



# Embedding Intuitionistic into Classical Logic

Alexander Pluska and Florian Zuleger

TU Wien, Vienna, Austria

## Abstract

The famous double negation translation [16, 17] establishes an embedding of classical into intuitionistic logic. Curiously, the reverse direction has not been covered in literature. Utilizing a normal form for intuitionistic logic [20], we establish a small model property for intuitionistic propositional logic. We use this property for a direct encoding of the Kripke semantics into classical propositional logic and quantified Boolean formulae. Next, we transfer the developed techniques to the first order case and provide an embedding of intuitionistic first-order logic into classical first-order-logic. Our goal here is an encoding that facilitates the use of state-of-the-art provers for classical first-order logic for determining intuitionistic validity. In an experimental evaluation, we show that our approach can compete with state-of-the-art provers for certain classes of benchmarks, in particular when the intuitionistic content is low. We further note that our constructions support the transfer of counter-models to validity, which is a desired feature in model checking applications.

## 1 Introduction

Constructive mathematics refers to a flavor of mathematics in which the existence of an object can only be established by explicit construction, as opposed to classical mathematics where existence can be shown implicitly, e.g. by assuming non-existence and deriving a contradiction. The formalism usually associated with constructive mathematics is intuitionistic logic, which essentially differentiates itself from classical logic by the fact that the law of excluded middle  $A \vee \neg A$  and the double negation shift  $\forall x \neg \neg P(x) \rightarrow \neg \neg \forall x P(x)$  are not valid. Besides philosophical considerations, most prominently advocated by Brouwer [10] and Bishop [9], there is a particular motivation for studying constructive mathematics from the perspective of computer science in that proofs in intuitionistic logic directly correspond to computer programs — as expressed in the Curry–Howard correspondence [18].

The interest in intuitionistic logic has lead to the development of a number of automated theorem proving systems and a collection of benchmark problems (see e.g. the ILTP library website [3]). The progress in automated reasoning for intuitionistic logic, however, has been slower than the impressive advances in solvers for classic logics — evidenced, e.g., by the CASC [2] and SAT [1] competitions. This difference can partially be explained by fundamental differences between the logics. First of all, determining intuitionistic validity is computationally harder, e.g., in the propositional case intuitionistic validity is PSPACE-complete [26], whereas classical validity is coNP-complete [11]. A further advantage of classical logic is the existence

of calculi that are particularly suited for automation, such as superposition [6], which rely on the existence of convenient normal forms such as CNF, and the duality between validity and satisfiability (i.e., in order to show the validity of a formula it suffices to show the unsatisfiability of the negated formula, which is insufficient in intuitionistic logic). The first dedicated intuitionistic theorem provers [22, 27] used the naïve inverse method, i.e., a direct search for a cut-free proof by applying the rules from some proof calculus inversely, which generally leads to a very complex search. More recently, connection-based methods have been applied to various non-classical logics [23, 25], including intuitionistic logic. There have also been some successful attempts to study intuitionistic validity via embedding into higher-order classical logic [4]. We finally add that, in contrast to intuitionistic provers, a tremendous amount of work has been put into optimizing provers for classical logic.

With this work we want to leverage the progress in classical reasoning for intuitionistic logic. To this end, we study embeddings of intuitionistic into classical logic and propose the following methodology:

- Give for each  $\varphi$  a formula  $\varphi^\#$ , which is classically valid iff  $\varphi$  is intuitionistically valid.
- Determine the classical validity of  $\varphi^\#$  using a prover for classical logic.
- Translate the classical proof (resp. counter-model) of  $\varphi^\#$  given by the prover to a corresponding intuitionistic proof (resp. counter-model) of  $\varphi$ .

## Contributions

The following main contributions will be discussed in more detail in Sections 3 and 8:

- We establish a small model property for counter-models to validity in intuitionistic propositional logic.
- Based on the small model property, we provide an embedding of intuitionistic propositional logic into classical propositional logic and into quantified Boolean formulae.
- We discuss the reduction of counter-models in intuitionistic first-order logic and, based on this reduction, present an embedding of intuitionistic into classical first-order logic.
- We provide an experimental evaluation on the use of first-order provers for classical logic for determining intuitionistic validity. We show that our approach can compete with state-of-the-art provers for certain classes of benchmarks, in particular when the intuitionistic content is low.

## Related work

The famous double-negation translation establishes an embedding of classical into intuitionistic logic, going back to Glivenko [16] in the propositional case, and to Gödel [17] and Gentzen [15] in the first-order case. In the propositional case the embedding is particularly simple: A formula  $\varphi$  is classically valid if and only if  $\neg\neg\varphi$  is intuitionistically valid. Intuitively, the translation collapses for each subformula  $\psi$  of  $\varphi$  the truth values of  $\psi$  and  $\neg\neg\psi$ , which are classically but not intuitionistically equivalent. For embedding intuitionistic into classical logic, we need to do the opposite: We need to expand the truth values of  $\psi$  and  $\neg\neg\psi$ , i.e., if they both occur in  $\varphi$ , we must have a way to (classically) assign different truth values to their respective counterparts in  $\varphi^\#$ . This necessitates the introduction of new propositional variables in our procedure, which marks a big difference to the double negation translation.

We further mention some translations between logics from the literature. Most relevant are the embedding of intuitionistic logic into S4 [28] and of various non-classical logics into

higher-order classical logic [4, 29]. Of special relevance are the normalization procedures for intuitionistic propositional logic given in [20] and examined from a proof-theoretic perspective in [12], which we further expand upon in this paper.

We finally mention a long-standing line of research that has investigated when classical validity implies intuitionistic validity. This most famously holds for  $\Pi_2$ -formulae in Peano Arithmetic [13], but also other conditions are known [7]. In contrast, we propose a translation of an intuitionistic formula to an equivalent classical formula.

## 2 Preliminaries

In this section we fix notation and recall the semantics for classical and intuitionistic logic. Furthermore we recapitulate some results from [23].

### 2.1 Syntax

For propositional logic we assume an ambient set of propositional variables. As usual a formula is either  $\perp$ , a propositional variable or, given formulae  $\varphi$  and  $\psi$  and  $\circ \in \{\wedge, \vee, \rightarrow\}$ ,  $\varphi \circ \psi$ .

For predicate logic we assume an ambient signature  $\Sigma$ , i.e., a set of function and predicate symbols with their respective arities, and a set of variables. A term is either a variable or, given an  $n$ -ary function symbol  $f$  and terms  $t_1, \dots, t_n$ ,  $f(t_1, \dots, t_n)$ . A formula is either  $\perp$  or, given a  $n$ -ary predicate symbol  $P$  and terms  $t_1, \dots, t_n$ ,  $P(t_1, \dots, t_n)$  or, given formulae  $\varphi$  and  $\psi$  and  $\circ \in \{\wedge, \vee, \rightarrow\}$ ,  $\varphi \circ \psi$  or, given a formula  $\varphi$  and a variable  $x$ ,  $\exists x\varphi$  or  $\forall x\varphi$ . Free and bound variables are defined as usual and there is a Barendregt's convention in place, i.e. we implicitly assume that they comprise distinct sets, such that bound variables are not affected by substitutions of free variables in any formula.

### 2.2 Propositional Semantics

**Definition 2.1.** A *valuation*  $v$  maps each propositional variable  $A$  to a *truth value*  $v(A) \in \{0, 1\}$ . We inductively define the model relation  $\models$  between  $v$  and formulae:

- $v \not\models \perp$
- $v \models A$  iff  $v(A) = 1$  for each propositional variable  $A$ .
- $v \models \varphi \wedge \psi$  iff  $v \models \varphi$  and  $v \models \psi$ .
- $v \models \varphi \vee \psi$  iff  $v \models \varphi$  or  $v \models \psi$ .
- $v \models \varphi \rightarrow \psi$  iff  $v \not\models \varphi$  or  $v \models \psi$ .

A valuation  $v$  is a *model* for  $\varphi$  if  $v \models \varphi$ . If every valuation is a model for  $\varphi$  then we say  $\varphi$  is *valid*. We denote the set of valid formulae with *CPC* (Classical Propositional Calculus).

One notable property of classical logic is that a formula  $\varphi$  is valid if and only if its negation  $\neg\varphi := \varphi \rightarrow \perp$  is not satisfiable, i.e. there does not exist a model for it. The same does not hold true for intuitionistic logic as we shall see.

**Definition 2.2.** A *Kripke structure*  $\mathcal{K} = (\mathbf{W}, (v_w)_{w \in W})$  consists of a partially ordered set  $\mathbf{W} = (W, \leq)$  of *worlds*, called the *Kripke frame*, and an indexed family of valuations  $(v_w)_{w \in W}$  such that  $v_u(A) = 1$  entails  $v_w(A) = 1$  for all worlds  $u \leq w$  and propositional variables  $A$  (this is called the *persistence condition*). The model relation between the  $v_u$  and formulae  $\varphi$  is defined as before, except in the case of implications, where we set

- $v_u \models \varphi \rightarrow \psi$  iff for all  $w \geq u$  we have  $w \not\models \varphi$  or  $w \models \psi$ .

We say that  $\varphi$  is satisfied at a world  $u$  if  $v_u \models \varphi$  and write  $u \models \varphi$ . If a formula is satisfied at every world then  $\mathcal{K}$  is a *model* for  $\varphi$ . A formula is valid if every Kripke structure is a model of it. We denote the set of valid formulae with *IPC* (Intuitionistic Propositional Calculus).

There are many classical tautologies which are not intuitionistically valid, e.g. the law of excluded middle  $A \vee \neg A$ . On the other hand any intuitionistic theorem is a classical one.

## 2.3 Predicate Semantics

We now recall the semantics of first-order logic:

**Definition 2.3.** Let  $\Sigma$  be a signature. A  $\Sigma$ -structure  $\mathcal{M}$  consists of a non-empty set  $M$ , the *domain* of  $\mathcal{M}$ , and an *interpretation*  $I$  that assigns

- to each  $n$ -ary function symbol  $f \in \Sigma$  a  $n$ -ary function  $f^I : M^n \rightarrow M$ .
- to each  $n$ -ary predicate symbol  $R \in \Sigma$  a  $n$ -ary predicate  $R^I \subseteq M^n$ .

A variable assignment  $v$  is a function that assigns to each free variable an element  $m \in M$ . For each free variable  $a$  and  $m \in M$  we define

$$v[m/a](b) = \begin{cases} m, & \text{if } b = a, \\ v(b), & \text{otherwise.} \end{cases}$$

Then terms are interpreted as follows:

- $a^{I,v} = v(a)$  for each free variable  $a$ .
- $f(t_1, \dots, t_n)^{I,v} = f^I(t_1^{I,v}, \dots, t_n^{I,v})$  for each  $n$ -ary function symbol  $f \in \Sigma$ .

We define a model relation between pairs  $\mathcal{M}, v$  and formulae  $\varphi$  as follows

- $\mathcal{M}, v \not\models \perp$
- $\mathcal{M}, v \models R(t_1, \dots, t_n)$  iff  $(t_1^{I,v}, \dots, t_n^{I,v}) \in R^I$ ,  $\mathcal{M}, v \models s = t$  iff  $s^{I,v} = t^{I,v}$ .
- $\mathcal{M}, v \models \varphi \wedge \psi$  iff  $\mathcal{M}, v \models \varphi$  and  $\mathcal{M}, v \models \psi$
- $\mathcal{M}, v \models \varphi \vee \psi$  iff  $\mathcal{M}, v \models \varphi$  or  $\mathcal{M}, v \models \psi$ .
- $\mathcal{M}, v \models \varphi \rightarrow \psi$  iff  $\mathcal{M}, v \not\models \varphi$  or  $\mathcal{M}, v \models \psi$ .
- $\mathcal{M}, v \models \exists x \varphi$  iff there exists  $m \in M$  such that  $\mathcal{M}, v[m/x] \models \varphi$ .
- $\mathcal{M}, v \models \forall x \varphi$  iff for all  $m \in M$  we have  $\mathcal{M}, v[m/x] \models \varphi$ .

$\varphi$  is satisfiable if  $\mathcal{M}, v \models \varphi$  for some  $\mathcal{M}, v$ . It is valid if  $\mathcal{M}, v \models \varphi$  for every  $\mathcal{M}, v$ . We denote the set of valid formulae with *CQC* (Classical Quantified Calculus).

**Definition 2.4.** A  $\Sigma$ -Kripke structure  $\mathcal{K}$  is a partially ordered set  $(W, \leq)$ , called the Kripke frame, and an indexed family of  $\Sigma$ -structures  $(\mathcal{M}_w)_{w \in W}$  such that for  $u \leq w$  we have  $M_u \subseteq M_w$ ,  $f^{I_w}|_{M_u} = f^{I_u}$  and  $R^{I_w}|_{M_u} = R^{I_u}$ .<sup>1</sup> The model relation between  $\mathcal{M}_u, v$  (mapping into  $M_u$ ) and formulae  $\varphi$  is defined as before, except in the cases of implication and  $\forall$ -quantification, where

- $\mathcal{M}_u, v \models \varphi \rightarrow \psi$  iff for every  $w \geq u$  we have  $\mathcal{M}_w, v \not\models \varphi$  or  $\mathcal{M}_w, v \models \psi$ .
- $\mathcal{M}_u, v \models \forall x \varphi$  iff for every  $w \geq u$  and  $m \in M_w$  we have  $\mathcal{M}_w, v[m/x] \models \varphi$ .

We write  $u, v \models \varphi$  if  $\mathcal{M}_u, v \models \varphi$ .  $\mathcal{K}$  satisfies  $\varphi$  if  $u, v \models \varphi$  holds for every world  $u \in W$  and variable assignment  $v$ .  $\varphi$  is valid if it is satisfied by every Kripke structure. We denote the set of valid formulae with *IQC* (Intuitionistic Quantified Calculus).

In addition to propositional tautologies there are now quantified formulae which are classically valid but not intuitionistically, e.g.,  $\neg \forall x A(x) \rightarrow \exists x \neg A(x)$ .

## 2.4 Skolemization and Herbrandization

An important step in the embedding will be the elimination of quantifiers via Herbrandization. In this process we introduce fresh variables and add additional function symbols to the signature. A fresh variable is any variable that does not occur in any considered formula. Whenever we add a function symbol we implicitly extend the signature by a not previously contained symbol.

**Definition 2.5.** For formulae  $\varphi$  we define the Skolemization  $\varphi_Z^S$  and Herbrandization  $\varphi_Z^H$  with respect to a set of free variables  $Z$  by simultaneous induction:

- $A_Z^S = A_Z^H = A$  for each atomic  $A$ .
- $(\varphi \circ \psi)_Z^X = \varphi_Z^X \circ \psi_Z^X$  for  $\circ \in \{\wedge, \vee\}$ ,  $X \in \{S, H\}$ .
- $(\varphi \rightarrow \psi)_Z^S = \varphi_Z^H \rightarrow \psi_Z^S$
- $(\varphi \rightarrow \psi)_Z^H = \varphi_Z^S \rightarrow \psi_Z^H$ .
- $(\forall x \varphi)_Z^S = \forall x (\varphi_{Z \cup \{x\}}^S)$ .
- $(\forall x \varphi)_Z^H = \varphi[s(z_1, \dots, z_n)/x]_Z^H$  where  $s$  is a new function,  $\{z_1 \dots z_n\} = Z$ .
- $(\exists x \varphi)_Z^S = \varphi[s(z_1, \dots, z_n)/x]_Z^S$  where  $s$  is a new function,  $\{z_1 \dots z_n\} = Z$ .
- $(\exists x \varphi)_Z^H = \exists x (\varphi_{Z \cup \{x\}}^H)$ .

Let  $\varphi^S = (\exists x_1 \dots \exists x_n \varphi[x_1/a_1 \dots x_n/a_n])_0^S$  and  $\varphi^H = (\forall x_1 \dots \forall x_n \varphi[x_1/a_1 \dots x_n/a_n])_0^H$  where  $a_1, \dots, a_n$  are the free variables occurring in  $\varphi$ .

**Theorem 2.6.** For every formula  $\varphi$

- $\varphi$  and  $\varphi^S$  are classically equisatisfiable.
- $\varphi$  and  $\varphi^H$  are classically equivalent.

<sup>1</sup>Here  $f|_M$  denotes the restriction of  $f$  to  $M$

## 2.5 Intuitionistic normal forms

Finally we recall a result from [20], on which we build in this paper.

**Lemma 2.7.** For every propositional formula  $\varphi$  there exists an atom  $P$  as well as sets of clauses  $\mathcal{R}, \mathcal{X}$ , with  $\mathcal{R}$  containing *flat clauses* of the form

$$\bigwedge_i A_i \rightarrow \bigvee_j B_j$$

and  $\mathcal{X}$  containing *implication clauses* of the form

$$(A \rightarrow B) \rightarrow C,$$

such that  $\varphi$  is intuitionistically equivalent to

$$\left( \bigwedge \mathcal{R} \wedge \bigwedge \mathcal{X} \right) \rightarrow P$$

where  $A_i, B_i, A, B, C$  are atomic. In general these formulae will contain additional propositional variables. There is a constructive algorithm to obtain such a normal form with linear time complexity in the size of  $\varphi$ .

**Lemma 2.8.** For every predicate formula  $\varphi$  there exists a nullary predicate symbol  $P$  as well as sets of clauses  $\mathcal{R}, \mathcal{X}, \mathcal{Q}$ , with  $\mathcal{R}$  containing *flat clauses* of the form

$$\forall \vec{x} \left( \bigwedge_i A_i \rightarrow \bigvee_j B_j \right),$$

$\mathcal{X}$  containing *implication clauses* of the form

$$\forall \vec{x} ((A \rightarrow B) \rightarrow C)$$

and  $\mathcal{Q}$  containing *quantification clauses* of the form

$$\forall \vec{x} ((\forall y A) \rightarrow B),$$

such that  $\varphi$  is intuitionistically equivalent to

$$\left( \bigwedge \mathcal{R} \wedge \bigwedge \mathcal{X} \wedge \bigwedge \mathcal{Q} \right) \rightarrow P$$

where  $A_i, B_i, A, B, C$  are atomic. In general these formulae will contain new function and predicate symbols. There is a constructive algorithm to obtain such a normal form with linear time complexity in the size of  $\varphi$ .

Since we could not find a proof for the first-order case, we provide one in the appendix [A.1](#).

## 3 Overview

We shall now give an overview of our main results and highlight the key arguments. We start with the propositional case, as the important insights are the same as in the first-order case, while it is technically simpler and less cluttered.

The first step is to express the Kripke semantics of intuitionistic logic in classical first-order logic. That is, for each  $\varphi$  we give a  $\varphi^\#$  that is classically valid iff  $\varphi$  is intuitionistically valid, establishing a dual to the double negation translation.

Then, the key to reducing the first-order logic encoding to a propositional logic encoding is a small model property that supports the reduction of counter-models. In particular, given

$$\varphi = \left( \bigwedge \mathcal{R} \wedge \mathcal{X} \right) \rightarrow P$$

as in Lemma 2.7 we establish:

**Lemma 3.1.** If  $\varphi$  is intuitionistically invalid then there exists a Kripke counter-model which is a rooted tree with height and degree bounded by  $|\mathcal{X}|$ .

This allows us to embed intuitionistic propositional logic into classical propositional logic by replacing the quantification over all possible worlds by an explicit enumeration:

The details can be found in Section 5. Building on the same idea, we can also give a quantified Boolean formula (QBF) of polynomial size, where the branches of the tree-model can be enumerated implicitly using (alternating) quantifiers. We obtain the following result:

**Lemma 3.2.** For each propositional formula  $\varphi$  there is effectively a QBF  $\varphi^Q$  with  $|\varphi^Q| \in \mathcal{O}(|\mathcal{R}| \cdot |\mathcal{X}| + |\mathcal{X}|^3)$  such that  $\varphi$  is intuitionistically invalid if and only if  $\varphi^Q$  is a satisfiable QBF.

Again, the details can be found in Section 6. Next we attempt to transport these results to the first order case. However, there are formulae for which only counter-examples with an infinite Kripke frame exist. Therefore, there is no hope of completely eliminating terms representing worlds from the final embedding. This also means that our language necessarily becomes more complex. Still, a number of reductions are possible.

**Theorem 3.3.** There exists a linear-time procedure that gives for every first-order formula  $\varphi$  a formula  $\varphi^\#$  such that  $\varphi$  is intuitionistically valid if and only if  $\varphi^\#$  is classically valid. The size of  $\varphi^\#$  is in linear in the size of  $\varphi$ , however, for each  $n$ -ary predicate symbol in  $\varphi$  there is a corresponding  $n + 1$ -ary predicate symbol in  $\varphi^\#$ , and  $\varphi^\#$  contains a new binary predicate  $E$  as well as a number of new function symbols. Furthermore, there is an effective translation between intuitionistic counter-models of  $\varphi$  and classical counter-models of  $\varphi^\#$ .

A detailed account of the translation can be found in Definition 7.1. Finally, we benchmark our translation by translating all problems in the ILTP problem set [3] and applying the Vampire theorem prover [21] to the translated instances. While this method falls short of the absolute state-of-the-art in general, for certain subsets of benchmarks it beats existing provers, in particular when the intuitionistic content is low. Since there is still a lot of room for improvement this is a hopeful first sign.

## 4 Bounded Intuitionistic Counter-models

The most obvious approach to embedding intuitionistic into classical logic is to examine intuitionistic semantics, e.g. Kripke frames, as a classical first order theory. For every propositional variable  $A$  consider a unary predicate  $A$  of the same name where  $A(u)$  expresses that  $A$  is true at some world  $u$ . We can then consider the following naïve encoding:

**Definition 4.1.** Let  $\varphi$  be a propositional formula. Define  $\varphi^b$  inductively:

- $\perp^b = \perp$ .
- $A^b = A(b)$  for every propositional variable  $A$ .
- $(\varphi \circ \psi)^b = \varphi^b \circ \psi^b$  for  $\circ \in \{\wedge, \vee\}$ .
- $(\varphi \rightarrow \psi)^b = \forall w (b \preceq w \rightarrow \varphi^w \rightarrow \psi^w)$  where  $w$  is some new bound variable.

Let  $K(\varphi)$  encode the theory of Kripke structures, i.e.

$$K(\varphi) := \text{PartialOrder}(\preceq) \wedge \forall u \forall w (u \preceq w \rightarrow \text{Persistent}(u, w))$$

with e.g.

$$\text{PartialOrder}(\preceq) := \forall u (u \preceq u) \wedge \forall u \forall w (u \preceq w \rightarrow w \preceq u \rightarrow u = w) \wedge$$

$$\forall u \forall v \forall w (u \preceq v \rightarrow v \preceq w \rightarrow u \preceq w)$$

$$\text{Persistent}(u, w) := \bigwedge \{A(u) \rightarrow A(w) \mid A \text{ occurs in } \varphi\}$$

Then define

$$\varphi^C := K(\varphi) \rightarrow \varphi^b$$

where  $b$  is a new free variable.

All we have done is model Kripke semantics as a first-order theory, so we obtain:

**Lemma 4.2.**  $\varphi$  is intuitionistically valid if and only if  $\varphi^C$  is classically valid.

This is somewhat unsatisfying as we end up in a much more complex logic. However using the normal form from Lemma 2.7, i.e. assuming that  $\varphi$  is of the form

$$\left( \bigwedge \mathcal{R} \wedge \bigwedge \mathcal{X} \right) \rightarrow P$$

for some flat clauses  $\mathcal{R}$  and implicational clauses  $\mathcal{X}$ , further simplification is possible:

- For flat clauses  $(\bigwedge A_i \rightarrow \bigvee B_j)^b$  is

$$\forall w \left( b \preceq w \rightarrow \bigwedge_i A_i(w) \rightarrow \bigvee_j B_j(w) \right).$$

- For implication clauses  $((A \rightarrow B) \rightarrow C)^b$  is

$$\forall w (b \preceq w \rightarrow (\forall k (w \preceq k \rightarrow A(k) \rightarrow B(k))) \rightarrow C(w))$$

and after applying Herbrandization we end up with a formula

$$\varphi^{CH} = K(\varphi) \rightarrow \left( \bigwedge \mathcal{R}^{bS} \wedge \bigwedge \mathcal{X}^{bS} \right) \rightarrow P(s)$$

where  $s$  is a new constant symbol, each formula in  $\mathcal{R}^{bS}$  is of the form

$$\forall w \left( s \preceq w \rightarrow \bigwedge_i A_i(w) \rightarrow \bigvee_j B_j(w) \right)$$

and each formula  $\psi \in \mathcal{X}^{bS}$  is of the form

$$\forall w (s \preceq w \rightarrow (w \preceq f_\psi(w) \rightarrow A(f_\psi(w)) \rightarrow B(f_\psi(w))) \rightarrow C(w))$$

with  $f_\psi$  being a new function symbol.



**Definition 4.3.** For  $\psi \in \mathcal{X}$  let  $[\psi]$  denote the corresponding clause in  $\mathcal{X}^{bS}$ . Suppose we are given a counter-model  $\mathcal{M} = (M, I)$  for  $\varphi^{CH}$ . We say  $(A \rightarrow B) \rightarrow C \in \mathcal{X}$  is *fulfilled* at  $m \in M$  iff  $A^I(m) \rightarrow B^I(m)$  is false or  $C^I(m)$  is true. For  $\psi \in \mathcal{X}$  define

$$g_\psi : M \rightarrow M, m \mapsto \begin{cases} m, & \text{if } \psi \text{ is fulfilled at } m \text{ or } m \not\leq f_{[\psi]}^I(m), \\ f_{[\psi]}^I(m), & \text{else.} \end{cases}$$

Define a model  $\mathcal{M}_T = (M_T, I_T)$  as follows:

- $M_T$  is the set of sequences without repetition on  $\mathcal{X}$ .
- $s^{I_T} = \epsilon$ .
- Interpret  $\preceq$  as the prefix-order.
- We set

$$f_\psi^{I_T}(\psi_1 \dots \psi_n) = \begin{cases} \psi_1 \dots \psi_n, & \text{if } \psi \text{ occurs in } [\psi_1] \dots [\psi_n], \\ \psi_1 \dots \psi_n \psi, & \text{else.} \end{cases}$$

- For propositional variables  $P$ , we set

$$P^{I_T}(\psi_1 \dots \psi_n) = P^I(g_{\psi_n}(\dots(g_{\psi_1}(s^I))\dots)).$$

**Lemma 4.4.** Suppose  $\mathcal{M} = (M, T)$  is a counter-model to  $\varphi^{CH}$ , then so is  $\mathcal{M}_T = (M_T, I_T)$ .

A proof can be found in Appendix [A.2](#).

**Remark 4.5.** It is well known that Kripke frames which are trees are complete with regard to intuitionistic logic. Our contribution is giving an explicit bound to height and degree that depends only on a simple property of the considered formula.

**Corollary 4.6** (Small model property). If a formula in normal form as in Lemma [2.7](#) is intuitionistically invalid then there exists a Kripke counter-model which is a rooted tree of height and degree bounded by  $|\mathcal{X}|$ .

## 5 IPC to CPC

We can apply the previous results to give a direct embedding from intuitionistic into classical propositional logic. We know that  $\varphi^{CH}$  is valid if and only if it is valid for structures that have as a domain the sequences without repetition over  $\mathcal{X}$ , ordered by the prefix-relation and

$$s^I = \epsilon, f_\psi^I(x) = \begin{cases} x, & \text{if } \psi \text{ occurs in } x, \\ x\psi, & \text{otherwise.} \end{cases}$$

Now we can replace every  $\forall$ -quantifier in  $\varphi^{CH}$  by enumerating all the ground terms, and then replace all distinct ground instances of relations by new propositional variables yielding a propositional formula. Note that the encoding of the partial order becomes redundant, so we drop it. We need to keep the persistency constraints. This gives us the following theorem:

**Theorem 5.1.** Suppose we are given a formula as in Lemma 2.7, i.e. of the form

$$\left( \bigwedge \mathcal{R} \wedge \bigwedge \mathcal{S} \right) \rightarrow P$$

with flat  $\mathcal{R}$  and implicational  $\mathcal{S}$ . Let  $\Lambda$  denote the set of sequences without repetition over  $\mathcal{X}$ . For every propositional variable  $A$  and  $\lambda \in \Lambda$  consider a new propositional variable  $A^\lambda$ . Define  $\mathcal{R}^\#$  as the union over all sets

$$\left\{ \bigwedge_i A_i^\lambda \rightarrow \bigvee_j B_j^\lambda \mid \lambda \in \Lambda \right\}$$

for  $\bigwedge_i A_i \rightarrow \bigvee_j B_j \in \mathcal{R}$  and  $\mathcal{X}^\#$  as the union over all sets

$$\{(A^{\lambda\psi} \rightarrow B^{\lambda\psi}) \rightarrow C^\lambda \mid \lambda \in \Lambda, \psi \text{ does not occur in } \lambda\}$$

for  $\psi = (A \rightarrow B) \rightarrow C \in \mathcal{X}$  as well as  $\mathcal{P}^\#$  as the union over all sets

$$\{A^\lambda \rightarrow B^{\lambda\psi} \mid \lambda \in \Lambda, \psi \text{ does not occur in } \lambda\}$$

for each propositional variable  $A$  occurring in  $\varphi$ , encoding persistency. Then

$$\varphi^\# := \left( \bigwedge \mathcal{R}^\# \wedge \bigwedge \mathcal{X}^\# \wedge \bigwedge \mathcal{P}^\# \right) \rightarrow P^\epsilon$$

is classically valid if and only if  $\varphi$  is intuitionistically valid.

**Example 5.2.** Let us consider the law of excluded middle  $\varphi = A \vee \neg A$  for which there is the equivalid normal form

$$((A \rightarrow P) \wedge ((A \rightarrow \perp) \rightarrow P)) \rightarrow P.$$

Applying the above lemma this is intuitionistically valid if and only if

$$((A^\epsilon \rightarrow A^\psi) \wedge (P^\epsilon \rightarrow P^\psi) \wedge (A^\epsilon \rightarrow P^\epsilon) \wedge (A^\psi \rightarrow P^\psi) \wedge ((A^\psi \rightarrow \perp) \rightarrow P^\epsilon)) \rightarrow P^\epsilon$$

is classically valid where  $\psi := (A \rightarrow \perp) \rightarrow P$ . We note that  $A^\epsilon = P^\epsilon = 0$ ,  $A^\psi = P^\psi = 1$  indeed defines a counter-model.

Note that it is straightforward to obtain an intuitionistic counter-model to  $\varphi$  with frame  $\Lambda$  from a classical counter-model to  $\varphi^\#$ .

Since  $|\Lambda| \in \mathcal{O}(2^{|\mathcal{X}| \log |\mathcal{X}|})$  a simple counting argument shows that  $|\mathcal{P}^\#| \in \mathcal{O}(|\varphi| \cdot 2^{|\mathcal{X}| \log |\mathcal{X}|})$ ,  $|\mathcal{R}^\#| \in \mathcal{O}(|\mathcal{R}| \cdot 2^{|\mathcal{X}| \log |\mathcal{X}|})$  and  $|\mathcal{X}^\#| \in \mathcal{O}(|\mathcal{X}| \cdot 2^{|\mathcal{X}| \log |\mathcal{X}|})$ . This gives a bound on the size of  $\varphi^\#$ .

**Lemma 5.3.** The size of  $\varphi^\#$  is in  $\mathcal{O}(|\varphi| \cdot 2^{|\mathcal{X}| \log |\mathcal{X}|})$ .

## 6 IPC to QBF

It is known that a polynomial time translation between QBF and IPC must exist, since both problems are PSPACE complete [14, 26], but we make it explicit. In particular, our translation enables a direct translation of counter-models. For lack of space, we do not give the semantics of QBF in this paper and instead refer the reader to standard sources e.g. [8]. The translation from the last section will serve as a starting point for this section.

Suppose we are given a formula  $\varphi$  in normal form

$$\left( \bigwedge \mathcal{R} \wedge \bigwedge \mathcal{X} \right) \rightarrow P$$

as in Lemma 2.7. We now give an intuition on how to obtain a QBF of polynomial size that expresses that  $\varphi$  has an intuitionistic counter-model: Instead of expressing validity at every node of  $\mathcal{M}_T$  — which was done in the CPC translation — we express that on each path in  $\mathcal{M}_T$  the nodes satisfy all conditions for  $\mathcal{M}_T$  being a counter-model. The universally quantified variables  $X_\psi$  in the next definition handle the branching, i.e. express which path is considered.

**Definition 6.1.** For every propositional variable  $A$  occurring in  $\varphi$  and non-negative integer  $n$  (representing the  $n$ -th position of a node on a path) consider a new propositional variable  $A^n$  and for every formula  $\psi \in \mathcal{X}$  consider a new propositional variable  $X_\psi^n$ . Let  $\vec{A}^n$  range over all propositional variables  $A^n$  and  $\vec{X}^n$  range over all  $X_\psi^n$ . Define

- $\text{Valid}(n)$  which encodes that  $\vec{X}_\psi^n$  represents a valid next element of a sequence without repetition, i.e. if we interpret the  $\vec{X}^i$  as bit-vectors, then  $\vec{X}^n$  has exactly one bit set to 1, indicating the formula  $\psi$  at the  $n$ -th position of the sequence, and it is at a position that is 0 in all  $\vec{X}^i$  with  $i < n$  (i.e., it did not occur in the sequence previously).
- $\text{Persistent}(n)$  which encodes that the persistency condition holds.
- $\text{Sat}_{\mathcal{R}}(n)$  encoding that the formulae in  $\mathcal{R}$  hold at the  $n$ -th world of the path.
- $\text{Sat}_{\mathcal{X}}(n)$  encoding that the formula in  $\mathcal{X}$  that is represented by  $\vec{X}^{n-1}$  holds.

Define

$$\varphi_i^Q := \exists \vec{A}^i \forall \vec{X}^i \left( \text{Persistent}(i) \wedge \text{Sat}_{\mathcal{R}}(i) \wedge \text{Sat}_{\mathcal{X}}(i) \wedge \left( \text{Valid}(i) \rightarrow \varphi_{i+1}^Q \right) \right)$$

for  $0 < i < |\mathcal{X}|$ , as well as the special cases

$$\varphi_{|\mathcal{X}|}^Q := \exists \vec{A}^{|\mathcal{X}|} \left( \text{Persistent}(|\mathcal{X}|) \wedge \text{Sat}_{\mathcal{R}}(|\mathcal{X}|) \wedge \text{Sat}_{\mathcal{X}}(|\mathcal{X}|) \right)$$

for leaves, and

$$\varphi^Q = \varphi_0^Q := \exists \vec{A}^0 \forall \vec{X}^0 \left( \neg P^0 \wedge \text{Sat}_{\mathcal{R}}(0) \wedge \left( \text{Valid}(0) \rightarrow \varphi_1^Q \right) \right)$$

for the root. Example encodings of the above formulae are:

$$\begin{aligned} \text{Valid}(n) := & \left( \bigvee \{ X_\psi^n \mid \psi \in \mathcal{X} \} \right) \wedge \left( \bigwedge \{ \neg (X_{\psi_1}^n \wedge X_{\psi_2}^n) \mid \psi_1 \neq \psi_2 \in \mathcal{X} \} \right) \wedge \\ & \left( \bigwedge \{ \bigwedge \{ X_\psi^i \rightarrow \neg X_\psi^n \mid i < n \} \mid \psi \in \mathcal{X} \} \right) \end{aligned}$$

$$\text{Persistent}(n) := \bigwedge \left\{ A^{n-1} \rightarrow A^n \mid A \text{ prop. variable with } A^{n-1} \in \vec{A}^{n-1}, A^n \in \vec{A}^n \right\}$$

$$\text{Sat}_{\mathcal{R}}(n) := \bigwedge \left\{ \bigwedge A_i^n \rightarrow \bigvee B_j^n \mid \bigwedge A_i \rightarrow \bigvee B_j \in \mathcal{R} \right\}$$

$$\text{Sat}_{\mathcal{X}}(n) := \bigwedge \left\{ X_\psi^{n-1} \rightarrow (A^n \rightarrow B^n) \rightarrow C^{n-1} \mid \psi = (A \rightarrow B) \rightarrow C \in \mathcal{X} \right\}$$

**Example 6.2.** With the previous encoding for double negation elimination  $\varphi = ((A \rightarrow \perp) \rightarrow \perp) \rightarrow A$  we have

$$\varphi^Q = \exists A^0 \forall X^0 \left( \neg A^0 \wedge (X^0 \rightarrow \exists A^1 \left( (A^0 \rightarrow A^1) \wedge (X_0 \rightarrow (A^1 \rightarrow \perp) \rightarrow \perp) \right)) \right),$$

which is a satisfiable QBF since  $A^0 = 0, A^1 = 1$  satisfies it for any choice of  $X^0$ .

**Lemma 6.3.**  $\varphi$  is not intuitionistically valid if and only if  $\varphi^Q$  is a satisfiable QBF.

A proof can be found in the appendix as Lemma A.3.

A simple counting argument shows that for each  $n < |\mathcal{X}|$  we have  $|\text{Valid}(n)| \in \mathcal{O}(|\mathcal{X}|^2)$ ,  $|\text{Persistent}(n)|, |\text{Sat}_{\mathcal{R}}(n)| \in \mathcal{O}(|\mathcal{R}|)$  and  $|\text{Sat}_{\mathcal{X}}(n)| \in \mathcal{O}(|\mathcal{X}|)$  so overall we get:

**Lemma 6.4.** The size of  $\varphi^Q$  is in  $\mathcal{O}(|\mathcal{R}| \cdot |\mathcal{X}| + |\mathcal{X}|^3)$ .

Note that the last set of universally quantified variables can be avoided because there is only one assignment that has the chance to falsify  $\varphi^Q$ , namely the assignment that assigns 1 to the single  $X_{\psi}^n$  such that  $X_{\psi}^i = 0$  for all  $i < n$ . Hence, in our final formula we can replace every  $X_{\psi}^n$  with  $\bigwedge\{\neg X_{\psi}^i \mid i < n\}$  and remove that quantification over  $X_{\psi}^n$ . With that we get the following:

**Corollary 6.5.** Let  $N\text{-IntInvalid}$  be the problem of deciding if a formula  $(\bigwedge \mathcal{R} \wedge \bigwedge \mathcal{X}) \rightarrow P$  as in Lemma 2.7 with  $|\mathcal{X}| \leq N$  is not intuitionistically valid. Then  $N\text{-IntInvalid}$  is in  $\Sigma_{2N-1}^P$ . The dual problem  $N\text{-IntValid}$  is in  $\Pi_{2N-1}^P$ .

## 7 CQC to IQC

We now give an analogous transformation for first order logic. Our trick of lifting the sentence to first-order logic to encode the Kripke semantics no longer works in such a straightforward manner, since we already are in first-order logic. However we do it nonetheless! The domain of our classical model will feature on the one hand elements representing worlds in a Kripke frame and on the other hand the domain-elements from the Kripke model. We reconcile these notions by introducing a special binary predicate  $E$ , first considered in [5] and expanded onto in [19], where  $E(x, u)$  encodes that  $x$  is an element of the domain of the world  $u$ . To encode that some  $n$ -ary predicate  $A$  holds at a world  $u$  we extend each  $n$ -ary predicate symbol to a  $n+1$ -ary predicate symbol of the same name, interpreting  $A(\vec{x}, u)$  as " $A(\vec{x})$  holds at  $u$ ". Write  $\vec{E}(\vec{x}, u)$  for  $\bigwedge\{E(x_i, u) \mid x_i \in \vec{x}\}$ . With this we obtain the following encoding of Kripke Semantics:

**Definition 7.1.** Let  $\varphi$  be a predicate formula. Define  $\varphi^b$  inductively:

- $\perp^b = \perp$ .
- $A(\vec{t})^b = A(\vec{t}, b)$  for every predicate symbol  $A$ .
- $(\varphi \circ \psi)^b = \varphi^b \circ \psi^b$  for  $\circ \in \{\wedge, \vee\}$ .
- $(\varphi \rightarrow \psi)^b = \forall w(b \preceq w \rightarrow \varphi^w \rightarrow \psi^w)$  where  $w$  is a new bound variable.
- $(\exists x \varphi)^b = \exists x(E(x, b) \wedge \varphi^b)$ .
- $(\forall x \varphi)^b = \forall w(b \preceq w \rightarrow \forall x(E(x, w) \rightarrow \varphi^w))$  where  $w$  is a new bound variable.

Let  $K(\varphi)$  encode the theory of Kripke structures as well as the predicate  $E$ , i.e.

$$K(\varphi) := \text{PartialOrder}(\preceq) \wedge \forall u \forall w (u \preceq w \rightarrow \text{DomainSubset}(u, w)) \wedge \\ \forall u (\text{World}(u) \rightarrow \text{DomainClosed}(u)) \wedge \forall u \forall w (u \preceq w \rightarrow \text{Persistent}(u, w)).$$

with e.g.:

$$\begin{aligned}
 \text{PartialOrder}(\preceq) &:= \forall u(u \preceq u) \wedge \forall u \forall w(u \preceq w \rightarrow w \preceq u \rightarrow u = w) \wedge \\
 &\quad \forall u \forall v \forall w(u \preceq v \rightarrow v \preceq w \rightarrow u \preceq w) \\
 \text{DomainSubset}(u, w) &:= \forall x(E(x, u) \rightarrow E(x, w)) \\
 \text{World}(u) &:= \exists x E(x, u) \\
 \text{DomainClosed}(u) &:= \bigwedge \left\{ \forall \vec{x}(\vec{E}(\vec{x}, u) \rightarrow E(f(\vec{x}), u)) \mid f \text{ a function in } \varphi \right\} \\
 \text{Persistent}(u, w) &:= \bigwedge \left\{ \forall \vec{x}(\vec{A}(\vec{x}, u) \rightarrow A(\vec{x}, w)) \mid A \text{ a predicate in } \varphi \right\}
 \end{aligned}$$

Then define

$$\varphi^C := K(\varphi) \rightarrow \text{World}(b) \rightarrow \vec{E}(\vec{a}, b) \rightarrow \varphi^b$$

where  $\vec{a}$  contains all free variables occurring in  $\varphi$  and  $b$  is a new free variable.

**Lemma 7.2.**  $\varphi$  is intuitionistically valid if and only if  $\varphi^C$  is classically valid.

A proof can be found in the appendix A.4.

Now again due to Lemma 2.8 it is sufficient to consider formulae  $\varphi$  of the form

$$\left( \bigwedge \mathcal{R} \wedge \bigwedge \mathcal{X} \wedge \bigwedge \mathcal{Q} \right) \rightarrow P$$

with flat clauses  $\mathcal{R}$ , implication clauses  $\mathcal{X}$  and quantification clauses  $\mathcal{Q}$ , which after Herbrandization leaves us with

$$\varphi^{CH} = K(\varphi) \rightarrow \text{World}(s) \rightarrow \vec{E}(\vec{r}, s) \rightarrow \left( \bigwedge R^{bS} \wedge \bigwedge X^{bS} \wedge \bigwedge Q^{bS} \right) \rightarrow P(s)$$

where  $\vec{r}, s$  are new constant symbols, each formula in  $\mathcal{R}^{bS}$  is of the form

$$\forall w \forall \vec{x}(s \preceq w \rightarrow \vec{E}(x, w) \rightarrow \bigwedge_i A_i(\vec{a}_i, u) \rightarrow \bigvee_j B_j(\vec{b}_j, u)),$$

each formula  $\psi \in \mathcal{X}^{bS}$  is of the form

$$\forall w \forall \vec{x}(s \preceq w \rightarrow \vec{E}(\vec{x}, w) \rightarrow (w \preceq f_\psi(\vec{x}, w) \rightarrow A(\vec{a}, f_\psi(\vec{x}, w)) \rightarrow B(\vec{b}, f_\psi(\vec{x}, w))) \rightarrow C(\vec{c}, w))$$

and each formula  $\psi \in \mathcal{Q}^{bS}$  is of the form

$$\forall w \forall \vec{x}(s \preceq w \rightarrow \vec{E}(\vec{x}, w) \rightarrow \forall y(E(y, f_\psi(\vec{x}, w)) \rightarrow A(\vec{a}, f_\psi(\vec{x}, w))) \rightarrow B(\vec{b}, w))$$

with  $f_\psi$  being a new function symbol.

As in the previous section we can define a tree counter-model for this.

**Definition 7.3.** Suppose we are given a counter-model  $\mathcal{M} = (M, I)$  for  $\varphi^{CH}$ . Similarly to the propositional case for  $\psi \in \mathcal{X}$  define

$$g_\psi : M^n \rightarrow M, (\vec{z}, m) \mapsto \begin{cases} m, & \text{if } m \not\leq f_{[\psi]}^I(\vec{z}, m) \\ f_{[\psi]}^I(\vec{z}, m) & \text{otherwise} \end{cases}$$

where  $[\psi]$  denotes the corresponding element of  $\mathcal{X}^{bS}$ . Define  $\mathcal{M}_T = (M_T, I_T)$  as follows:

- $M_T$  is the union of  $\{m \in M \mid \neg \text{World}^I(m)\}$  and sequences  $(\vec{m}_1, \psi_1) \dots (\vec{m}_n, \psi_n)$  such that  $\vec{E}^I(\vec{m}_1, s^I)$  and  $\vec{E}^I(\vec{m}_i, g_{\psi_{i-1}}(\vec{m}_{i-1}, g_{\psi_{i-2}}(\dots g_{\psi_1}(\vec{m}_1, s^I) \dots)))$  for  $i > 0$  hold.
- $s^{I_T} = s^I, \vec{r}^{I_T} = \vec{r}^I$ .
- We interpret  $\preceq$  as the prefix-order on the set of sequences and trivially elsewhere.
- Have  $E^{I_T}(m, (\vec{m}_1, \psi_1) \dots (\vec{m}_n, \psi_n))$  iff  $E^I(m, g_{\psi_n}(\vec{m}_n, \dots g_{\psi_1}(\vec{m}_1, s^I) \dots))$ .
- We set  $f_{\psi}^{I_T}(\vec{m}, (\vec{m}_1, \psi_1) \dots (\vec{m}_n, \psi_n)) = (\vec{m}_1, \psi_1) \dots (\vec{m}_n, \psi_n)(\vec{m}, \psi)$  and arbitrary else.
- For predicate symbols  $R$ , we set

$$R^{I_T}(\vec{m}, (\vec{m}_1, \psi_1) \dots (\vec{m}_n, \psi_n)) = R^I(\vec{m}, g_{\psi_n}(\vec{m}_n, \dots (g_{\psi_1}(\vec{m}_1, s^I) \dots)))$$

and arbitrarily elsewhere, for function symbols  $f$  we set  $f^{I_T}(\vec{m}) = f^I(\vec{m})$  so long as  $\vec{m} \in \{m \in M \mid \neg \text{World}^I(m)\}^n$  and arbitrarily elsewhere.

**Lemma 7.4.** If  $\mathcal{M}$  is a counter-model to  $\varphi^{CH}$  then so is  $\mathcal{M}_T$ .

Even if we do not have the finiteness property of the propositional case, we have still managed to reduce the complexity of possible counter-models because we know that considering worlds of the form

$$g_{\psi_1}(\vec{m}_1, g_{\psi_2}(\vec{m}_2, \dots g_{\psi_n}(\vec{m}_n, s) \dots))$$

is sufficient. In particular, this allows us to completely eliminate  $\preceq$ , that is we remove the  $\text{PartialOrder}(\preceq)$  from  $K(\varphi)$  and adjust  $\text{DomainSubset}(u, w)$  and  $\text{Persistent}(u, w)$ , e.g. remove quantification over  $w$  and the precondition  $u \preceq w$  from  $K(\varphi)$  and replace  $\text{DomainSubset}(u, w)$  with

$$\bigwedge \{ \forall u \forall x \forall \vec{y} (E(x, u) \rightarrow E(x, f_{\psi}(\vec{y}, u))) \mid \psi \in \mathcal{X} \cup \mathcal{Q} \}$$

and  $\text{Persistent}(u, w)$  with

$$\bigwedge \{ \forall \vec{x} \forall \vec{y} \forall u (A(\vec{x}, u) \rightarrow A(\vec{x}, f_{\psi}(\vec{y}, u))) \mid \psi \in \mathcal{X} \cup \mathcal{Q}, A \text{ a predicate in } \varphi \}.$$

Putting all of this together yields the following translation:

**Theorem 7.5.** Let  $\mathcal{R}, \mathcal{X}, \mathcal{Q}$  be as above. For each  $n$ -ary predicate  $A$  define a new  $n + 1$ -ary predicate of the same name. For every  $\psi \in \mathcal{X} \cup \mathcal{Q}$  define a new function symbol  $f_{\psi}$ . Consider a new binary predicate  $E$  and define  $\vec{E}(\vec{x}, u) := \bigwedge \{ E(x, u) \mid x \in \vec{x} \}$ . Obtain  $\mathcal{R}^{\#}$  by including

$$\forall u \forall \vec{x} \left( \vec{E}(\vec{x}, u) \rightarrow \bigwedge_i A_i(\vec{a}_i, u) \rightarrow \bigvee_j B_j(\vec{b}_j, u) \right)$$

for each flat clause  $\forall \vec{x} \left( \bigwedge_i A_i(\vec{a}_i) \rightarrow \bigvee_j B_j(\vec{b}_j) \right) \in \mathcal{R}, \mathcal{X}^{\#}$  by including

$$\forall u \forall \vec{x} \left( A(\vec{a}, f_{\psi}(\vec{x}, u)) \rightarrow B(\vec{b}, f_{\psi}(\vec{x}, u)) \rightarrow C(\vec{c}, u) \right)$$

for each implication clause  $\psi = \forall \vec{x} \left( \left( A(\vec{a}) \rightarrow B(\vec{b}) \right) \rightarrow C(\vec{c}) \right) \in \mathcal{X}$ , and  $\mathcal{Q}^{\#}$  by including

$$\forall u \forall \vec{x} \left( \forall y (A(\vec{a}, f_{\psi}(\vec{x}, u))) \rightarrow B(\vec{b}, u) \right)$$

for each implication clause  $\psi = \forall \vec{x} (\forall y (A(\vec{a})) \rightarrow B(\vec{b})) \in \mathcal{X}$ . Then  $\varphi$  is intuitionistically valid if and only if

$$\varphi^\# := K(\varphi) \rightarrow \text{World}(s) \rightarrow \vec{E}(\vec{r}, s) \rightarrow \left( \bigwedge \mathcal{R}^\# \wedge \bigwedge \mathcal{X}^\# \wedge \bigwedge \mathcal{Q}^\# \right) \rightarrow P^\#(s)$$

is classically valid, where  $s, \vec{r}$  are new constants introduced by Herbrandization of free variables,

$$\begin{aligned} K(\varphi) := & \bigwedge \{ \forall \vec{z} \forall x \forall u ((E(x, u) \rightarrow E(x, f_\psi(\vec{z}, u))) \mid \psi \in \mathcal{X} \cup \mathcal{Q} \} \wedge \\ & \forall u \left( \exists x E(x, u) \rightarrow \bigwedge \{ \forall \vec{z} (\vec{E}(\vec{z}, u) \rightarrow E(f(\vec{z}, u))) \mid f \text{ a function in } \varphi \} \right) \wedge \\ & \bigwedge \{ \forall \vec{z}_1 \forall \vec{z}_2 \forall u (A(\vec{z}_1, u) \rightarrow A(\vec{z}_1, f_\psi(\vec{z}_2, u))) \mid \psi \in \mathcal{X} \cup \mathcal{Q}, A \text{ a predicate in } \varphi \}. \end{aligned}$$

The size of the obtained formula is linear in the input. However  $|\mathcal{X}| + |\mathcal{Q}| + 1$  new function symbols and the new binary predicate  $E$  were introduced and each  $n$ -ary predicate has been extended to a  $n + 1$ -ary predicate.

It is not possible to completely eliminate worlds from the domain by enumerating them since there are formulae with no finite counter-examples. An analogous procedure to the propositional case fails, however we can still retrieve additional axioms that restrict the interpretation of functions. That is, we may assume that  $f_\psi(\vec{t}, w) = w$ , if  $\exists u (w = f_\psi(\vec{t}, u))$  for all  $\psi \in \mathcal{X} \cup \mathcal{Q}$  (i.e. that  $f_\psi$  parametrized by  $\vec{t}$  is idempotent), and  $f_\psi(\vec{t}, w) = w$ , if  $\exists u (w \geq f_\psi(\vec{t}, u))$  for all  $\psi \in \mathcal{X}$ .

## 8 Experimental Evaluation

We have benchmarked our first-order translation on version 1.1.2 of the ILTP problem set [3] by performing our translation and then examining the generated classical problem using Vampire [21]. For comparison the state-of-the-art intuitionistic provers ileanCoP 1.2 [24] and nanoCoP-i 2.0 [25] were used. All benchmarks were conducted on an Intel Core i5-8400 2.80GHz machine running Linux with a time limit of 10s per problem. The problem SYN007+1 was removed as our normal form procedure would explode the size of the instance.

Two versions of the translation were tested, one with and one without the additional axioms encoding idempotence. The results are stated in Figure 1.

In general, our approach falls short of the state of the art. However when grouping the benchmarks by  $|\mathcal{Q}|$  and  $|\mathcal{X}|$ , see Figure 2, we observe good performance of our approach for  $|\mathcal{Q}| = 0$  and  $|\mathcal{X}| \leq 2$ . Also, there were some instances with high  $|\mathcal{Q}|$  and low  $|\mathcal{X}|$  that were solved by our approach but not the previous solvers.

$ \mathcal{Q} $	$ \mathcal{X} $	Embedding		Embedding (idem.)		ileanCoP		nanoCoP-i 2.0		Total
0	0	157	88.2%	157	88.2%	128	71.9%	61	38.1%	178
0	1-2	119	29.8%	130	32.5%	108	27.0%	96	24.0%	400
0	3+	33	18.9%	38	21.7%	70	40.0%	51	29.1%	175
1-2	0-2	84	29.3%	81	28.2%	227	79.1%	221	77.0%	287
1-2	3+	102	24.7%	98	27.3%	180	43.6%	157	39.1%	413
3+	0+	66	5.4%	63	5.2%	262	20.7%	270	22.2%	1216

Figure 2: Solved problems grouped by  $|\mathcal{Q}|$  and  $|\mathcal{X}|$ .

Embedding	571	21.4%
Embedding (idem.)	557	20.9%
ileanCoP 1.2	875	32.8%
nanoCoP-i 2.0	858	32.1%
Total	2669	100%

Figure 1: Solved problems.

Figure 3 shows that our approach solved 52 problems not solved by the state-of-the-art provers. Note that ileanCoP and nanoCop were combined for presentation purposes but they also have a significant symmetric difference, e.g. of the 417 problem not solved by our approach there were 68 solved only by ileanCoP and 55 solved only by nanoCoP.

We conclude that our approach already has merits when  $|\mathcal{Q}| = 0$  and  $|\mathcal{X}|$  is small. We see that encoding idempotence does help to solve some additional benchmarks, but in general does not improve performance. However, idempotence is a quite powerful concept from a model-theoretic perspective. A more clever encoding of this property could improve results. There are other restrictions on models we could consider, based on structural properties of the formula, e.g. the relation between occurrences of different atoms.

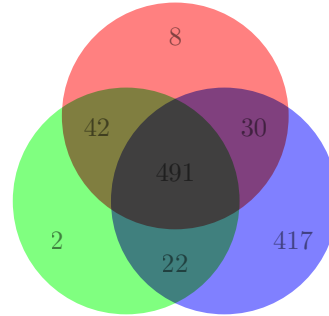


Figure 3: Instances proven by combinations of Embedding (red), Embedding (idem.) (green), ileanCoP or nanoCoP (blue)

## 9 Conclusion

We have presented embeddings of intuitionistic into classical logic for the propositional and the predicate case. Using Kripke semantics and model reduction we have established a complexity overhead of intuitionistic logic over classical logic parametrized by  $|\mathcal{X}|$  and  $|\mathcal{Q}|$ . This is directly reflected in the exponential blow-up parametrized by  $|\mathcal{X}|$  in the propositional case and more indirectly in our benchmarks in the predicate case.

A key motivation for our work has been the potential of using classical provers for intuitionistic proofs. Our tests have shown that there are cases where this is already reasonable. However, the current complexity considerations are direct consequences of our straightforward constructions. We plan to establish better bounds in future work by utilizing structural properties of the input formula. We hope that this will allow the construction of smaller counter-examples  $\mathcal{M}_T$  and the use of redundant axioms to limit search in a more clever way.

On the theoretical side, we also hope to give a new translation from QBF to IPC that improves our understanding of the relationship between intuitionistic propositional logic and the polynomial hierarchy. Finally, we plan to complement the counter-model translation presented in this paper by a proof translation. We will target the particular calculi used by state-of-the-art provers, e.g. superposition in the case of Vampire. Ultimately we hope that this will also open a new path for program extraction via the Curry–Howard Correspondence.



## References

- [1] The international sat competition. <http://www.satcompetition.org/>. Accessed: 2022-07-15.
- [2] The CADE ATP System Competition (CASC). <https://www.tptp.org/CASC/28/>. Accessed: 2022-07-14.
- [3] The Intuitionistic Logic Theorem Proving (ILTP) library. <http://iltp.de/results.html>. Accessed: 2022-07-14.
- [4] The Leo-III project. <http://page.mi.fu-berlin.de/lex/leo3/>. Accessed: 2022-10-6.
- [5] Matthias Baaz and Rosalie Iemhoff. The Skolemization of existential quantifiers in intuitionistic logic. *Annals of Pure and Applied Logic*, 142(1-3):269–295, 2006.
- [6] Leo Bachmair and Harald Ganzinger. Resolution Theorem Proving. *Handbook of automated reasoning*, 1(02), 2001.
- [7] Ulrich Berger, Wilfried Buchholz, and Helmut Schwichtenberg. Refined program extraction from classical proofs. *Annals of Pure and Applied Logic*, 114(1-3):3–25, 2002.
- [8] Armin Biere, Marijn Heule, Hans van Maaren, and Toby Walsh, editors. *Handbook of Satisfiability*, volume 185 of *Frontiers in Artificial Intelligence and Applications*. IOS Press, 2009.
- [9] Errett Bishop. Foundations of constructive analysis. 1967.
- [10] Luitzen Egbertus Jan Brouwer. *Over de grondslagen der wiskunde*. Maas & van Suchtelen, 1907.
- [11] Stephen A Cook. The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, pages 151–158, 1971.
- [12] Camillo Fiorentini, Rajeev Goré, and Stéphane Graham-Lengrand. A proof-theoretic perspective on SMT-solving for intuitionistic propositional logic. In *International Conference on Automated Reasoning with Analytic Tableaux and Related Methods*, pages 111–129. Springer, 2019.
- [13] Harvey Friedman. Classically and intuitionistically provably recursive functions. In *Higher set theory*, pages 21–27. Springer, 1978.
- [14] Michael R Garey and David S Johnson. *Computers and intractability*, volume 174. freeman San Francisco, 1979.
- [15] Gerhard Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische annalen*, 112(1):493–565, 1936.
- [16] Valery Glivenko. Sur quelques points de la logique de M. Brouwer. *Bulletins de la classe des sciences*, 15(5):183–188, 1929.
- [17] Kurt Gödel. Zur intuitionistischen Arithmetik und Zahlentheorie. *Ergebnisse eines mathematischen Kolloquiums*, 4(1933):34–38, 1933.
- [18] William A Howard. The formulae-as-types notion of construction. *To HB Curry: essays on combinatory logic, lambda calculus and formalism*, 44:479–490, 1980.
- [19] Rosalie Iemhoff. The eskolemization of universal quantifiers. *Annals of Pure and Applied Logic*, 162(3):201–212, 2010.
- [20] Dan Rosén Koen Claessen. SAT modulo intuitionistic implications. In *Logic for Programming, Artificial Intelligence, and Reasoning*, pages 622–637. Springer, 2015.
- [21] Laura Kovács and Andrei Voronkov. First-order theorem proving and Vampire. In *International Conference on Computer Aided Verification*, pages 1–35. Springer, 2013.
- [22] Sean McLaughlin and Frank Pfenning. Efficient intuitionistic theorem proving with the polarized inverse method. In *International Conference on Automated Deduction*, pages 230–244. Springer, 2009.
- [23] Jens Otten. Clausal connection-based theorem proving in intuitionistic first-order logic. In *International Conference on Automated Reasoning with Analytic Tableaux and Related Methods*, pages 245–261. Springer, 2005.
- [24] Jens Otten. leanCoP 2.0 and ileanCoP 1.2: high performance lean theorem proving in classical and intuitionistic logic (system descriptions). In *International Joint Conference on Automated*

- Reasoning*, pages 283–291. Springer, 2008.
- [25] Jens Otten. The nanoCoP 2.0 Connection Provers for Classical, Intuitionistic and Modal Logics. In *International Conference on Automated Reasoning with Analytic Tableaux and Related Methods*, pages 236–249. Springer, 2021.
  - [26] Richard Statman. Intuitionistic propositional logic is polynomial-space complete. *Theoretical Computer Science*, 9(1):67–72, 1979.
  - [27] Tanel Tammet. A resolution theorem prover for intuitionistic logic. In *International Conference on Automated Deduction*, pages 2–16. Springer, 1996.
  - [28] Anne Sjerp Troelstra and Helmut Schwichtenberg. *Basic proof theory*. Number 43. Cambridge University Press, 2000.
  - [29] Max Wisniewski, Alexander Steen, and Christoph Benzmüller. TPTP and beyond: Representation of quantified non-classical logics. *ARQNL 2016. Automated Reasoning in Quantified Non-Classical Logics*, pages 51–65, 2016.

## A Omitted proofs

**Lemma A.1.** For every predicate formula  $\varphi$  there exists a nullary predicate symbol  $P$  as well as sets of clauses  $\mathcal{R}, \mathcal{X}, \mathcal{Q}$ , with  $\mathcal{R}$  containing *flat clauses* of the form

$$\forall \vec{x} \left( \bigwedge_i A_i \rightarrow \bigvee_j B_j \right),$$

$\mathcal{X}$  containing *implication clauses* of the form  $\forall \vec{x} ((A \rightarrow B) \rightarrow C)$  and  $\mathcal{Q}$  containing *quantification clauses* of the form  $\forall \vec{x} ((\forall y A) \rightarrow B)$ , such that  $\varphi$  is intuitionistically equivalent to

$$\left( \bigwedge \mathcal{R} \wedge \bigwedge \mathcal{X} \wedge \bigwedge \mathcal{Q} \right) \rightarrow P$$

where  $A_i, B_i, A, B, C$  are atomic. This size of  $\mathcal{R}, \mathcal{X}, \mathcal{Q}$  is linear in the size of  $\varphi$ . In general these formulae will contain new function and predicate symbols.

*Proof.* As in [23] we may assume that  $\varphi$  is of the form  $\psi \rightarrow P$ . Recall that the transformation works by transforming  $\psi$  into a set of equivalent assumptions  $[\psi]$  of the desired form. We extend the given procedure with rules for quantifiers as follows:

$$\begin{aligned} [(\forall x \psi) \rightarrow A] &:= [\psi[s(\vec{z})/x] \rightarrow A] \\ [A \rightarrow (\forall x \psi)] &:= \{\forall x \chi \mid \chi \in [A \rightarrow \psi]\} \\ [(\exists x \psi) \rightarrow A] &:= \{\forall x \chi \mid \chi \in [A \rightarrow \psi]\} \\ A \rightarrow (\exists x \psi) &:= [A \rightarrow \psi[s(\vec{z})/x]] \end{aligned}$$

where  $s$  is a new function symbol,  $a$  is a new free variable and  $\vec{z}$  ranges over the free variables in  $\psi$ . While the generated set of clauses is in general not equivalent to  $\psi$ , section 2.4 guarantees that it is at least equivalent.  $\square$

**Lemma A.2.** Suppose  $\mathcal{M} = (M, T)$  is a counter-model to  $\varphi^{CH}$ , then so is  $\mathcal{M}_T = (M_T, I_T)$ .

*Proof.* Let  $v$  be any variable assignment. We show that  $\mathcal{M}_T, v \not\models \varphi$ . It follows directly from the definitions and the same facts for  $\mathcal{M}$  that  $\mathcal{M}_T, v \not\models P(s)$  and  $\mathcal{M}_T, v \models K(\varphi)$ . It remains to show that  $\mathcal{M}_T, v \models \bigwedge \mathcal{R}^{bS}$  and  $\mathcal{M}_T, v \models \bigwedge \mathcal{X}^{bS}$ .

Let

$$\forall w \phi = \forall w \left( s \preceq w \rightarrow \bigwedge_i A_i(w) \rightarrow \bigvee_j B_j(w) \right) \in \mathcal{R}^{bS}$$

and  $\psi_1 \dots \psi_n \in M_T$ . Since  $\mathcal{M}$  is a counter-model to  $\varphi$  we know that  $\mathcal{M}, v \models \forall w \phi$  and thus  $\mathcal{M}, v[g_{\psi_n}(\dots(g_{\psi_1}(s^I))\dots)]/w \models \phi$ . Then by definition  $\mathcal{M}_T, v[\psi_1 \dots \psi_n/w] \models \phi$ . This shows that  $\mathcal{M}_T, v \models \forall w \phi$ .

On the other hand consider

$$\forall w \phi = \forall w (s \preceq w \rightarrow (w \preceq f_\psi(w) \rightarrow A(f_\psi(w)) \rightarrow B(f_\psi(w))) \rightarrow C(w)) \in \mathcal{X}^{bS}$$

and  $\psi_1 \dots \psi_n \in M_T$ . Let  $m = g_{\psi_n}(\dots(g_{\psi_1}(s^I))\dots)$ . If  $\mathcal{M}, v[m/w] \models C(w)$  it follows directly that  $\mathcal{M}_T, v[\psi_1 \dots \psi_n/w] \models C(w)$  and we are done. Otherwise  $\mathcal{M}, v[m/w] \not\models s \preceq w$  or  $\mathcal{M}, v[m/w] \not\models w \preceq f_\psi(w) \rightarrow A(f_\psi(w)) \rightarrow B(f_\psi(w))$ . The former is impossible since  $g_\psi$  is increasing for all  $\psi$ . For the latter to hold we must have  $\mathcal{M}, v[m/w] \models w \preceq f_\psi(w)$  as well as

$\mathcal{M}, v[m/w] \not\models A(f_\psi(w)) \rightarrow B(f_\psi(w))$ . Since by definition  $\mathcal{M}_T, v[\psi_1 \dots \psi_m/w] \models s \preceq w \wedge w \preceq f_\psi(w)$  we must show  $\mathcal{M}_T, v[\psi_1 \dots \psi_m/w] \not\models A(f_\psi(w)) \rightarrow B(f_\psi(w))$ .

There are now 2 cases. Suppose  $\psi$  is fulfilled at  $m$ . Then we have  $\mathcal{M}, v[m/w] \not\models A(w) \rightarrow B(w)$  and  $g_\psi(m) = m$  and therefore  $\mathcal{M}, v[\psi_1 \dots \psi_n/w] \not\models A(f_\psi(w)) \rightarrow B(f_\psi(w))$ , i.e. indeed  $\mathcal{M}, v[\psi_1 \dots \psi_n/w] \models \phi$ .

On the other hand suppose  $\psi$  is not fulfilled at  $m$ , i.e.  $g_\psi(m) = f_\psi^I(m)$ . Since  $\mathcal{M}, v[m/w] \not\models A(f_\psi(w)) \rightarrow B(f_\psi(w))$  it directly follows that  $\mathcal{M}_T, v[\psi_1 \dots \psi_n/w] \not\models A(f_\psi(w)) \rightarrow B(f_\psi(w))$  and thus  $\mathcal{M}_T, v[\psi_1 \dots \psi_n/w] \not\models A(f_\psi(w)) \rightarrow B(f_\psi(w))$ , i.e. indeed  $\mathcal{M}_T, v[\psi_1 \dots \psi_n/w] \models \phi$ .

In any case  $\mathcal{M}_T, v \models \forall w \phi$ .  $\square$

**Lemma A.3.**  $\varphi$  is not intuitionistically valid iff  $\varphi^Q$  is a satisfiable QBF.

*Proof.* Suppose  $\varphi$  is not intuitionistically valid. Then,  $(\bigwedge \mathcal{R} \wedge \bigwedge \mathcal{X}) \rightarrow P$  has an intuitionistic counter-model. By theorem 5.1 there exists a classical counter-model  $\mathcal{M}$  for  $(\bigwedge \mathcal{R}^\# \wedge \bigwedge \mathcal{X}^\#) \rightarrow P^\epsilon$ . We now define a QBF model iteratively. For each atom  $A$  interpret  $A^0$  such as  $\mathcal{M}$  interprets  $A^\epsilon$ . Suppose we are given interpretations of all atoms  $A^i$  for  $i < n$  and a sequence with no repetitions  $\psi_1 \dots \psi_{n-1}$  over  $\mathcal{X}$  such that  $\psi_i$  is exactly the  $\psi \in \mathcal{X}$  for which  $X_\psi^i$  is true and  $A^i$  is interpreted as  $A^{\psi_1 \dots \psi_i}$  in  $\mathcal{M}$ . Let the  $X_\psi^n$  be arbitrarily interpreted (since they are  $\forall$ -quantified). If not exactly one of the  $X_\psi^n$  is interpreted as true, then  $\text{valid}(n)$  fails and the remaining propositions can be chosen arbitrarily. If  $X_{\psi_n}^i = 1$  for some  $i < n$ , then  $\text{valid}(n)$  also fails and the remaining propositions can be chosen arbitrarily. So we may assume that  $\psi_1 \dots \psi_n$  is a sequence with no repetitions. Interpret the atoms  $A^n$  as  $\mathcal{M}$  interprets  $A^{\psi_1 \dots \psi_n}$ . Continue this construction until  $n = |\mathcal{X}|$ . Then from  $\mathcal{M}$  being a counter-example to  $(\bigwedge \mathcal{R}^\# \wedge \bigwedge \mathcal{X}^\#) \rightarrow P^\epsilon$  it directly follows that this interpretation satisfies  $\varphi^Q$ .

On the other hand suppose  $\varphi^Q$  is satisfiable. We construct a counter-example to  $\varphi^\#$ . Again we proceed iteratively. Interpret  $A^\epsilon$  such as  $A^0$  is interpreted in some satisfying interpretation of  $\varphi^Q$ . Suppose we are given a sequence  $\psi_1 \dots \psi_{n-1}$  such that for  $i < n$  having  $A^i = A^{\psi_0 \dots \psi_{n-1}}$  is part of a satisfying interpretation of  $\varphi^Q$ , in which  $X_\psi^i$  is chosen true iff  $\psi = \psi_i$ . Let  $\psi_n \in \mathcal{X}$ . Consider some interpretation of the  $A^n$  that are part of a satisfying assignment where  $X_\psi^n$  is true iff  $\psi = \psi_n$  and all variables quantified above are chosen as before. Have  $A^{\psi_0 \dots \psi_n} = A^n$  for each propositional variable  $A$ . From the definitions it directly follows that construction yields a counter-model for  $\varphi^\#$ .  $\square$

**Lemma A.4.**  $\varphi$  is intuitionistically valid if and only if  $\varphi^C$  is classically valid.

*Proof.* Since we are mixing worlds and proper intuitionistic domain elements in a single classical domain there are some details to be taken care of. We proceed by translation of counter-models.

Suppose first we have a counter-model  $\mathcal{M} = (M, I)$  to  $\varphi^C$ . As a Kripke frame  $(W, \leq)$  take  $W = \{m \in M \mid \text{World}^I(m)\}$  and let  $\leq$  be  $\preceq^I$  restricted to  $W$ . Then let  $M_u = \{m \in M \mid E(m, u)\}$  and let  $f^{Iu}$  be  $f^I$  restricted to  $M_u$  and  $A^{Iu}(\vec{x}) := A(\vec{x}, u)$ . It is then a straightforward check of definitions that this defines a Kripke counter-model to  $\varphi$ .

The other direction is a bit more involved. Suppose we have a Kripke counter-model to  $\varphi$  with frame  $(W, \preceq)$  and family of  $\Sigma$ -structures  $(M_w, I_w)_{w \in W}$ . In particular, since it is a counter-model there exists  $w_0 \in W$  and a variable assignment  $v$  with  $w_0, v \not\models \varphi$ . Let  $W_0 = \{w \in W \mid w_0 \preceq w\}$  and define an equivalence relation  $\sim$  on  $\{(x, u) \mid u \in W_0, x \in M_u\}$  via  $(x, u) \sim (y, w)$  iff  $x = y$  and there exists  $v \in W_0$  comparable with both  $u, w$  such that  $x \in v$  and denote the equivalence class of  $(x, u)$  with  $[x, u]$ . Let  $M = W_0 \cup \{[x, u] \mid u \in W_0, x \in M_u\}$ . Now have

- $\preceq^I = \leq \cup \{(m, m) : m \in M \setminus W_0\}$ , i.e.  $\preceq$  corresponds to  $\leq$  on  $W_0$  and is trivial elsewhere.

- $E^I(m, w)$  iff  $w \in W_0$  and  $m = [x, w]$  for some  $x \in M_w$ .
- $f^I(m_1 \dots m_n) = \begin{cases} f^{I_u}(x_1 \dots x_n), & \text{if there are } u \in W_0, x_i \in M_u \\ & \text{with } m_i = [x_i, u] \text{ for all } i, \\ w_0, & \text{otherwise.} \end{cases}$
- $A^I(m_1 \dots m_n, u) \Leftrightarrow \begin{cases} A^{I_u}(x_1 \dots x_n), & \text{if } u \in W_0 \text{ and } \exists x_i \in M_u \\ & \text{with } m_i = [x_i, u] \text{ for all } i, \\ \perp, & \text{otherwise} \end{cases}$

and  $\mathcal{M} = (M, I)$ . One easily verifies that these are well-defined. We now show by induction on the height of the formula the following:

*Claim.* For any formula  $\psi$  in which the free variable  $b$  doesn't occur, world  $w_k \in W_0$  and variable assignments  $v_c, v_i$  such that  $v_c(a) = [v_i(a), w_k]$  for all free variables  $a$  occurring in  $\psi$  it holds that

$$\mathcal{M}, v_c[w_k/b] \models \psi^b \text{ if and only if } w_k, v_i \models \psi.$$

For the base case, i.e. if  $\psi = \perp$  or  $\psi = P(t_1, \dots, t_n)$  for some  $n$ -ary predicate symbol  $P$ , the claim follows directly from the definitions.

For the induction step we must consider the cases where  $\psi$  is of the form  $\psi_0 \wedge \psi_1$ ,  $\psi_0 \vee \psi_1$ ,  $\psi_0 \rightarrow \psi_1$ ,  $\exists x\psi'$ ,  $\forall x\psi'$ , there former two are straightforward and are left out.

*Case 1:*  $\psi = \psi_0 \rightarrow \psi_1$ .

Suppose first that  $w_k, v_i \not\models \psi_0 \rightarrow \psi_1$ . Then there exists  $w_l \geq w_k$  such that  $w_l, v_i \models \psi_0$  and  $w_l, v_i \not\models \psi_1$ . Then by induction hypothesis  $\mathcal{M}, v_c[w_l/b] \models \psi_0^b$  and  $\mathcal{M}, v_c[w_l/b] \not\models \psi_1^b$ , i.e.  $\mathcal{M}, v_c[w_l/b] \not\models \psi$ . Suppose on the other hand that  $w_l, v_i \models \psi_0 \rightarrow \psi_1$ . Then for all  $w_k \geq w_l$  we have  $w_k \models \psi_0 \rightarrow \psi_1$ , i.e.  $w_k, v_i \not\models \psi_0$  or  $w_k, v_i \models \psi_1$ . Then by induction hypothesis  $\mathcal{M}, v_c[w_k/b] \not\models \psi_0$  or  $\mathcal{M}, v_c[w_k/b] \models \psi_1$ , i.e.  $\mathcal{M}, v_c[w_k/b] \models \psi_0 \rightarrow \psi_1$ .

*Case 2:*  $\psi = \exists x\psi'$ .

Suppose first that  $w_k, v_i \not\models \exists x\psi'$ , i.e. for all  $m_k \in M_{w_k}$  we have  $w_k, v_i[m_k/x] \not\models \psi'$  and by induction hypothesis  $\mathcal{M}, v_i[[m_k, w_k]/x, w_k/b] \not\models \psi'^b$ . Every other substitution for  $x$  violates  $E(x, b)$ . Therefore  $\mathcal{M}, v_i[w_k/b] \models \forall x(\neg E(x, b) \vee \neg \psi')$ , i.e.  $\mathcal{M}, v_i[w_k/b] \not\models \psi^b$ . Suppose on the other hand that  $w_k, v_i \models \exists x\psi'$  then there exists  $m_k \in M_{w_k}$  such that  $w_k, v_i[m_k/x] \models \psi'$ . Then by induction hypothesis  $\mathcal{M}, v_i[[m_k, w_k]/x, w_k/b] \models \psi'^b$ . By definition  $\mathcal{M}, v_i[[m_k, w_k]/x, w_k/b] \models E(x, b)$  holds. Therefore  $\mathcal{M}, v_i[w_k/b] \models \exists x(E(x, b) \wedge \psi')$ , i.e.  $\mathcal{M}, v_i[w_k/b] \models \psi^b$ .

*Case 3:*  $\psi = \forall x\psi'$ .

Suppose first that  $w_k, v_i \not\models \forall x\psi'$ , i.e. there exists  $w_l \geq w_k$  and  $m_l \in M_{w_l}$  such that  $w_l, v_i[m_l/x] \not\models \psi'$ . Then by the induction hypothesis  $\mathcal{M}, v_c[[m_l, w_l]/x, w_l/c] \not\models \psi'^c$ . Furthermore by definition  $\mathcal{M}, v_c[w_l/c, w_k/b] \models b \preceq c$  and  $\mathcal{M}, v_c[[m_l, w_l]/x, w_l/c] \models E(x, c)$ . Hence  $\mathcal{M}, v_c[w_k/b] \not\models \psi^b$ . Suppose on the other hand that  $w_k, v_i \models \forall x\psi'$ , i.e. for all  $w_l \geq w_k$  and  $m_l \in M_{w_l}$  it holds that  $w_l, v_i[m_l/x] \models \psi'$ . Then by the induction hypothesis  $\mathcal{M}, v_c[[m_l, w_l]/x, w_l/c] \models \psi'^c$  for all  $w_l \geq w_k$  and  $m_l \in M_{w_l}$ . On the other hand for every other assignment of  $c$  and  $x$  either  $b \not\preceq c$  or  $E(x, c)$  is violated. Hence  $\mathcal{M}, v_c[w_k/b] \models \psi^b$ .

This concludes our proof of the claim.

Now to show that  $\mathcal{M}$  is a counter-model to  $\varphi^C$ : Recall that  $w_0, v \not\models \varphi$ . From the claim it follows that

$$\mathcal{M}, v'[[v(a_0), w_0]/a_0 \dots [v(a_n), w_0]/a_n, w_0/b] \not\models \varphi^b$$

where  $a_0 \dots a_n$  are the free variables occurring in  $\varphi$  and  $v'$  is any variable assignment. Furthermore a straightforward check of definitions yields

$$\mathcal{M}, v'[[v(a_0), w_0]/a_0 \dots [v(a_n), w_0]/a_n, w_0/b] \models K(\varphi) \wedge \text{World}(b) \wedge \vec{E}(\vec{a}, b).$$

Therefore  $\mathcal{M}, v' \not\models \varphi^C$  and we are done.  $\square$