightarrow \Box B\top) \to (A\Box B\top \wedge \Box A\Box B\top \leftrightarrow A(A\Box B\top \rhd B\top) \wedge \Box A(A\Box B\top \rhd B\top))$
$A\Box B\top \wedge \Box A\Box B\top \to A(A\Box B\top \rhd B\top) \wedge \Box A(A\Box B\top \rhd B\top)$

For the right to left direction, it suffices to show: $\vdash \Box A(A\Box B\top \rhd B\top) \to \Box A\Box B\top$. Reason in iPL:

$\Box A\Box B\top \to \boxdot(A\Box B\top \rhd B\top \leftrightarrow \Box B\top)$
$A(A\Box B\top \rhd B\top) \to (\Box A\Box B\top \to A\Box B\top)$

$\Box A(A\Box B\top \rhd B\top) \to \Box(\Box A\Box B\top \to A\Box B\top)$
$\Box A(A\Box B\top \rhd B\top) \to \Box A\Box B\top$

■

LEMMA 4.13. $A\Box B\top \equiv A(A\Box B\top \rhd B\top)$.

PROOF. Obviously, $\vdash A\Box B\top \equiv A\Box B\top \wedge \Box A\Box B\top$ (from Lemma 4.6). It follows from Theorem 4.12 that $\vdash A\Box B\top \wedge \Box A\Box B\top \equiv A(A\Box B\top \rhd B\top) \wedge \Box A(A\Box B\top \rhd B\top)$. In addition, $\vdash A(A\Box B\top \rhd B\top) \equiv A(A\Box B\top \rhd B\top) \wedge \Box(A(A\Box B\top \rhd B\top))$.

■

THEOREM 4.14. (Fixed Point Theorem for $A(p) \rhd B(p)$) $\vdash A\Box B\top \rhd B\top \leftrightarrow A(A\Box B\top \rhd B\top) \rhd B(A\Box B\top \rhd B\top)$.[18]

PROOF. This is just a combination of Lemmas 4.9. and 4.13. ■

We may consider boxed formulas to be defined of course, but we can also rely on the fact that the proof of fixed point theorem for such formulas in iPL is the same as that in iL.

THEOREM 4.15. (The Fixed Point Theorem for $\Box$-formulas in $L_\rhd$) $\vdash_{iPL} \Box A\top \leftrightarrow \Box A\Box A\top$ for all formulas A in $L_\rhd$.

We can get a symmetric form of the fixed point for $Ap \rhd Bp$.

THEOREM 4.16. $\vdash A\Box B\top \rhd B\top \leftrightarrow A\Box B\top \rhd B\Box B\top$

PROOF. Since $\vdash_{iPL} \Box B\top \to \Box B\top$, $B\top \equiv B\Box B\top$.

■

Since we have now proved FIX of Theorem 4.2 we can conclude

THEOREM 4.17. (Fixed Point Theorem) For every formula Ap with p modalized, there is formula J such that p does not occur in J and $\vdash_{iPL} J \leftrightarrow AJ$.

PROOF. It has to be checked that
$\vdash_{iPL} \Box(A \leftrightarrow B) \to (A \rhd C \leftrightarrow B \rhd C)$
$\vdash_{iPL} \Box(A \leftrightarrow B) \to (C \rhd A \leftrightarrow C \rhd B)$.
But that is just the Substitution lemma (Lemma 2.8). So iPL satisfies FIX in Theorem 4.2.

■

[18] Following de Jongh-Visser [1991] we can also get an interesting dual result: $\vdash_{iPL} (A\top \rhd B\Box A\top) \leftrightarrow A(B\Box A\top \rhd A\top) \rhd B(B\Box A\top \rhd A\top)$.

In iPW, we have a simpler form of fixed point for $Ap \rhd Bp$.

THEOREM 4.18. In iPW, the fixed point of $Ap \rhd Bp$ is $A\top \rhd B\top$.

PROOF. Reason in iPW:
$\Box B\top \to (\Box B\top \leftrightarrow \top)$
$\Box B\top \to \boxdot(\Box B\top \leftrightarrow \top)$
$\Box B\top \to (A\top \leftrightarrow A\Box B\top)$
In other words, $\vdash A\Box B\top \wedge \Box B\top \leftrightarrow A\top \wedge \Box B\top$. Therefore we can proceed in iPW as follows:
$A\Box B\top \rhd B\top \leftrightarrow (A\Box B\top \wedge \Box B\top) \rhd B\top$
$A\Box B\top \rhd B\top \leftrightarrow (A\top \wedge \Box B\top) \rhd B\top$
$A\Box B\top \rhd B\top \leftrightarrow A\top \rhd B\top$

■

However, in iPL, we can't get such a simpler form. Consider the formula $p \rhd q$. Suppose that the fixed point for formulas $Ap \rhd Bp$ were $A\top \rhd B\top$. Then $\Box q$ would be the fixed point of $p \rhd q$, i.e. $\vdash_{iPL} (\Box q \rhd q) \leftrightarrow \Box q$. It is easy to see that one direction is correct: $\vdash_{iPL} \Box q \to (\Box q \rhd q)$. But for the other direction it is not difficult to construct a countermodel.
Actually the fixed point theorems for IL and ILW (de Jongh and Visser [1991]) may be seen as a consequence of Theorems 4.17 and 4.18.

COROLLARY 4.19. *For every formula Ap with p modalized, there is formula J such that: p does not occur in J, and $\vdash_{IL} J \leftrightarrow AJ$.*

PROOF. Just use the translation discussed at the start of this subsection and note that the principle Dp (which, dually, is not available in IL) has not been used in the above proof. Clearly ILW can be treated similarly.

■

4.3. Beth Definability and Fixed Points

In the following, we will show for a general class of intuitionistic modal logics two theorems (Theorem 4.24 and Theorem 4.25) about the interderivability of the Beth property (Definition 4.22) and the fixed point property (Definition 4.23). The theorem applies to logics in an extended language as e.g. preservativity logics. The theorems and their proofs are an adaptation of the corresponding theorems and proofs of Areces et al. [2000] concerning interpretability logic.

The essential difference lies in an adaptation of Maximova's trick to obtain the Beth property from the existence of fixed points. The problem is of course that fixed points are there for formulas with p modalized only, and Beth's property is supposed to apply to all formulas. Maximova's trick (Maximova[1989]) that was applied in the proof in Areces et al.[2000] relies on the fact that $A(\bar{p}, r)$ is equivalent to $(A_1(\bar{p}, r) \wedge r) \vee (A_2(\bar{p}, r) \wedge \neg r)$ for some A_1, A_2 with r modalized. But this presupposes the existence of a disjunctive normal form unavailable in intuitionistic logic. However (skipping the $\bar{p}$ from here onwards), A_2 is not used in the proof and the role of A_1 can be taken over by the formula arising from the substitution of $\top$ for all the *nonmodalized* occurrences of r. It is easy to see that, for A_1 thus defined, $A_1(r)$ is modalized in r and thus $\vdash_{iPL} \Box(r \leftrightarrow r') \rightarrow (A_1(r) \leftrightarrow A_1(r'))$. The following straightforward lemma about the relation between $A(r)$ and $A_1(r)$ is all we need.

LEMMA 4.20. *For any intuitionistic logic $\mathcal{T}$ with modal operators, if A_1 arises from A by the substitution of $\top$ for all nonmodalized occurrences of r, then* $\vdash_T r \rightarrow (A(r) \leftrightarrow A_1(r))$.

DEFINITION 4.21. (Beth Definability Property) A logic $\mathcal{L}$ has *the Beth Property* iff for all formulas $A(\bar{p}, r)$ the following holds:

- If $\vdash_L \boxdot A(\bar{p}, r) \wedge \boxdot A(\bar{p}, r') \rightarrow (r \leftrightarrow r')$, then there exists a formula $C(\bar{p})$ such that $\vdash_L \boxdot A(\bar{p}, r) \rightarrow (C(\bar{p}) \leftrightarrow r)$.

DEFINITION 4.22. (Fixed Point Property) A logic $\mathcal{L}$ has the *fixed point property* iff, for any formula $A(\bar{p}, r)$ which is modalized in r, there exists a formula $F(\bar{p})$ such that

- (existence) $\vdash_L F(\bar{p}) \leftrightarrow A(\bar{p}, F(\bar{p}))$
- (uniqueness) $\vdash_L \boxdot(r \leftrightarrow A(\bar{p}, r)) \wedge \boxdot(r' \leftrightarrow A(\bar{p}, r')) \rightarrow (r \leftrightarrow r')$.

We now state the theorems in a form that seems more perspicuous than the formulation in Areces et al.[2000]. The properties we require in our formulations for the logics $\mathcal{L}$ are clearly strong enough to ensure the properties of Areces et al.[2000]. (This is because, just as in the classical case, over $iK4$ the rule LR is equivalent to the axiom scheme of iL.)

THEOREM 4.23. (From Beth Definability to Fixed Points) Let $\mathcal{L}$ be an intuitionistic logic with modal operators that extends iL and obeys the substition lemmas and for which the Beth theorem holds.
Then $\mathcal{L}$ has the fixed point property.

PROOF. It is easy to check that the proof for the classical case in Hoogland [2001] is intuitionistically acceptable. ■

THEOREM 4.24. (From Fixed Points to Beth Definability) Let $\mathcal{L}$ be an intuitionistic logic with modal operators that extends iL and obeys the substition lemmas and for which the fixed point theorem holds.
Then $\mathcal{L}$ has the Beth property.

PROOF. Again, the proof is similar to that in Areces et al.[2000]. The only difference is that we will not use Maksimova's lemma (see in Hoogland [2001]) to reduce arbitrary formulas to ones that are "largely modalized" but apply Lemma 4.20 directly.

■

We have shown the fixed point theorem for iL and iPL. Since any extension $\mathcal{L}$ of iL or iPL will have the fixed point property, it should also have the Beth property according to the above theorem.

COROLLARY 4.25. *Let $\mathcal{T}$ be an extension of iL or iPL (of course in the appropriate language). Then $\mathcal{T}$ has the Beth property.*

Acknowledgements

The first author is supported by a Marie Curie fellowship of the European Union under grant HPMF-CT-2001-01383. The third author is very grateful to Prof. Johan van Benthem for his encouragement and his financial support through the Spinoza project "Logic in Action" during his master study in the Graduate Programme in Logic at ILLC.

References

[1] Carlos Areces, Eva Hoogland and Dick de Jongh [2000]: Interpolation, definability and fixed points in interpretability logics, in M. Zakharyaschev, K. Segerberg, M. de Rijke and H. Wansing (eds.), *Advances in Modal Logic, Volume 2*, Lecture Notes, pp. 53–76, Stanford, CA. CSLI Publications.

[2] Milan Bozic and Kosta Došen [1984]: Models for normal intuitionistic modal logics, *Studia Logica*, XLIII, 3, pp. 217-245.

[3] Eva Hoogland [2001]: *Definability and Interpolation*, Ph.D thesis, ILLC-dissertation series, DS-2001-05, Universiteit van Amsterdam.

[4] Rosalie Iemhoff [2001]: *Provability Logic and Admissible Rules*, Ph.D thesis, ILLC-dissertation series, DS-2001-04, Universiteit van Amsterdam.

[5] Rosalie Iemhoff [2003]: Preservativity Logic (An analogue of interpretability logic for constructive theories), *Mathematical Logic Quarterly*, 49(3), 2003, pp. 1-21.

[6] Rosalie Iemhoff, Dick de Jongh and Chunlai Zhou [2004]: Properties of Intuitionistic Provability and Preservativity Logics, *Proceedings of CombLog'04, workshop on Combinations of Logics: Theory and Applications, July 28-30, 2004, CLC IST-Lisbon, Portugal, Walter A. Carnielli, F. Miguel Dionísio, Paul Mateus (eds.)*, CLC, Center for Logic and Computation, Department of Mathematics, Instituto Superior Técnico.

[7] Dick de Jongh and Albert Visser [1991]: Explicit Fixed Points in Interpretability logic, *Studia Logica*, 50, pp. 39-50.

[8] Maximova [1989]: Definability theroems in normal extensions of the provability logic. *Studia Logica*, 50, pp. 495-507

[9] Craig Smoryński [1985]: *Self-reference and Modal Logic*, Springer-Verlag, Berlin.

[10] Robert Solovay [1976]: Provability Interpretations of modal logic, *Israel Journal of Mathematics*, 25, pp, 287-304.

[11] Albert Visser [2002]: Substitutions of Σ^0_1 sentences, *Annals of Pure and Applied Logic* 114(1-3), pp. 227-271.

[12] Chunlai Zhou [2003]: *Some Intuitionistic preservativity Logics and Provability Logics*, Master of Logic thesis, ILLC-MoL series, MoL-2003-01, Universiteit van Amsterdam.

Institute for Discrete Mathematics and Geometry
Technical University Vienna
Wiedner Hauptstrasse
8-10 A-1040 Wien, Austria
iemhoff@logic.at

Institute for Logic, Language and Computation
University of Amsterdam

Plantage Muidergracht 24 1018 TV Amsterdam
The Netherlands
dickdj@science.uva.nl

Department of Mathematics
Indiana University, Bloomington
IN 47405, U.S.A.
czhou@indiana.edu