Abstract. We show that the polynomial translation of the classical propositional normal modal logic $S4$ into the intuitionistic propositional logic $Int$ from Fernández is incorrect. We give a modified translation and prove its correctness, and provide implementations of both translations to allow others to test our results.

§1. Introduction. It is well known that the validity and satisfiability problems for the classical propositional normal modal logic $S4$ and the intuitionistic propositional logic $Int$ are PSPACE-complete and thus there must exist a polynomial translation from each into the other. The Gödel translation [2] provides a translation from $Int$ into $S4$, but the only published polynomial translation from $S4$ into $Int$ we could find is by Fernández [1]. Here, we first show that the translation is incorrect. By pinpointing the flaws, we give a correct polynomial translation from $S4$ into $Int$.

The paper is structured as follows. In Section 2 we define the syntax and Kripke semantics of the propositional intuitionistic logic $Int$ and of the propositional normal modal logic $S4$. In Section 3 we show that the original translation is incorrect. In Section 4 we give our solution and prove it correct.

§2. Semantic Preliminaries. We define $Int$-formulae from an infinite set $Prop$ of propositional variables using the following BNF grammar where $p \in Prop$ and $\bot$ is the falsum constant:

$$\varphi = \bot \mid p \mid \varphi \land \varphi \mid \varphi \lor \varphi \mid \varphi \rightarrow \varphi$$

We also define $\neg \varphi = (\varphi \rightarrow \bot)$. We use rooted Kripke models of $Int$ which are structures $M = (W, R, L, r)$ where: $W$ is a non-empty set of possible worlds; $R$ is a reflexive, transitive and antisymmetric binary relation on $W$; the valuation $L : Prop \rightarrow 2^W$ obeys persistence: if $w \in L(p)$ and $R(w, v)$ then $v \in L(p)$; and $r \in W$ is a root world such that $\forall w \in W.R(r, w)$ holds. Since $Int$ enjoys the finite model property, we can restrict ourselves to models where $W$ is finite.

The semantics of $Int$ are given in Figure 1. An $Int$-formula $\varphi$ is $Int$-satisfiable if there exists some $Int$-model $M$ and some world $w$ in that $Int$-model such that $M, w \models \varphi$. An $Int$-formula is $Int$-valid if $\neg \varphi$ is not $Int$-satisfiable. That is, an $Int$-formula is $Int$-valid if every world $w$ in every $Int$-model $M$ obeys $M, w \models \varphi$. 

1
\[ M, w \not\models \perp \]
\[ M, w \models p \quad \text{iff} \quad w \in L(p) \]
\[ M, w \models \varphi \land \psi \quad \text{iff} \quad M, w \models \varphi \text{ and } M, w \models \psi \]
\[ M, w \models \varphi \lor \psi \quad \text{iff} \quad M, w \models \varphi \text{ or } M, w \models \psi \]
\[ M, w \models \varphi \rightarrow \psi \quad \text{iff} \quad \forall v. \text{ if } R(w, v) \text{ then } M, v \not\models \varphi \text{ or } M, v \models \psi \]

\textbf{Figure 1.} Kripke semantics for \( \mathbf{Int} \)

We define S4-formulae over an infinite set \( Prop \) of propositional variables using the following BNF grammar where \( p \in Prop \) and \( \perp \) is the falsum constant:

\[ \varphi = \perp | p | \varphi \land \varphi | \varphi \lor \varphi | \varphi \rightarrow \varphi | \Box \varphi \]

Again we define \( \neg \varphi = (\varphi \rightarrow \perp) \). We can also define \( \Diamond \varphi = (\neg \Box \neg \varphi) \). For S4, Kripke models are structures \( M = (W, R, L, r) \) where: \( W \) is a non-empty set of possible worlds; \( R \) is a reflexive and transitive binary relation on \( W \); \( L : Prop \rightarrow 2^W \) is a valuation; and \( r \in W \) is a root world obeying \( \forall w \in W. R(r, w) \).

The semantics of S4 are given in Figure 2. An S4-formula \( \varphi \) is S4-satisfiable if there exists some S4-model \( M \) and some world \( w \) in that S4-model such that \( M, w \models \varphi \). An S4-formula is S4-valid if \( \neg \varphi \) is not S4-satisfiable. That is, an S4-formula is S4-valid if every world \( w \) in every S4-model \( M \) obeys \( M, w \models \varphi \).

We can restrict this class further because S4 is complete with respect to the class of binary, reflexive and transitive Kripke frames which are rooted finite trees of finite clusters of worlds where, within a cluster, all worlds are related to each other.

\section*{3. Translating S4-formulae into Int-formulae}

If \( N \) is the number of \( \Box \)-symbols that appear in an S4 formula \( \varphi \), then we can restrict ourselves to those frames with at most \( N + 1 \) distinct clusters along any branch, and each cluster has at most \( N + 1 \) worlds. In such a frame, we say that the level of a world is the number of clusters between the root and the cluster containing that world. If the world is in the root cluster, then it has level 0.

To represent such S4-frames, the translation of Fernández [1] creates multiple Int-propositions \( p_i^j \) for each S4-proposition \( p \) in the S4-formula, intended to represent the valuation of the S4-proposition \( p \) in a world with level \( i \), using \( j \) to

\[ M, w \not\models \perp \]
\[ M, w \models p \quad \text{iff} \quad w \in L(p) \]
\[ M, w \models \varphi \land \psi \quad \text{iff} \quad M, w \models \varphi \text{ and } M, w \models \psi \]
\[ M, w \models \varphi \lor \psi \quad \text{iff} \quad M, w \models \varphi \text{ or } M, w \models \psi \]
\[ M, w \models \varphi \rightarrow \psi \quad \text{iff} \quad \forall v. \text{ if } R(w, v) \text{ then } M, v \not\models \varphi \text{ or } M, v \models \psi \]

\textbf{Figure 2.} Kripke semantics for \( \mathbf{S4} \)
\begin{align*}
(p)^n_m &= p^n_m \\
(\bot)^n_m &= \bot \\
(\psi_1 \cdot \psi_2)^n_m &= (\psi_1)^n_m \cdot (\psi_2)^n_m \\
(\square \psi)^n_m &= b^n_\psi.
\end{align*}

**Figure 3.** Translation from Fernández [1] of an S4-formula \( \psi \) at the \( m \)th world in a cluster at level \( n \) in an S4-model to an \( \text{Int} \)-formula \( (\psi)^n_m \). The propositional variables \( b^n_\psi \) are disjoint from \( p^n_m \) and "·" on the left/right hand side of the equal sign represents the same binary connective of S4/Int respectively.

distinguish between worlds within a cluster in the given S4-model. Figure 3 gives the \( \text{Int} \)-formula \( (\psi)^n_m \) which represents the valuation of the S4-formula \( \psi \) at the \( m \)th world in a cluster at level \( n \) in the S4-model. The branching of the \( \text{Int} \)-model allows for multiple clusters with level \( n \) in the S4-model.

The translation also makes use of new \( \text{Int} \)-propositions \( l^i \) intended to indicate the level of an S4-cluster, and new \( \text{Int} \)-propositions \( b^n_\psi \) to indicate when the S4-formula \( \square \psi \) holds at a cluster of level \( i \).

To determine S4-validity, Fernández [1] defines the translation \( \varphi^\text{Int}_o \), written with an extra subscript \( o \) for "original", as shown in Figure 4. The claim is that \( \varphi^\text{Int}_o \) is \( \text{Int} \)-valid iff \( \varphi \) is S4-valid.

However, there are two errors in this encoding. First, consider a formula \( \varphi \) of S4 with no \( \square \)-formulae. In this case, \( N = 0 \), and so Lev(\( \varphi \)) = \( (l^0 \land \neg l^N) \), leading

\[
\text{Lev}(\varphi) = l^0 \land \neg l^N \land \bigwedge_{k=0}^{N-1} (l^{k+1} \rightarrow l^k)
\]

\[
\text{Mid}^n(\varphi) = l^n \rightarrow \left( \bigwedge_{\square \psi \in \text{sub}(\varphi)} (b^n_\psi \lor \neg b^n_\psi) \land \bigwedge_{p \in \text{sub}(\varphi)} \bigwedge_{0 \leq m \leq N-1} (p)^n_m \lor \neg (p)^n_m \right)
\]

\[
A^n_{o,\psi}(\varphi) = \bigwedge_{n \leq k < N} (l^k \rightarrow l^{k+1} \lor \bigwedge_{m=0}^{N-1} (\psi)^k_m)
\]

\[
\text{Box}^n_{o,\psi}(\varphi) = l^n \rightarrow \left( (b^n_\psi \rightarrow l^{n+1} \lor A^n_{o,\psi}(\varphi)) \land (A^n_{o,\psi}(\varphi) \rightarrow l^{n+1} \lor b^n_\psi) \right)
\]

\[
P(\varphi) = \text{Lev}(\varphi) \land \bigwedge_{0 \leq n \leq N-1} \text{Mid}^n(\varphi) \land \bigwedge_{0 \leq n \leq N-1} \bigwedge_{\square \psi \in \text{sub}(\varphi)} \text{Box}^n_{o,\psi}(\varphi)
\]

\[
\varphi^\text{Int}_o = P(\varphi) \rightarrow (\varphi)^0_0
\]

**Figure 4.** S4 to \( \text{Int} \) translation \( \varphi^\text{Int}_o \) of Fernández [1]
\[ \varphi = (\Box p \rightarrow p) \quad \text{modified } N = 2 \quad \text{sub}(\varphi) = \{p, \Box p\} \]

\[ \text{Lev}(\varphi) = l^0 \land \neg l^N \land \bigwedge_{k=0}^{N-1} (l^{k+1} \rightarrow l^k) = l^0 \land \neg l^2 \land (l^1 \rightarrow l^0) \land (l^2 \rightarrow l^1) \]

\[ \text{Mid}^n(\varphi) = l^n \rightarrow \left( \bigwedge_{\psi \in \text{sub}(\varphi)} (b^n_\psi \lor \neg b^n_\psi) \land \bigwedge_{p \in \text{sub}(\varphi)} \bigwedge_{0 \leq m \leq N-1} (p)^n_m \land \neg (p)^n_m \right) \]

\[ \Box^0_{o,p}(\varphi) = l^0 \rightarrow \left( (b^0_\psi \rightarrow l^1 \lor A^0_{o,p}(\varphi)) \land (A^0_{o,p}(\varphi) \rightarrow l^1 \lor b^0_\psi) \right) \]

\[ A^1_{o,p}(\varphi) = \bigwedge_{1 \leq k < 2} \left( (l^k \rightarrow l^{k+1} \lor \bigwedge_{m=0}^{k-1} (p)^k_m) = l^1 \rightarrow l^{1+1} \lor (b^1_p \land p^1_l) \right) \]

\[ \Box^1_{o,p}(\varphi) = l^1 \rightarrow \left( (b^1_p \rightarrow l^{1+1} \lor A^1_{o,p}(\varphi)) \land (A^1_{o,p}(\varphi) \rightarrow l^{1+1} \lor b^1_p) \right) \]

\[ P(\varphi) = \text{Lev}(\varphi) \land \bigwedge_{0 \leq n \leq 1} \text{Mid}^n(\varphi) \land \Box^0_{o,p}(\varphi) \land \Box^1_{o,p}(\varphi) \]

\[ \varphi^\text{Int}_o = P(\varphi) \rightarrow (\varphi)^0_0 \]

\[ w \models l^0, l^1, p^0_0, p^0_1, p^1_0, b^0_p, b^1_p \]

**Figure 5.** Computation of \( \varphi^\text{Int}_o \) using the Fernández translation from Example 1. Underlines indicate the formulae that are “true” at \( w \) in the given model and which directly influence the truth value of the larger formula.

to \( \varphi^\text{Int}_o = \bot \rightarrow \varphi^0_0 \), which is \( \text{Int} \)-valid regardless of \( \varphi \). The obvious solution here is to modify \( N \) to be one more than the number of \( \Box \) symbols in \( \varphi \).

The second error is more subtle, and we demonstrate it via an example.

**Example 1.** Consider the \( \text{S}4 \)-valid formula \( \varphi = (\Box p \rightarrow p) \) which has only one \( \Box \)-symbol and thus a modified \( N \) of 2. The translation given in Figure 4 requires the following \( \text{Int} \)-formulae as new propositions: \( l^0, l^1, l^2 \) for level formulae, \( b^0_p, b^1_p \) to represent the formula \( \Box p \), and \( p^0_0, p^0_1, p^1_0, p^1_1 \) to represent the value of \( p \) in up to two worlds and up to two levels.

Now consider the \( \text{Int} \)-model \( M = \{ (w), \{(w, w)\}, L, w \} \) with a single reflexive world \( w \), and \( w \in L(\psi) \) for \( \psi \in \{ l^0, l^1, p^0_0, p^0_1, p^1_0, p^1_1, b^0_p, b^1_p \} \), and \( w \notin L(\psi) \) for \( \psi \in \{ l^2, p^0_0 \} \).

Referring to Figure 5 where underlines indicate the parts that are “true” at \( w \) and which directly affect the truth value of the larger formulae, we obviously have \( M, w \models \text{Lev}(\varphi) \). We have \( M, w \models \text{Mid}^n(\varphi) \) because in a single-world model, \( \psi \lor \neg \psi \) is intuitionistically true for all \( \psi \). Since \( M, w \models l^1 \), we have \( M, w \models \Box^0_{o,p}(\varphi) \), because the inner implications are made true by the “escape hatch” provided by \( l^{n+1} = l^1 \). We also have \( M, w \models p^0_0 \land p^1_1 \), thus \( M, w \models A^1_{o,p}(\varphi) \), and since \( M, w \models b^1_p \), we have \( M, w \models \Box^1_{o,p}(\varphi) \). Thus we have \( M, w \models P(\varphi) \). However, \( M, w \not\models \varphi^0_0 = b^1_p \rightarrow p^0_0 \), and so \( M \) is an \( \text{Int} \)-countermodel to \( \varphi^\text{Int}_o \), despite \( \varphi \) being \( \text{S}4 \)-valid.
The culprits are the “escape hatches” \( l^{n+1} \) in Box\(^n\)\(_{o,\psi} (\varphi) \) and \( l^{k+1} \) in \( A_{o,\psi}^{n} (\varphi) \), which allow us to ignore constraints imposed by \( l^{0} \) by jumping straight to \( l^{1} \).

What happens if we keep the original definition of \( N \) and just drop the \( l^{N} \) part from \( \text{Lev}(\varphi) \)? Example \([\text{II}]\) is no longer a counter-example, but using \( \varphi = (\Box p \rightarrow p) \land (\Box q \rightarrow q) \) with exactly the same structure does give a counterexample.

\[ \text{§4. Solution.} \] As mentioned before, the first step of the solution is to modify \( N \) to be one more than the number of \( \Box \)-formulae in the given S4-formula \( \varphi \). This is probably what was intended, as no proofs need to change and there are no 0-standard models, as defined by Fernández [1]. To avoid confusion, we will retain \( N \) as the number of \( \Box \)-symbols, as used by Fernández [1], and use \( M = N + 1 \) for the modified value.

The second change is to modify the definition of \( A_{o,\psi}^{n} \) and Box\(^n\)\(_{o,\psi} \) as follows:

\[
A_{o,\psi}^{n} (\varphi) = \bigwedge_{n \leq k < M} \left( l^{k} \rightarrow \bigwedge_{m=0}^{M-1} (\psi)_{m}^{k} \right)
\]

\[
\text{Box}_{o,\psi}^{n} (\varphi) = l^{n} \rightarrow ((b_{\psi}^{n} \rightarrow A_{o,\psi}^{n} (\varphi)) \land (A_{o,\psi}^{n} (\varphi) \rightarrow b_{\psi}^{n}))
\]

This removes the “escape hatches” in the formula in the case where a higher \( l \) proposition was true. All conditions imposed on formulae by some \( l^{i} \) must be met, regardless of whether other \( l^{j} \) formulae are true. For example we no longer have \( \mathcal{M}, w \models P(\varphi) \) in Example \([\text{II}]\) because \( \mathcal{M}, w \not\models \text{Box}_{o,\psi}^{n} (\varphi) \): that is, we have \( \mathcal{M}, w \not\models b_{\psi}^{n} \) and \( \mathcal{M}, w \models l^{0} \), but \( \mathcal{M}, w \not\models A_{o,\psi}^{0} (\varphi) \) because \( \mathcal{M}, w \not\models p_{0}^{0} \).

We write \( \varphi_{\text{c}}^{\text{Int}} \) for our “correct” translation (we cannot use \( n \) for “new” as it clashes with the integers used as subscripts). Note that our translation \( \varphi_{\text{c}}^{\text{Int}} \) is actually smaller than the translation \( \varphi_{o}^{\text{Int}} \) of Fernández [1] since all we have done is remove some disjunctions, and so the translation remains polynomial.

\[ \text{4.1. Converting Int-models to S4-models.} \] We work with rooted and finite \( \text{Int} \)-models \( \mathcal{M} = (W, R, L, r) \). We intend to show that the modified \( \varphi_{\text{c}}^{\text{Int}} \) has an \( \text{Int} \)-countermodel iff \( \varphi \) has an S4-countermodel.

First, we prove some lemmas about small modifications to \( \text{Int} \)-models.

\[ \text{Definition 2.} \] Given an \( \text{Int} \)-model \( \mathcal{M} = (W, R, L, r) \), a world \( u \in W \), and a finite set \( L \) of propositional variables such that \( \forall w \in W, \forall p \in L, \text{if } R(w, u) \text{ and } w \neq u \text{ then } w \not\in L(p) \). Define insert(\( L, u, \mathcal{M} \) = (\( W', R', L', r' \)) as follows:

1. let \( v \) be a new world not in \( W \)
2. if \( r = u \) then \( r' = v \) otherwise \( r' = r \)
3. \( W' = W \cup \{ v \} \)
4. \( R' = R \cup \{ (v, v) \} \cup \{ (v, x) | (u, x) \in R \} \cup \{ (y, v) | (y, u) \in R \land y \neq u \} \)
5. for all \( p \) we have \( L'(p) \cap W = L(p) \)
6. for all \( p \not\in L \) we have \( v \in L'(p) \) iff \( u \in L'(p) \)
7. for all \( p \in L \) we have \( v \not\in L'(p) \).

That is, we insert a new world \( v \) as an immediate predecessor of \( u \), where all proper predecessors \( y \) of \( u \) are made proper predecessors of \( v \) and all successors \( x \) of \( u \) including \( u \) itself are made successors of \( v \).
LEMMA 3. If $\mathcal{M} = (W, R, L, v)$ is an \textit{Int}-model with a world $u$, and $\mathcal{L}$ is a set of propositional variables falsified at all $y$ such that $R(y, u)$ and $y \neq u$, then $\mathcal{M} = \text{insert}(\mathcal{L}, u, \mathcal{M})$ is an \textit{Int}-model and for all \textit{Int}-formulae $\psi$ which do not include propositions from $\mathcal{L}$ and for all $w \in W$, we have $\mathcal{M}', w \models \psi$ iff $\mathcal{M}, w \models \psi$, and additionally we have $\mathcal{M}', v \models \psi$ iff $\mathcal{M}, u \models \psi$.

PROOF. We first prove that $\mathcal{M}'$ is still an \textit{Int} -model. That is, we have to prove that $\mathcal{M}'$ is transitive, reflexive, antisymmetry and persistence. Of these, we deal only with the non-trivial cases.

Transitivity still holds: the only case that could possibly fail is $R'(a, b)$ and $R'(b, v)$ but not $R'(a, v)$ for some $a \neq v$ and $b \neq v$. Since both $a$ and $b$ are in the original model, the edge $R'(a, b)$ is from the original model, hence $R(a, b)$. Since $R'(b, v)$, we must have $R(b, u)$ and $b \neq u$ by definition of $R'$. By the transitivity of $R$ we must have $R(a, u)$, and thus by definition of $R'$ we must have $R'(a, v)$ as required.

The valuation $L'$ obeys the persistence property: because the original model had a persistent valuation, the only way for $\mathcal{M}'$ to not have a persistent valuation is if the introduction of $v$ changed something. Suppose for a contradiction that $\mathcal{M}'$ does not have a persistent valuation.

Now the step cases, using the following inductive hypotheses:

IH1: for all subformulae $\phi$ of $\psi$ we have $\mathcal{M}, u \models \phi$ iff $\mathcal{M}', v \models \phi$.

IH2: for all subformulae $\phi$ of $\psi$ and for all worlds $w \in W$ we have $\mathcal{M}, w \models \phi$ iff $\mathcal{M}', w \models \phi$.

...
Suppose instead that $\mathcal{M}, u \not\models \psi_1 \to \psi_2$. Then there must exist a witness $w \in W$ such that $R(u, w)$ and $\mathcal{M}, w \models \psi_1$ and $\mathcal{M}, w \not\models \psi_2$. But this same witness will also exist in $\mathcal{M}'$ by IH2, thus $\mathcal{M}', w \models \psi_1$ and $\mathcal{M}', w \not\models \psi_2$. Since $w$ is reachable from $u$, and $v$ is a predecessor of $u$, we must also have $w$ reachable from $v$, and thus $\mathcal{M}', w \not\models \psi_1 \to \psi_2$.

For any $w \in W$, suppose that $\mathcal{M}, w \models \psi_1 \to \psi_2$. Then for all successors $x$ of $w$, if $\mathcal{M}, x \models \psi_1$ then $\mathcal{M}, x \models \psi_2$. Thus by IH2, for all $x \neq v$ with $R(w, x)$, if $\mathcal{M}', x \models \psi_1$ then $\mathcal{M}', x \models \psi_2$. If $v$ is a successor of $w$ in $\mathcal{M}'$ then $u$ must also be a successor of $w$ in $\mathcal{M}'$, and so by IH1, if $\mathcal{M}', v \models \psi_1$ then $\mathcal{M}', v \models \psi_2$. Thus $\mathcal{M}', w \models \psi_1 \to \psi_2$ as required.

If instead $\mathcal{M}, w \not\models \psi_1 \to \psi_2$ then there is some successor $x$ of $w$ such that $\mathcal{M}, x \models \psi_1$ and $\mathcal{M}, x \not\models \psi_2$. By IH2, we have $\mathcal{M}', x \models \psi_1$ and $\mathcal{M}', x \not\models \psi_2$, and thus $\mathcal{M}', w \not\models \psi_1 \to \psi_2$.

Effectively Lemma 3 states that we can insert “copies” of worlds with minor changes to some atomic propositions $\mathcal{L}$ without changing the truth values of formulae which do not refer to those atomic propositions.

Next we prove that if our amended $\varphi_c^{\text{Int}}$ has an Int-countermodel, then $\varphi$ has an $S4$-countermodel.

**Definition 4.** If $\mathcal{M} = (W, R, L, r)$ is an Int-model such that $\mathcal{M}, r \models P(\varphi)$, then for $w \in W$, let $Lv(w)$ be defined as the index $i$ such that $w \in L(l^i)$ and $w \notin L(l^{i+1})$.

As long as $\mathcal{M}, w \models \text{Lev}(\varphi)$ then $Lv(w)$ has a unique definition because then we must have $\mathcal{M}, w \models l^0$ and thus $Lv(w) \geq 0$, and we must have $\mathcal{M}, w \not\models l^M$ and thus $Lv(w) < M$, and we must have that if $\mathcal{M}, w \models l^k$ then $\mathcal{M}, w \models l^j$ for all $j < k$.

**Definition 5.** A model $\mathcal{M} = (W, R, L, r)$ which falsifies $\varphi_c^{\text{Int}}$ is stratified if:

1. $Lv(r) = 0$;
2. for any two worlds $w, v \in W$, if $R(w, v)$ and $Lv(w) > Lv(w) + 1$ then there is another (necessarily different) world $u$ such that $R(w, u)$ and $R(u, v)$ with $Lv(u) = Lv(w) + 1$; and
3. if for some $w, u \in W$ we have $R(w, u)$ and $Lv(w) = Lv(u)$ then $w = u$.

We now prove that there must be a stratified Int-countermodel to $\varphi_c^{\text{Int}}$ if there is any Int-countermodel of $\varphi_c^{\text{Int}}$.

**Lemma 6.** If a countermodel to $\varphi_c^{\text{Int}}$ exists, then one satisfying condition 3 of Definition 2 exists.

**Proof.** Let $\mathcal{M} = (W, R, L, r)$ be an Int-countermodel of $\varphi_c^{\text{Int}}$. Without loss of generality, assume $\mathcal{M}, r \models P(\varphi)$ and $\mathcal{M}, r \not\models (\varphi)_0^0$. If $Lv(r) = 0$ then the lemma holds immediately. Otherwise $Lv(r) \geq 1$ and so we have $\mathcal{M}, r \models l^1$ and $\mathcal{M}, r \models l^0$. Create a new Int-model $\mathcal{M}' = \text{insert}(\mathcal{L}, r, \mathcal{M}) = (W', R', L', r')$ according to Lemma 3 using $\mathcal{L} = \{l^i \mid 0 < i \leq M\}$.

The new model $\mathcal{M}'$ still falsifies $\varphi_0^0$ at the new root $r'$ according to Lemma 3 because $\varphi_0^0$ does not refer to any proposition in $\mathcal{L}$. Note that $Lv(r') = 0$ by the definition of $L'$ as required. It remains to show that $\mathcal{M}', r' \models P(\varphi)$. 
We obviously have $\mathcal{M}', r' \vDash \text{Lev}(\varphi)$. The successors of $r$ are also successors of $r'$, so the only way for $r'$ to fail $\text{Mid}^l(\varphi)$ would be to fail locally. Since $\mathcal{M}', r' \vDash \text{I}'$ only for $i = 0$, we have $\mathcal{M}', r' \vDash \text{Mid}^l(\varphi)$ for $i > 0$. For $i = 0$, $\text{Mid}^0(\varphi)$ does not refer to any propositions in $L$ and thus by Lemma 8 we must also have $\mathcal{M}', r' \vDash \text{Mid}^0(\varphi)$.

Finally, we show that $r'$ satisfies $\text{Box}^n(\varphi)$. For $n > 0$ it satisfies $\text{Box}^n(\varphi)$ vacuously because $\mathcal{M}', r' \not\vDash \text{I}'$, and all strict successors of $r'$ satisfy $\text{Box}^n(\varphi)$ because they did in $\mathcal{M}$. For $n = 0$, we have $\mathcal{M}, r \vDash b^0_\psi \leftrightarrow A^0_\psi(\varphi)$, and we want to show that $\mathcal{M}', r' \vDash b^0_\psi \leftrightarrow A^0_\psi(\varphi)$. Because $b^0_\psi \not\in L$, we have $\mathcal{M}, r' \vDash b^0_\psi$ iff $\mathcal{M}, r \vDash b^0_\psi$, so it remains to show that $\mathcal{M}, r \vDash A^0_\psi(\varphi)$ iff $\mathcal{M}', r' \vDash A^0_\psi(\varphi)$.

Suppose that $\mathcal{M}, r \not\vDash A^0_\psi(\varphi)$. Then there must be some successor which satisfies $\psi^k$ and falsifies $\psi^k_m$ for some $k$ and $m$, and such a successor is also a successor of $r'$ thus $\mathcal{M}', r' \not\vDash A^0_\psi(\varphi)$.

Suppose instead that $\mathcal{M}, r \vDash A^0_\psi(\varphi)$ and thus since $\mathcal{M}, r \vDash \text{I}'$ we must have $\mathcal{M}, r \vDash (\psi)^0_m$ for all $0 \leq m < 0$. The only way that $r'$ could fail to satisfy $A^0_\psi(\varphi)$ is to do so locally, and with $k = 0$. However, since $(\cdot)^0_m$ does not refer to any $l^i$, we must also have $\mathcal{M}', r' \vDash (\psi)^0_m$ iff $\mathcal{M}, r \vDash (\psi)^0_m$ using Lemma 8 so $\mathcal{M}', r' \vDash (\psi)^0_m$ and thus $\mathcal{M}', r' \vDash A^0_\psi(\varphi)$.

Thus $\mathcal{M}', r' \vDash P(\varphi)$, and $\mathcal{M}', r' \not\vDash \varphi^0_0$, hence $\mathcal{M}'$ is a countermodel to $\varphi^\text{Int}_c$ with $\text{Lev}(r') = 0$ as required.

Note that Lemma 8 does not hold for the original specification of $\varphi^\text{Int}_c$ from Fernández [1]: the counterexample we gave cannot be converted to one with $\text{Lev}(r) = 0$ while still satisfying the original $P(\varphi)$. In particular $\text{Box}^n(\varphi)_\psi(\varphi)$ will fail to hold if $l^i$ is false at the root as required by $\text{Lev}(r) = 0$.

**Lemma 7.** If an $\text{Int}$-countermodel of $\varphi^\text{Int}_c$ exists, then one satisfying conditions 2 and 3 of Definition 5 exists.

**Proof.** Let $\mathcal{M} = (W, R, L, r)$ be an $\text{Int}$-countermodel of $\varphi^\text{Int}_c$ after applying Lemma 8 with $w, u \in W$ such that $\text{Lev}(w) = j, \text{Lev}(v) > j + 1, R(w, v)$. Thus we have $\mathcal{M}, w \vDash \text{I}'$ and $\mathcal{M}, w \not\vDash \text{I}'+1$, and $\mathcal{M}, v \vDash \text{I}'+2$. Suppose that there is no $u$ such that $R(u, v), R(u, v)$ and $\text{Lev}(u) = j + 1$, and thus condition 2 does not hold. Let $\mathcal{L} = \{l^i | j + 1 < i \leq \text{Lev}(v)\}$, and consider $\mathcal{M}' = \text{insert}(\mathcal{L}, v, \mathcal{M})$ where the newly introduced world is $u$.

That is, $u$ is a copy of $v$, added between $w$ and $v$ with the valuation only differing on the level variables in $\mathcal{L}$. Note that $\text{Lev}(u) = j + 1$ because $\text{I}'+1 \not\in \mathcal{L}$ and so $\mathcal{M}'$, $u \vDash \text{I}'+1$, but $\text{I}'+2 \in \mathcal{L}$ so $\mathcal{M}'$, $u \not\vDash \text{I}'+2$.

A similar argument to Lemma 8 applies, again using Lemma 8. The structure $\mathcal{M}'$ is an $\text{Int}$-model, the truth of formulae which do not refer to $l^k \in \mathcal{L}$ does not change between $\mathcal{M}$ and $\mathcal{M}'$, and the truth of the formulae which do refer to $l^k \in \mathcal{L}$ is preserved because the $l^k$ are falsified on the left of an implication.

Let the “gap” between a world $x$ and one of its immediate successors $y$ be defined as $\text{Lev}(y) - \text{Lev}(x) - 1$ if $\text{Lev}(y) > \text{Lev}(x)$, and 0 if $\text{Lev}(y) = \text{Lev}(x)$. The sum of these gaps is unchanged between $\mathcal{M}$ and $\mathcal{M}'$ except that for the gaps between $v$ and the immediate predecessors of $v$ the gap between $w$ and $u$ and the previous immediate successors of $w$ is decreased by 1, so the total sum of the gaps decreases through this process. Since our
Int-models are finite we repeat the process until condition 2 holds. Note that because the original model satisfies Condition 1, and because we do not change the root (we add a world in between two other existing worlds) the model \( \mathcal{M}' \) must still satisfy Condition 1.

Note that this may break Condition 3, since the world \( v \) may already have a predecessor \( x \) with level \( j + 1 \), but \( x \) is not a successor of \( w \). When we introduce the new world \( u \) we make \( u \) a successor of \( x \), which causes Condition 3 to fail.

**Lemma 8.** If an Int-countermodel to \( \varphi^\text{Int}_c \) exists, then one satisfying all three conditions of Definition 3 exists.

**Proof.** Let \( \mathcal{M} = (W, R, L, r) \) be an Int-countermodel of \( \varphi^\text{Int}_c \) satisfying conditions 1 and 2 after applying Lemma 7 with worlds \( a, b \in W \) such that \( \text{Lv}(a) = \text{Lv}(b), R(a, b) \) and \( a \neq b \), thus breaking condition 3.

There must be a pair of “adjacent” worlds \( w \) and \( u \) such that \( \text{Lv}(w) = \text{Lv}(u) \), \( R(w, u) \) and \( w \neq u \) and there is no distinct \( v \) such that \( R(w, v) \) and \( R(v, u) \). We show that we get closer to satisfying condition 3 by removing the edge \( R(w, u) \).

Let \( \mathcal{M}' = (W', R', L, r) \) where \( R' = R \setminus \{(w, u)\} \).

The relation \( R' \) is still transitive because \( R \) was, and there is no “intermediate” world \( v \) that could require the removed edge. Reflexivity and antisymmetry are also preserved.

Suppose that \( \mathcal{M}, r \models P(\varphi) \), but \( \mathcal{M}', r \not\models P(\varphi) \). The only change is the removal of \( R(w, u) \), so it is simple to see that \( \mathcal{M}' \models \text{Lev}(\varphi) \) and \( \mathcal{M}' \models \text{Mid}^n(\varphi) \). Therefore we must have \( \mathcal{M}' \not\models \text{Box}^n(\varphi) \). Thus there must be some world \( x \) such that \( \mathcal{M}' \models \text{Lev}(\varphi) \) and \( \mathcal{M}', x \not\models \text{Box}^n(\varphi) \) or \( \mathcal{M}' \models \text{Mid}^n(\varphi) \). We consider each case to obtain a contradiction.

Suppose that \( \mathcal{M}', x \not\models \text{Box}^n(\varphi) \). Expanding the semantics, there must therefore be some indices \( k \) and \( m \) and some world \( y \) such that \( \mathcal{M}', y \models \text{Box}^k(\varphi) \) and \( \mathcal{M}', y \not\models \text{Box}^m(\varphi) \). All propositional variables referred to by \( (\psi)_m^k \) will have superscript \( k \), and since \( \mathcal{M}', y \models \text{Mid}^k(\varphi) \) we must have \( \mathcal{M}', y \models \varphi_k \) or \( \mathcal{M}', y \models \varphi_k \rightarrow \bot \) for all propositional variables \( \varphi_k \), thus the valuations are fixed in all successors. The valuations are common between \( \mathcal{M} \) and \( \mathcal{M}' \), thus \( \mathcal{M}, y \models (\psi)_m^k \) as well, and so \( \mathcal{M} \not\models \text{Box}^n(\varphi) \), a contradiction.

Suppose instead that \( \mathcal{M}', x \not\models \text{Box}^n(\varphi) \rightarrow b_m^n(\varphi) \). There must therefore be a world \( y \) such that \( \mathcal{M}, y \models \text{Box}^n(\varphi) \) and \( \mathcal{M}', y \models b_m^n(\varphi) \). Because \( \mathcal{M} \models \text{Box}^n(\varphi) \) and \( \mathcal{M}, y \models \text{Box}^n(\varphi) \), we must have \( \mathcal{M}, y \models \varphi_k \) or \( \mathcal{M}, y \models \varphi_k \rightarrow \bot \), and because \( \mathcal{M}, y \models \varphi_k \) we must have \( \mathcal{M}, y \models \text{Box}^n(\varphi) \rightarrow b_m^n(\varphi) \). Thus the witness falsifying \( \text{Box}^n(\varphi) \) must be \( u \) and \( w \) (otherwise the witness would still exist in \( \mathcal{M}' \)); that is, \( \mathcal{M}, u \models \text{Box}^n(\varphi) \) and \( \mathcal{M}, u \models \text{Box}^n(\varphi) \rightarrow b_m^n(\varphi) \). However, this means that \( \text{Lv}(u) \geq k \), and thus \( \text{Lv}(w) \geq k \). Since \( \mathcal{M}, w \models \text{Mid}^k(\varphi) \) we have \( \mathcal{M}, w \models \varphi_k \) or \( \mathcal{M}, w \models \varphi_k \rightarrow \bot \), and since \( R(w, u) \) we must have \( \mathcal{M}, w \models \varphi_k \) iff \( \mathcal{M}, u \models \varphi_k \). Thus we must have \( \mathcal{M}', w \not\models \text{Box}^n(\varphi) \), a contradiction.

Thus \( \mathcal{M}', r \not\models P(\varphi) \), and \( \mathcal{M}', r \not\models \varphi_0^\text{Int} \), and so \( \mathcal{M}' \) is a countermodel with at least one fewer instance of Condition 3 failing. Since Int has the finite model property we can begin with a finite model (and a finite number of failures of condition 3) and repeat the process until condition 3 holds. Since we only remove
edges between worlds with the same level, we do not break either Condition 1
or Condition 2 if they hold initially.

COROLLARY 9. If there is some \( \text{Int} \)-countermodel to \( \varphi^\text{Int}_c \) then there is a stratified \( \text{Int} \)-countermodel to \( \varphi^\text{Int}_c \).

PROOF. Given an arbitrary finite rooted \( \text{Int} \)-countermodel to \( \varphi^\text{Int}_c \), apply Lemma 6 to obtain a model satisfying Condition 1, then Lemma 7 to introduce worlds to satisfy Condition 2 without destroying Condition 1. Finally we use Lemma 8 to combine worlds with the same level to satisfy condition 3 without breaking condition 2 or condition 1. Finally we use Lemma 6 to obtain a model satisfying Condition 1, then Lemma 7 to introduce worlds to satisfy Condition 2 without destroying Condition 1. Finally we use Lemma 8 to combine worlds with the same level to satisfy condition 3 without breaking condition 2 or condition 1.

We now show how \( \text{Int} \)-countermodels of \( \varphi^\text{Int}_c \) correspond to \( \text{S4} \)-countermodels of \( \varphi \) following Fernández [1] but being mindful of our modifications.

DEFINITION 10. Let \( \mathcal{M}^\text{Int} = (W^\text{Int}, R^\text{Int}, L^\text{Int}, P^\text{Int}) \) be a stratified \( \text{Int} \) countermodel for \( \varphi^\text{Int}_c \), such that \( \mathcal{M}^\text{Int}, x^\text{Int} \models \varphi^\text{Int}_0 \) and \( \mathcal{M}^\text{Int}, x^\text{Int} \models P(\varphi) \).

For each \( x \in W^\text{Int} \), let \( \mathcal{I} = \{ x_0, \ldots, x_{M-1} \} \) be a set of \( M \) distinct worlds, and let \( W^{\text{Int} \rightarrow \text{S4}} \) be the disjoint union of all \( \mathcal{I} \). Let \( R^{\text{Int} \rightarrow \text{S4}} = \{ (x_m, y_n) \mid R^\text{Int}(x, y) \} \),

Define \( \mathcal{M}^{\text{Int} \rightarrow \text{S4}} = (W^{\text{Int} \rightarrow \text{S4}}, R^{\text{Int} \rightarrow \text{S4}}, L^{\text{Int} \rightarrow \text{S4}}, I^{\text{Int}}) \).

LEMMA 11. If \( \psi \) is a subformula of \( \varphi \), then \( \mathcal{M}^{\text{Int} \rightarrow \text{S4}}, x_m \models \psi \) iff \( \mathcal{M}^{\text{Int}}, x \models \psi \).

PROOF. We proceed by induction on the structure of \( \psi \). First the base cases:

\[ \psi = \bot: \text{Trivially true.} \]
\[ \psi = p: \text{By the definition of } L^{\text{Int} \rightarrow \text{S4}} \text{ the lemma holds.} \]

Now the step cases, using the inductive hypothesis that for all formulae smaller than \( \psi \) the property already holds.

\( \psi = \psi_1 \land \psi_2 \): By definition, we have \( \mathcal{M}^{\text{Int} \rightarrow \text{S4}}, x_m \models \psi_1 \land \psi_2 \) iff \( \mathcal{M}^{\text{Int} \rightarrow \text{S4}}, x_m \models \psi_i \) for all \( i \in \{1, 2\} \). By the induction hypothesis, \( \mathcal{M}^{\text{Int} \rightarrow \text{S4}}, x_m \models \psi_i \) iff \( \mathcal{M}^{\text{Int}}, x \models (\psi_i)^{L_v(x)} \), and thus \( \mathcal{M}^{\text{Int}}, x \models (\psi_1 \land \psi_2)^{L_v(x)} \) as required.

\( \psi = \psi_1 \lor \psi_2 \): As above.

\( \psi = \psi_1 \rightarrow \psi_2 \): If \( \mathcal{M}^{\text{Int} \rightarrow \text{S4}}, x_m \models \psi_1 \rightarrow \psi_2 \) then \( x_m \) either satisfies \( \psi_2 \) or falsifies \( \psi_1 \). By induction this translates to \( \mathcal{M}^{\text{Int}} \), thus \( \mathcal{M}^{\text{Int}}, x \not\models (\psi_1)^{L_v(x)} \) or \( \mathcal{M}^{\text{Int}}, x \models (\psi_2)^{L_v(x)} \). Both of these formulae refer to only propositional atoms indexed by \( m \), and so because \( \text{Mid}^{L_v(x)}(\varphi) \) holds, all successors of \( x \) will give the same valuation, and thus either satisfy \( (\psi_2)^{L_v(x)} \) or falsify \( (\psi_1)^{L_v(x)} \), and so \( \mathcal{M}^{\text{Int}}, x \models (\psi_1 \rightarrow \psi_2)^{L_v(x)} \).

If instead \( \mathcal{M}^{\text{Int}}, x \not\models (\psi_1 \rightarrow \psi_2)^{L_v(x)} \), then because \( R^{\text{Int}} \) is reflexive we must have \( \mathcal{M}^{\text{Int}}, x \not\models (\psi_1)^{L_v(x)} \) or \( \mathcal{M}^{\text{Int}}, x \models (\psi_2)^{L_v(x)} \). Using the inductive hypothesis, we thus have \( \mathcal{M}^{\text{Int} \rightarrow \text{S4}}, x_m \models \psi_1 \rightarrow \psi_2 \) as required.

\( \psi = \Box \psi_1 \): Because Box \( (\psi_1)^{L_v(x)} \) (\( \varphi \) holds, we have \( \mathcal{M}^{\text{Int}}, x \not\models b^{L_v(x)} \) iff \( \forall y. R^{\text{Int}}(x, y) \) implies \( \forall k. \mathcal{M}^{\text{Int}}, y \models (\psi_1)^{L_v(y)} \). By induction, for each of these worlds \( y \) we have \( \mathcal{M}^{\text{Int} \rightarrow \text{S4}}, y_k \models \psi_1 \). By the definition of \( R^{\text{Int} \rightarrow \text{S4}} \), these \( y_k \) are exactly the worlds such that \( R^{\text{Int} \rightarrow \text{S4}}(x_m, y_k) \), thus we have \( \mathcal{M}^{\text{Int}}, x \models b^{L_v(x)} \) if
\[ \forall y_k, R(x_m, y_k) \] implies \( \mathcal{M}_{\text{Int}\rightarrow S4}, y_k \models \psi_1 \). This is exactly the definition of \( \mathcal{M}_{\text{Int}\rightarrow S4}, x_m \models \Box \psi_1 \).

\[ \neg \]

**COROLLARY 12.** If there is an \( \text{Int} \)-countermodel to \( \varphi_{c}^{\text{Int}} \) then there is an \( S4 \)-countermodel to \( \varphi \). Equivalently, if \( \varphi \) is \( S4 \)-valid, then \( \varphi_{c}^{\text{Int}} \) is \( \text{Int} \)-valid.

**PROOF.** By Corollary 11 if there is an \( \text{Int} \)-countermodel to \( \varphi_{c}^{\text{Int}} \) then there must be a stratified \( \text{Int} \)-countermodel \( \mathcal{M}_{\text{Int}} \) as well. Construct \( \mathcal{M}_{\text{Int}\rightarrow S4} \) as described in Definition 10. Applying Lemma 11 to \( \mathcal{M}_{\text{Int}\rightarrow S4} \) and choosing \( \psi = \varphi \), we find that because \( \mathcal{M}_{\text{Int}} \models \varphi_{c}^{\text{Int}} \), we must have \( \mathcal{M}_{\text{Int}\rightarrow S4}, r_{0}^{\text{Int}} \not\models \varphi \), as required.

**4.2. Converting \( S4 \)-models to \( \text{Int} \)-models.** It remains to show that the converse holds, that if there is an \( S4 \)-countermodel to \( \varphi \) then there is an \( \text{Int} \)-countermodel to \( \varphi_{c}^{\text{Int}} \).

We will use the same notion of \( N \)-standard frames as Fernández [1], though we refer to it as \( M \)-standard to avoid confusion between the \( N \) used by Fernández [1] and the \( M = N + 1 \) that we use. If \( K = (W, R) \) is an \( S4 \)-frame, then let \( \pi \) denote the \( R \)-equivalence class of worlds \( \{ y \mid (x, y) \in R \} \). The quotient \( W/R \) with induced relation \( \overline{R} \) forms a partial order since \( R \) is transitive and reflexive, and taking the quotient ensures that it is antisymmetric as well.

**DEFINITION 13.** An \( S4 \) Kripke frame \( K = (W, R) \) is \( M \)-standard if:

1. Any strictly ascending chain in \( \overline{R} \) has length shorter than \( M \);
2. For all \( x \in W \), \( \pi \) has exactly \( M \) elements, \( \{ x_0, \ldots, x_{M-1} \} \);
3. \( (W/R, \overline{R}) \) forms a tree.

Fernández [1] proves the following theorem:

**THEOREM 14 (Theorem 5.1 of [1]).** If \( \mathcal{M} = (W, R, L, r) \) is an \( S4 \)-model, and \( \varphi \) is a formula of \( S4 \), then there is an \( M \)-standard model \( \mathcal{M}^\varphi \), such that for all subformulas \( \psi \) of \( \varphi \), we have \( \mathcal{M}^\varphi, r^\varphi \models \psi \iff \mathcal{M}, r \models \psi \).

Thus if there is a countermodel to \( \varphi \), then there is an \( M \)-standard countermodel to \( \varphi \). Let \( \mathcal{M}_{S4} = (W_{S4}, R_{S4}, L_{S4}, r_{S4}) \) be such a model. Let \( L_{\pi}(\overline{x}) \) be the length of the shortest chain \( \overline{R}(\overline{w}_1, \overline{w}_2), \ldots, \overline{R}(\overline{w}_{n-1}, \overline{x}) \) where each \( w_i \) is distinct, and there is no intermediate such that \( \overline{R}(\overline{w}_i, \overline{u}) \) and \( \overline{R}(\overline{u}, \overline{w}_{i+1}) \). We now define an \( \text{Int} \)-model which is a countermodel to \( \varphi_{c}^{\text{Int}} \).

**DEFINITION 15.** Define \( \mathcal{M}_{S4\rightarrow \text{Int}} = (W_{S4\rightarrow \text{Int}}, R_{S4\rightarrow \text{Int}}, L_{S4\rightarrow \text{Int}}, r_{S4\rightarrow \text{Int}}) \), where

- \( W_{S4\rightarrow \text{Int}} = \frac{W_{S4}}{R_{S4}} \)
- \( R_{S4\rightarrow \text{Int}} = \overline{R_{S4}} \)
- \( L_{S4\rightarrow \text{Int}} = \frac{L_{S4}}{R_{S4}} \)
- \( \overline{w} \in L_{S4\rightarrow \text{Int}}(l_i) \) iff \( L_{\pi}(\overline{w}) \geq i \)
- \( \overline{w} \in L_{S4\rightarrow \text{Int}}(p_{m}^{i}) \) iff \( L_{\pi}(\overline{w}) = i \) and \( w_{n} \in L_{S4}(p) \), or \( L_{\pi}(\overline{w}) > i \) and the immediate predecessor of \( \overline{w} \) in \( R_{S4\rightarrow \text{Int}} \) is \( \overline{v} \) with \( \overline{v} \in L_{S4\rightarrow \text{Int}}(p_{m}^{i}) \)
- \( \overline{w} \in L_{S4\rightarrow \text{Int}}(b_{p}^{i}) \) iff \( L_{\pi}(\overline{w}) = i \) and \( \mathcal{M}_{S4}, w_{0} \models \Box \psi \), or \( L_{\pi}(\overline{w}) > i \) and the immediate predecessor of \( \overline{w} \) in \( R_{S4\rightarrow \text{Int}} \) is \( \overline{v} \) with \( \overline{v} \in L_{S4\rightarrow \text{Int}}(p_{m}^{i}) \).

Now we prove that \( \mathcal{M}_{S4\rightarrow \text{Int}} \) is in fact an \( \text{Int} \)-model, \( \mathcal{M}_{S4\rightarrow \text{Int}}, r_{S4\rightarrow \text{Int}} \models P(\varphi) \), and \( \mathcal{M}_{S4\rightarrow \text{Int}}, r_{S4\rightarrow \text{Int}} \not\models \varphi_{0} \).
Lemma 16. \( M^{S_4 \rightarrow \text{Int}} \) is an \( \text{Int} \)-model.

Proof. First, \( R^{S_4 \rightarrow \text{Int}} \) is transitive, reflexive, because \( R^{S_4} \) was, and it is anti-symmetric because clusters have been collapsed to their equivalence class. We must show that \( L^{S_4 \rightarrow \text{Int}} \) is persistent.

If \( R^{S_4 \rightarrow \text{Int}}(\overline{m}, \overline{n}) \), then \( L^i(\overline{m}) \leq L^i(\overline{n}) \) from the definition of \( L^i \). Thus if \( \overline{m} \in L^{S_4 \rightarrow \text{Int}}(b) \), then \( L^i(\overline{m}) \geq i \) and so \( \overline{n} \in L^{S_4 \rightarrow \text{Int}}(b) \) as required.

For the other propositions, the truth is defined inductively based on the truth at predecessors, so if \( \overline{m} \in L^{S_4 \rightarrow \text{Int}}(p_m) \) then any successor \( \overline{n} \) will also be in \( L^{S_4 \rightarrow \text{Int}}(p_m) \), as required. \( \Box \)

Lemma 17. For all subformulae \( \psi \) of \( \phi \), \( M^{S_4}, w_m \models \psi \) iff \( M^{S_4 \rightarrow \text{Int}}, \overline{m} \models (\psi)^{L^{\text{Int}}(w_m)} \).

Proof. Much of the proof is the same as for Lemma 11. The only difference is for \( \Box \)-formulae.

By the definition of \( L^{S_4 \rightarrow \text{Int}}(b^i_0) \) we have \( \overline{m} \in L(b^i_0) \) iff \( M^{S_4}, w_0 \models \Box \psi_1 \), and thus \( M^{S_4}, w_m \models \Box \psi_1 \) since \( w_0 \) and \( w_m \) must have the same set of successors. Therefore \( M^{S_4}, w_m \models \Box \psi_1 \) iff \( M^{S_4 \rightarrow \text{Int}}, \overline{m} \models (\Box \psi_1)^{L^{\text{Int}}(w_m)} \), as required. \( \Box \)

Lemma 18. We have \( M^{S_4 \rightarrow \text{Int}}, r^{S_4 \rightarrow \text{Int}} \models \vdash \phi \) in the constructed intuitionistic model.

Proof. From the definition of \( L^{S_4 \rightarrow \text{Int}} \) we obviously have \( M^{S_4 \rightarrow \text{Int}}, \overline{m} \models \Box \phi \) if and only if \( M^{S_4}, w_m \models \phi \), since \( L^i(\overline{m}) \geq 0 \). Also, because the models are \( M \)-standard, the maximum chain length is \( M - 1 \), thus \( L^i(\overline{m}) < \overline{m} \) and so \( M^{S_4 \rightarrow \text{Int}}, \overline{m} \models \Box \phi \).

Next, if \( L^i(\overline{m}) = i \) then \( \overline{m} \in L^{S_4 \rightarrow \text{Int}}(p_i) \) iff \( w_m \in L^{S_4}(p) \) for all atomic propositions \( p \). All successors \( \overline{n} \) of \( \overline{m} \) must have \( L^i(\overline{n}) > i \) and thus if \( \overline{m} \notin L^{S_4 \rightarrow \text{Int}}(p_m) \) then \( M^{S_4 \rightarrow \text{Int}}, \overline{m} \models \neg(p_m) \). Thus we have \( M^{S_4 \rightarrow \text{Int}}, \overline{m} \models \Box \phi \) and \( M^{S_4 \rightarrow \text{Int}}, \overline{m} \models \neg(p_m) \) for all \( n, m \) and \( p \). A similar argument applies to \( b^i_\psi \). Thus we have \( M^{S_4 \rightarrow \text{Int}} \models \Box \phi \) for all \( n \).

The base case of the definition of \( \overline{m} \in L^{S_4 \rightarrow \text{Int}}(b_0^i) \) requires that \( M^{S_4}, w_0 \models \Box \phi \) which is exactly when all \( R^{S_4} \) successors \( v_m \) of \( w_0 \) satisfy \( \phi \). Any such \( v_m \) will correspond to a \( \overline{v}_m \) with \( L^i(\overline{v}_m) \geq L^i(\overline{m}) \), and it will satisfy \( v_m^{L^i(\overline{v}_m)} \) due to Lemma 17. Thus if \( M^{S_4 \rightarrow \text{Int}}, \overline{m} \models b_0^i \), then all successors \( \overline{n} \) will satisfy \( k \rightarrow (v_m^k) \) for any \( k \geq n \) and any \( m \). Similarly, if \( \overline{m} \notin L^{S_4 \rightarrow \text{Int}}(b_0^i) \) then there must be some successor \( v_m \) of \( w_0 \) such that \( M^{S_4}, v_m \not\models \phi \) and thus \( M^{S_4 \rightarrow \text{Int}}, \overline{m} \not\models (\phi)^k \) for \( k = L^i(\overline{v}_m) \). Thus \( M^{S_4 \rightarrow \text{Int}} \models \Box \phi \) as required. \( \Box \)

Corollary 19. If \( M^{S_4}, r^{S_4} \not\models \phi \) then we have \( M^{S_4 \rightarrow \text{Int}}, r^{S_4 \rightarrow \text{Int}} \not\models \phi^c \).

Proof. Lemma 18 gives us \( M^{S_4 \rightarrow \text{Int}}, r^{S_4 \rightarrow \text{Int}} \models \vdash \phi \), and then because \( M^{S_4}, r^{S_4} \not\models \phi \) and \( L^{S_4}(\overline{m}) = 0 \), using Lemma 17 we have \( M^{S_4 \rightarrow \text{Int}}, r^{S_4 \rightarrow \text{Int}} \not\models (\phi)^c \). Thus we have \( M^{S_4 \rightarrow \text{Int}}, r^{S_4 \rightarrow \text{Int}} \not\models \phi^c \) as required. \( \Box \)

Main theorem We have shown that there is an \( S_4 \)-countermodel to \( \phi \) if and only if there is an \( \text{Int} \)-countermodel to \( \varphi_c^\text{Int} \), and thus \( \varphi_c^\text{Int} \) is a faithful translation from \( S_4 \) to \( \text{Int} \).
Remark 20. We might ask where Fernández [1] goes awry; where does the purported proof fail? The theorems and lemmas presented there appear to be correct, and yet Example [1] demonstrates that the original translation is wrong. The problem is with his application of his Theorem 4.1, which is our Lemma [11]. The theorem states that worlds in the constructed $S_4$-model satisfy the same formulae as the worlds in the original $\text{Int}$-model of the same level. The assumption that Fernández [1] makes is effectively that the original $\text{Int}$-models are stratified, and in particular that the root of the $\text{Int}$-countermodel has level 0. If this is the case, then applying his Theorem 4.1 will indeed result in an $S_4$-model where the root falsifies $\varphi$, because the $\text{Int}$-model falsifies $(\varphi)_0^0$. What we illustrated with Example [1] was that this assumption does not always hold for the original definition of $\varphi_o^{\text{Int}}$, and indeed because his Theorem 4.1 is correct the example cannot be “fixed” into a stratified model.

By changing the translation as we have, we are able to prove that all models of the modified translation can be converted into stratified models according to Corollary [9] and then Fernández’s original proofs only require slight changes to account for the changed translation to prove that this new translation is in fact correct. Our Lemmas [6] to [8] which we use to prove Corollary [9] are the bulk of the new work here, and they do not hold for the original translation.

An implementation of our translation is available at the URL below:

http://users.cecs.anu.edu.au/~rpg/S4ToInt/

There are also options to apply the original translation $\varphi_o^{\text{Int}}$ of Fernández [1], as well as that translation using our $M$ instead of $N$. Thus the reader can test that: our translation $\varphi_o^{\text{Int}}$ is correct; the original translation $\varphi_o^{\text{Int}}$ is incorrect; and that even changing $N$ to $M$ in the original translation is still incorrect.

Acknowledgements: We thank the reviewer for the detailed comments on our previous submission(s) which have improved the paper immensely.

REFERENCES