GLIVENKO AND KURODA FOR SIMPLE TYPE THEORY

CHAD E. BROWN AND CHRISTINE RIZKALLAH

Abstract. Glivenko's theorem states that an arbitrary propositional formula is classically provable if and only if its double negation is intuitionistically provable. The result does not extend to full first-order predicate logic, but does extend to first-order predicate logic without the universal quantifier. A recent paper by Zdanowski shows that Glivenko's theorem also holds for second-order propositional logic without the universal quantifier. We prove that Glivenko's theorem extends to some versions of simple type theory without the universal quantifier. Moreover we prove that Kuroda's negative translation, which is known to embed classical first-order logic into intuitionistic first-order logic, extends to the same versions of simple type theory. We also prove that the Glivenko property fails for simple type theory once a weak form of functional extensionality is included.

1. Introduction

Glivenko's theorem states that an arbitrary propositional formula s is classically provable if and only if $\neg\neg s$ is intuitionistically provable [8]. Glivenko's theorem does not extend to first-order predicate logic, but does extend to first-order predicate logic without $\forall$ quantifiers [11]. A recent paper by Zdanowski [15] shows that Glivenko's theorem also holds for second-order propositional logic without the $\forall$ quantifier. In this paper we consider how Glivenko's result extends to simple type theories in the style of Church [4]. Such simple type theories are also referred to as higher-order logics.

A related result by Kuroda states that a first-order formula s is classically provable if and only if $\neg\neg s'$ is intuitionistically provable, where s' is obtained from s by adding double negations beneath each universal quantifier. The mapping from s to $\neg\neg s'$ is known as the *Kuroda negative translation* [11]. The Kuroda negative translation embeds classical first-order logic into intuitionistic first-order logic. Kuroda's translation extends in a natural way to higher-order formulas. We determine

This is a technical report version of a paper submitted in December 2011.

when this extension embeds the classical version of a simple type theory into the corresponding intuitionistic version.

There are other logical transformations that translate classical first-order logic into intuitionistic first-order logic including the *Kolmogorov negative translation* [10] and the *Gödel-Gentzen negative translation* [7, 9]. Unlike the Kuroda negative translation, the Kolmogorov and the Gödel-Gentzen negative translations do not extend in a natural way to higher-order logic. In both cases, propositional atoms are not mapped to themselves but to their double negations. In higher order logic a propositional atom may be a variable. If a translation adds a double negation to such variables, then the translation will not respect β-equivalence.

For the first part of the paper we consider simple type theory without extensionality. We prove the Glivenko and Kuroda results extend to the non-extensional case. We then consider simple type theory with the possible addition of three forms of extensionality: propositional extensionality, η-extensionality and ξ-extensionality. It turns out that the Glivenko and Kuroda results continue to hold if we include either propositional extensionality or η-extensionality (or both). Note that this implies the $\forall$ quantifier cannot be defined from the other logical constants in the intuitionistic versions of simple type theory without ξ-extensionality. On the other hand, it is well-known that all the logical constants (including $\forall$) can be defined from equality if one has all three forms of extensionality (see [2, 12]). Consequently, the Glivenko result fails in the presence of all three forms of extensionality. In fact, we prove the Glivenko result fails in each of the versions of simple type theory with ξ-extensionality. To summarize, the Glivenko property holds for simple type theory without $\forall$ and without ξ-extensionality, but does not hold for simple type theory with $\forall$ or with ξ-extensionality. Furthermore, the Kuroda translation preserves provability for simple type theory without ξ-extensionality, but not for simple type theory with ξ-extensionality.

In Section 2 we give a short presentation of simply typed λ-calculus. In Section 3 we present natural deduction calculi for a few variants of simple type theory. We prove the Kuroda and Glivenko results in Section 4. In Section 5 we consider what happens when extensionality principles are included.

2. Simply typed lambda calculus

We describe simply typed λ-calculus in the style of Church [4]. The set of *types* is given inductively: o and ι are types and $\sigma\tau$ is a type whenever σ and τ are types. The types o (the type of propositions) and ι (the type of individuals) are called *base types*. Types of the form $\sigma\tau$ are called *function types*. We use σ and τ to range over types.

For each type σ, let $\mathcal{N}_\sigma$ be an infinite set of *names* of type σ. Some of the names are *logical constants*:

- $\bot$ is a logical constant in $\mathcal{N}_o$.
- $\wedge$, $\vee$ and $\rightarrow$ are (distinct) logical constants in $\mathcal{N}_{ooo}$.
- For each σ, $=_\sigma$ is a logical constant in $\mathcal{N}_{\sigma\sigma o}$.
- For each σ, $\forall_\sigma$ and $\exists_\sigma$ are (distinct) logical constants in $\mathcal{N}_{(\sigma o)o}$.

The remaining names are *variables* . Let $\mathcal{C}_\sigma$ be the set of logical constants of type σ and $\mathcal{V}_\sigma$ be the (infinite) set of all variables of type σ. Let $\mathcal{N} = \bigcup_\sigma \mathcal{N}_\sigma$, $\mathcal{C} = \bigcup_\sigma \mathcal{C}_\sigma$ and $\mathcal{V} = \bigcup_\sigma \mathcal{V}_\sigma$.

For any $\mathcal{C}' \subseteq \mathcal{C}$ we define a family of sets of terms $\Lambda_\sigma^{\mathcal{C}'}$ for each type σ by induction.

- For every $x \in \mathcal{V}_\sigma$, $x \in \Lambda_\sigma^{\mathcal{C}'}$.
- For every $c \in \mathcal{C}'$, $c \in \Lambda_\sigma^{\mathcal{C}'}$.
- For every $x \in \mathcal{V}_\sigma$ and $s \in \Lambda_\tau^{\mathcal{C}'}$, $(\lambda x.s) \in \Lambda_{\sigma\tau}^{\mathcal{C}'}$.
- For every $s \in \Lambda_{\sigma\tau}^{\mathcal{C}'}$ and $t \in \Lambda_\sigma^{\mathcal{C}'}$, $(st) \in \Lambda_\tau^{\mathcal{C}'}$.

We defined $\Lambda_\sigma^{\mathcal{C}'}$ relative to a set of logical constants. There are two particular sets of logical constants of interest in this paper: the full set $\mathcal{C}$ of logical constants and the set

$$\mathcal{C}^- := \mathcal{C} \setminus \{\forall_\sigma | \sigma \text{ is a type}\}$$

omitting the $\forall$-quantifier. To simplify notation, we define Λ_σ to be $\Lambda_\sigma^{\mathcal{C}}$ and Λ_σ^- to be $\Lambda_\sigma^{\mathcal{C}^-}$. Note that $\Lambda_\sigma^{\mathcal{C}'} \subseteq \Lambda_\sigma$ for any $\mathcal{C}' \subseteq \mathcal{C}$. An element of Λ_σ is called a *term of type* σ. A *term* is an element of $\bigcup_\sigma \Lambda_\sigma$. Terms of the form $(\lambda x.s)$ are called λ-*abstractions*. Terms of the form (st) are called *applications*. A *formula* is a term of type o. Such formulas are sometimes called *higher-order formulas*.

We write $\neg s$ for $((\rightarrow s)\bot)$. We write stu for $(st)u$, except that $\neg st$ means $\neg(st)$. We use the infix notation $s \wedge t$, $s \vee t$, $s \rightarrow t$ and $s =_\sigma t$ as shorthand for $\wedge st$, $\vee st$, $\rightarrow st$ and $=_\sigma st$, respectively. Note that $\neg s$ is $s \rightarrow \bot$. We write $\forall x : \sigma.s$ and $\exists x : \sigma.s$ for $\forall_\sigma(\lambda x.s)$ and $\exists_\sigma(\lambda x.s)$, respectively. We may also omit the type entirely from a quantified formula or equation and write $\forall x.s$, $\exists x.s$ or $s = t$ when the types are clear in context.

When s is a term, $w \in \mathcal{N}_\sigma$ and t is a term of type σ, then s_t^w is defined to be the result of substituting t for w in s via a capture-avoiding substitution. A term of the form $(\lambda x.s)t$ is called a β-*redex* with β-*reduct* s_t^x. We say s β-*reduces to* t (and write $s \rightarrow_\beta t$) if a subterm of s is a β-redex such that t is the result of replacing this subterm by its β-reduct. We define $s \sim_\beta t$ to be the least equivalence relation containing $\to_\beta$. When $s \sim_\beta t$ holds, we say s and t are *β-equivalent*.

3. Natural deduction calculi

By a *simple type theory* we mean simply typed λ-terms with a notion of provability. Another phrase used in the literature for such a language and notion of provability is *higher-order logic*. In this section we define three notions of provability via natural deduction: classical non-extensional simple type theory (sometimes called *elementary type theory* [1]), classical non-extensional simple type theory without $\forall$, and intuitionistic non-extensional simple type theory. All the natural deduction rules we will need to define these calculi are given in Figure 1.

- $\Gamma \vdash_K s$ (where Γ is a finite subset of Λ_o and $s \in \Lambda_o$) holds when derivable using all the rules in Figure 1.
- $\Gamma \vdash_K^- s$ (where Γ is a finite subset of Λ_o^- and $s \in \Lambda_o^-$) holds when derivable using all the rules in Figure 1 (restricted to terms in Λ^-) except $\mathbf{N}_\forall^I$ and $\mathbf{N}_\forall^E$.
- $\Gamma \vdash_J s$ (where Γ is a finite subset of Λ_o and $s \in \Lambda_o$) holds when derivable using all the rules in Figure 1 except $\mathbf{N}_c$.

4. Glivenko and Kuroda for non-extensional simple type theory

We prove Glivenko's theorem extends to non-extensional simple type theory without the $\forall$ quantifier. Extensionality is handled separately in the next section. Furthermore, we provide a translation from classical simple type theory into intuitionistic simple type theory (see [13]). This translation turns out to be an extension of Kuroda's negative translation to simple type theory. Finally, we give an example illustrating Glivenko does not extend to full simple type theory (with $\forall$). This is not surprising since it does not extend to full first-order logic either.

We first discuss whether the Kolmogorov and the Gödel-Gentzen negative translations extend to higher-order logic. Assume that when a translation Ψ is extended to higher-order logic it translates $\Psi(st)$ as $\Psi(s)\Psi(t)$ and $\Psi(\lambda x.s)$ as $\lambda x.\Psi(s)$. In other words, assume the translation is compositional with respect to application and λ-abstraction. Under this assumption, there is no β-respecting extension of the Kolmogorov and the Gödel-Gentzen negative translations to higher-order logic. We provide an example showing that they do not respect β-equivalence. Let p be a variable of type o. The result of translating

$$\mathbf{N}_h \ \frac{}{\Gamma \vdash s}\ s \in \Gamma \qquad \mathbf{N}_w \ \frac{\Gamma \vdash s}{\Delta \vdash s}\ \Gamma \subseteq \Delta \qquad \mathbf{N}_\beta \ \frac{\Gamma \vdash s}{\Gamma \vdash t}\ s \sim_\beta t$$

$$\mathbf{N}^E_\bot \ \frac{\Gamma \vdash \bot}{\Gamma \vdash t} \qquad \mathbf{N}^I_\to \ \frac{\Gamma, s \vdash t}{\Gamma \vdash s \to t} \qquad \mathbf{N}^E_\to \ \frac{\Gamma \vdash s \to t \qquad \Gamma \vdash s}{\Gamma \vdash t}$$

$$\mathbf{N}^I_\wedge \ \frac{\Gamma \vdash s_1 \qquad \Gamma \vdash s_2}{\Gamma \vdash s_1 \wedge s_2} \qquad \mathbf{N}^{EL}_\wedge \ \frac{\Gamma \vdash s_1 \wedge s_2}{\Gamma \vdash s_1} \qquad \mathbf{N}^{ER}_\wedge \ \frac{\Gamma \vdash s_1 \wedge s_2}{\Gamma \vdash s_2}$$

$$\mathbf{N}^{IL}_\vee \ \frac{\Gamma \vdash s_1}{\Gamma \vdash s_1 \vee s_2} \qquad \mathbf{N}^{IR}_\vee \ \frac{\Gamma \vdash s_2}{\Gamma \vdash s_1 \vee s_2}$$

$$\mathbf{N}^E_\vee \ \frac{\Gamma \vdash s_1 \vee s_2 \qquad \Gamma, s_1 \vdash t \qquad \Gamma, s_2 \vdash t}{\Gamma \vdash t} \qquad \mathbf{N}^I_= \ \frac{}{\Gamma \vdash s =_\sigma s}$$

$$\mathbf{N}^E_= \ \frac{\Gamma \vdash s =_\sigma t \qquad \Gamma \vdash us}{\Gamma \vdash ut} \qquad \mathbf{N}^I_\exists \ \frac{\Gamma \vdash st}{\Gamma \vdash \exists_\sigma s}$$

$$\mathbf{N}^E_\exists \ \frac{\Gamma \vdash \exists_\sigma s \qquad \Gamma, sy \vdash t}{\Gamma \vdash t}\ y \notin \mathcal{N}\Gamma \cup \mathcal{N}s \cup \mathcal{N}t$$

$$\mathbf{N}^I_\forall \ \frac{\Gamma \vdash sy}{\Gamma \vdash \forall_\sigma s}\ y \notin \mathcal{N}\Gamma \cup \mathcal{N}s \qquad \mathbf{N}^E_\forall \ \frac{\Gamma \vdash \forall_\sigma s}{\Gamma \vdash st} \qquad \mathbf{N}_c \ \frac{\Gamma,\ s \to \bot \vdash \bot}{\Gamma \vdash s}$$

Figure 1. Natural deduction rules

p using either of the translations is $\neg\neg p$ [5]. Consider the two β-equivalent terms $(\lambda p.p)p$ and p. The translations of these two terms are $(\lambda p.\neg\neg p)(\neg\neg p)$ and $\neg\neg p$, respectively. The resulting terms are not β-equivalent. The reason for this is that the translations double negate the atoms in a formula.

We now introduce a lemma which provides some helpful formulas that we will make use of later on.

Lemma 4.1. *The following are derivable. Their derivation is left as an exercise for the reader.*

(1) $\vdash_J \neg\neg\bot \to \bot$
(2) $\vdash_J \forall pq.(p \to \neg\neg q) \to \neg\neg(p \to q)$
(3) $\vdash_J \forall pq.\neg\neg p \to \neg\neg(p \to q) \to \neg\neg q$
(4) $\vdash_J \forall pq.\neg\neg p \to \neg\neg q \to \neg\neg(p \wedge q)$

(5) $\vdash_J \forall pq.\neg\neg(p \wedge q) \to \neg\neg p$
(6) $\vdash_J \forall pq.\neg\neg(p \wedge q) \to \neg\neg q$
(7) $\vdash_J \forall pq.\neg\neg p \to \neg\neg(p \vee q)$
(8) $\vdash_J \forall pq.\neg\neg q \to \neg\neg(p \vee q)$
(9) $\vdash_J \forall p_1 p_2 q.\neg\neg(p_1 \vee p_2) \to (p_1 \to \neg\neg q) \to (p_2 \to \neg\neg q) \to \neg\neg q$
(10) $\vdash_J \forall p.p \to \neg\neg p$
(11) $\vdash_J \forall xy.\forall p.\neg\neg(x =_\sigma y) \to \neg\neg(px) \to \neg\neg(py)$
(12) $\vdash_J \forall p.\forall x.\neg\neg(px) \to \neg\neg\exists_\sigma p$
(13) $\vdash_J \forall pq.\neg\neg\exists_\sigma p \to (\forall x.px \to \neg\neg q) \to \neg\neg q$
(14) $\vdash_J \forall p.\neg\neg\neg p \to \neg p$

We prove that if a higher-order formula s with no $\forall$ quantifiers is classically provable using the rules in $\vdash_K^-$ then $\neg\neg s$ is intuitionistically provable using the rules in $\vdash^J$.

Theorem 4.2. *For all finite $\Gamma \subseteq \Lambda_o^-$ and $s \in \Lambda_o^-$, if $\Gamma \vdash_K^- s$, then $\Gamma \vdash_J \neg\neg s$.*

Proof. We prove the theorem by induction on the derivation of $\Gamma \vdash_K^- s$. Recall that the rules $\mathbf{N}_\forall^I$ and $\mathbf{N}_\forall^E$ can be used for a derivation in $\vdash_J$ but not for a derivation in $\vdash_K^-$. Conversely, $\mathbf{N}_c$ can be used for a derivation in $\vdash_K^-$ but not for a derivation in $\vdash_J$.

We first show the base cases of the induction, namely the case that the derivation of $\Gamma \vdash_K^- s$ consists of exactly one step, either the $\mathbf{N}_h$ rule or the $\mathbf{N}_=^I$ rule. If either one of the two rules is applied, we assume $\Gamma \vdash_K^- s$ and know $\Gamma \vdash_J \neg\neg s$ by the applied rule and by Lemma 4.1(10) (along with $\mathbf{N}_\forall^E$, $\mathbf{N}_\to^E$, and $\mathbf{N}_\beta$).

We consider the case that the derivation of $\Gamma \vdash_K^- s$ ends with the $\mathbf{N}_\exists^E$ rule. By the inductive hypothesis, we know $\Gamma \vdash_J \neg\neg\exists_\sigma t$ and $\Gamma, ty \vdash_J \neg\neg s$. We have $\Gamma \vdash_J \forall y.ty \to \neg\neg s$ using $\mathbf{N}_\to^I$ and $\mathbf{N}_\forall^I$. Using Lemma 4.1(13) (along with $\mathbf{N}_\to^E$, $\mathbf{N}_\forall^E$, $\mathbf{N}_w$, and $\mathbf{N}_\beta$) , we obtain $\Gamma \vdash_J \neg\neg t$.

If the derivation of $\Gamma \vdash_K^- s$ ends with the $\mathbf{N}_c$ rule. Then, by the inductive hypothesis we have $\Gamma, s \to \bot \vdash_J \neg\neg\bot$ and hence $\Gamma, s \to \bot \vdash_J \bot$ by Lemma 4.1(1). Thus $\Gamma \vdash_J \neg\neg s$ directly using $\mathbf{N}_\to^I$.

Moreover, the other cases of the induction step hold similarly. Namely, $\mathbf{N}_\bot^E$ follows from Lemma 4.1(1), $\mathbf{N}_\to^I$ from Lemma 4.1(2), $\mathbf{N}_\to^E$ from Lemma 4.1(3), $\mathbf{N}_\wedge^I$ from Lemma 4.1(4), $\mathbf{N}_\wedge^{EL}$ from Lemma 4.1(5), $\mathbf{N}_\wedge^{ER}$ from Lemma 4.1(6), $\mathbf{N}_\vee^{IL}$ from Lemma 4.1(7), $\mathbf{N}_\vee^{IR}$ from Lemma 4.1(8), $\mathbf{N}_\vee^E$ from Lemma 4.1(9), $\mathbf{N}_=^E$ from Lemma 4.1(11), and $\mathbf{N}_\exists^I$ from Lemma 4.1(12). The $\mathbf{N}_w$ case is obvious, and the $\mathbf{N}_\beta$ case follows from the fact that if $s \sim_\beta t$, then $\neg\neg s \sim_\beta \neg\neg t$. □

We now define a translation Φ from general higher-order formulas to formulas not containing $\forall$ quantifiers.

Definition 4.3. *We define $\Phi : \Lambda \to \Lambda^-$ by recursion as follows.*

$$\begin{array}{rcll} \Phi(x) & := & x & \textit{for variables } x \\ \Phi(c) & := & c & \textit{for } c \in \mathcal{C}^- \\ \Phi(\forall_\sigma) & := & \lambda p:\sigma.\neg\exists x.\neg px & \textit{where } p \in \mathcal{V}_{\sigma o} \textit{ and } x \in \mathcal{V}_\sigma \\ \Phi(st) & := & \Phi(s)\Phi(t) & \\ \Phi(\lambda x.s) & := & \lambda x.\Phi(s) & \end{array}$$

We call Φ compositional *since it respects application and λ-abstraction. For $\Gamma \subseteq \Lambda$ we write $\Phi(\Gamma)$ to mean $\{\Phi(t) | t \in \Gamma\}$.*

Note that $\neg\exists x.\neg px$ is intuitionistically equivalent to $\forall x.\neg\neg px$. Therefore, the translation taking a formula s to $\neg\neg\Phi(s)$ is equivalent to Kuroda's negative translation [11] in the sense of the equivalence defined in [5].

We now prove that if a general higher-order formula s is classically provable using the rules in $\vdash_K$, then its translation $\Phi(s)$ which does not have $\forall$ quantifiers is provable using the rules in $\vdash_K^-$.

Lemma 4.4. *For all finite $\Gamma \subseteq \Lambda_o$ and $s \in \Lambda_o$, if $\Gamma \vdash_K s$, then $\Phi(\Gamma) \vdash_K^- \Phi(s)$.*

Proof. We prove the theorem by induction on the derivation of $\Gamma \vdash_K s$. We consider the interesting cases where the derivation ends with the $\mathbf{N}_\forall^I$ or $\mathbf{N}_\forall^E$ rule. We also consider the rules $\mathbf{N}_=^E$ and $\mathbf{N}_\exists^I$. Similar to those two cases, for all the other cases the same rule that was used to derive $\Gamma \vdash_K s$ can be used to derive $\Phi(\Gamma) \vdash_K^- \Phi(s)$.

In the $\mathbf{N}_\forall^I$ case, we assume $\Gamma \vdash_K sy$ where $y \notin \mathcal{N}\Gamma \cup \mathcal{N}s$. By the induction hypothesis we have $\Phi(\Gamma) \vdash_K^- \Phi(sy)$. We need to prove that $\Phi(\Gamma) \vdash_K^- \Phi(\forall_\sigma s)$. Since Φ is compositional, we know $\Phi(sy) = \Phi(s)y$ and $y \notin \mathcal{N}\Phi(\Gamma) \cup \mathcal{N}\Phi(s)$. Without loss of generality assume $x \notin \mathcal{N}s$, by definition of Φ and the fact that it is compositional we know $\Phi(\forall_\sigma s) \sim_\beta \neg\exists x.\neg\Phi(s)x$. By $\mathbf{N}_\beta$ and $\mathbf{N}_\to^I$ it suffices to show $\Phi(\Gamma), \exists x.\neg\Phi(s)x \vdash_K^- \bot$. Since $y \notin \mathcal{N}\Phi(\Gamma) \cup \mathcal{N}\Phi(s) \cup \mathcal{N}(\exists x.\neg\Phi(s)x) \cup \mathcal{N}\bot$, by $\mathbf{N}_\exists^E$ and $\mathbf{N}_h$ it is enough to show $\Phi(\Gamma), \exists x.\neg\Phi(s)x, \neg\Phi(s)y \vdash_K^- \bot$. This we can obtain easily by using $\mathbf{N}_\to^E$ and the fact that $\Phi(\Gamma), \exists x.\neg\Phi(s)x, \neg\Phi(s)y \vdash_K^- \Phi(s)y$. This fact we know by applying the $\mathbf{N}_w$ rule to the induction hypothesis.

For the $\mathbf{N}_\forall^E$ rule, by the inductive hypothesis we have $\Phi(\Gamma) \vdash_K^- \Phi(\forall_\sigma s)$ and show $\Phi(\Gamma) \vdash_K^- \Phi(st)$. Without loss of generality, assume $x \notin \mathcal{N}s$, by the induction hypothesis and the definition of Φ we have $(\lambda f.\neg\exists x.\neg fx)\Phi(s)$ which is β-equivalent to $\neg\exists x.\neg(\Phi(s)x)$. By $\mathbf{N}_w$, we can obtain $\Phi(\Gamma), \neg\Phi(st) \vdash_K^- \neg\exists x.\neg(\Phi(s)x)$. Hence by using $\mathbf{N}_c$ and $\mathbf{N}_{\to}^E$ it is enough show $\Phi(\Gamma), \neg\Phi(st) \vdash_K^- \exists x.\neg(\Phi(s)x)$. This can be obtained by applying the $\mathbf{N}_{\exists}^I$ rule with the term $\Phi(t)$ then the $\mathbf{N}_h$ rule.

For $\mathbf{N}_{=}^E$ we know $\Phi(\Gamma) \vdash_K^- \Phi(s =_\sigma t)$ and $\Phi(\Gamma) \vdash_K^- \Phi(us)$ by the inductive hypothesis and show $\Phi(\Gamma) \vdash_K^- \Phi(ut)$. Using the definition of Φ and the fact that it is compositional we know that $\Phi(s =_\sigma t) = \Phi(s) =_\sigma \Phi(t)$, $\Phi(us) = \Phi(u)\Phi(s)$, and $\Phi(ut) = \Phi(u)\Phi(t)$. Hence we have $\Phi(\Gamma) \vdash_K^- \Phi(u)\Phi(t)$ by applying the $\mathbf{N}_{=}^E$ rule.

In the $\mathbf{N}_{\exists}^I$ case we have $\Phi(\Gamma) \vdash_K^- \Phi(st)$ by the inductive hypothesis and want to show that $\Phi(\Gamma) \vdash_K^- \Phi(\exists_\sigma s)$. Using the definition of Φ and the fact that it is compositional we know that $\Phi(st) = \Phi(s)\Phi(t)$, and that $\Phi(\exists_\sigma s) = \exists_\sigma \Phi(s)$. Thus using the assumption, we obtain $\Phi(\Gamma) \vdash_K^- \exists_\sigma \Phi(s)$ by applying the $\mathbf{N}_{\exists}^I$ with the term $\Phi(t)$. □

From Theorem 4.2 and Lemma 4.4 we can directly infer that Kuroda's negative translation extends to higher-order formulas moreover, that Glivenko's theorem extends to higher-order formulas with no $\forall$ quantifiers.

Theorem 4.5 (Kuroda). *For all finite $\Gamma \subseteq \Lambda_o$ and $s \in \Lambda_o$, if $\Gamma \vdash_K s$, then $\Phi(\Gamma) \vdash_J \neg\neg\Phi(s)$.*

Proof. Apply Theorem 4.2 and Lemma 4.4. □

Corollary 4.6 (Glivenko). *For all finite $\Gamma \subseteq \Lambda_o^-$ and $s \in \Lambda_o^-$, if $\Gamma \vdash_K s$, then $\Gamma \vdash_J \neg\neg s$.*

Proof. Apply Theorem 4.5 noting that $\Phi(\Gamma) = \Gamma$ and $\Phi(s) = s$ since $\Gamma \subseteq \Lambda_o^-$ and $s \in \Lambda_o^-$. □

We now prove that the following does not hold: For all finite $\Gamma \subseteq \Lambda_o$ and $s \in \Lambda_o$, $\Gamma \vdash_K s$ implies $\Gamma \vdash_J \neg\neg s$. We prove it does not hold by giving a formula s for which $\Gamma \vdash_K s$ and $\Gamma \not\vdash_J \neg\neg s$. Namely, we consider the formula $\forall x.fx \vee \neg fx$ where f is a variable of type ιo and x is a variable of type ι.

Lemma 4.7. *Let f be a variable of type ιo and x be a variable of type ι.*

$$\not\vdash_J \neg\neg\forall x.fx \vee \neg fx$$

Proof. Note that $\neg\neg\forall x.fx \vee \neg fx$ is a first-order formula. A first-order Kripke counter-model is given in Chapter 4 of [6]. In category theoretic terms, the counter-model is given by interpreting ι as a presheaf I over ω (ordered by $\leq$) where $In := \{1, \ldots, n+1\}$ and f is interpreted as the presheaf $fn := \{1, \ldots, n\}$ (both with the obvious inclusion maps). Since every presheaf category is a topos and hence a model of intuitionistic higher-order logic (see [12]), we also know $\not\vdash_J \neg\neg\forall x.fx \vee \neg fx$. □

Theorem 4.8 (Failure of Glivenko with $\forall$)**.** *There exists $s \in \Lambda_o$ such that $\vdash_K s$ and $\nvdash_J \neg\neg s$.*

Proof. Let f be a variable of type ιo and x be a variable of type ι. The statement $\vdash_K \forall x.fx \vee \neg fx$ is easily derivable using the $\mathbf{N}_c$ rule. By Lemma 4.7 we know $\nvdash_J \neg\neg\forall x.fx \vee \neg fx$. ☐

5. Extensionality

There are different extensionality properties one may or may not include in simple type theory. Four such properties are explored in [3]. We briefly review these properties.

η: *η-extensionality* is the property that every element of a function type is a λ-abstraction. The property can be expressed as a set of formulas $\forall f : \sigma\tau.f =_{\sigma\tau} \lambda x.fx$ where σ, τ range over types. An equivalent formulation is to use η-conversion. If x is not free in s, then the term $\lambda x.sx$ is called an *η-redex* with *η-reduct* s. We say s *η-reduces to* t (and write $s \rightarrow_\eta t$) if a subterm of s is an η-redex such that t is the result of replacing this subterm by its η-reduct. We define $s \sim_\eta t$ to be the least equivalence relation containing $\rightarrow_\eta$.

ξ: *ξ-extensionality* is the property that two λ-abstractions are equal if their bodies always have the same value. The property can be expressed as a set of formulas $\forall f : \sigma\tau.\forall g : \sigma\tau.(\forall x : \sigma.fx =_\tau gx) \rightarrow (\lambda x.fx) =_{\sigma\tau} (\lambda x.gx)$ where σ, τ range over types.

$\mathfrak{f}$: *Functional extensionality* is the property that two functions are equal if they are equal on all arguments. The property can be expressed as a set of formulas $\forall f : \sigma\tau.\forall g : \sigma\tau.(\forall x : \sigma.fx =_\tau gx) \rightarrow f =_{\sigma\tau} g$ where σ, τ range over types.

$\mathfrak{b}$: *Boolean extensionality* is the property that there are only two elements of type o.

In this paper we also consider a fifth form of extensionality.

$\mathfrak{p}$: *Propositional extensionality* is the property that two elements of type o are equal if they are equivalent. The property can be expressed as the formula $\forall p : o.\forall q : o.(p \rightarrow q) \rightarrow (q \rightarrow p) \rightarrow p = q$.

In a classical setting, Boolean extensionality and propositional extensionality are equivalent. Hence we can use $\mathfrak{b}$ and $\mathfrak{p}$ interchangeably for classical simple type theory. In particular, in all the results of [3] $\mathfrak{b}$ can be replaced by $\mathfrak{p}$. Furthermore, the assumption of Boolean extensionality in the intuitionistic setting forces the logic to be classical. Since our goal is to relate the intuitionistic and classical versions of simple type theory, only propositional extensionality will be considered in this paper.

$$\mathbf{N}_\eta \ \frac{\Gamma \vdash s \qquad s \sim_\eta t}{\Gamma \vdash t}$$

$$\mathbf{N}_\xi \ \frac{\Gamma \vdash s^x_y =_\tau t^x_y}{\Gamma \vdash (\lambda x.s) =_{\sigma\tau} (\lambda x.t)} \ y \notin \mathcal{N}\Gamma \cup \mathcal{N}(\lambda x.s) \cup \mathcal{N}(\lambda x.t)$$

$$\mathbf{N}_\mathfrak{f} \ \frac{\Gamma \vdash sy =_\tau ty}{\Gamma \vdash s =_{\sigma\tau} t} \ y \notin \mathcal{N}\Gamma \cup \mathcal{N}(s) \cup \mathcal{N}(t) \qquad \mathbf{N}_\mathfrak{p} \ \frac{\Gamma, s \vdash t \qquad \Gamma, t \vdash s}{\Gamma \vdash s =_o t}$$

FIGURE 2. Extensionality rules

In [3] it is proven that functional extensionality is exactly the combination of η-extensionality and ξ-extensionality. Thus there are eight versions of classical simple type theory relative to extensionality principles. In [3] the eight possibilities are named using the indexes β, $\beta\eta$, $\beta\xi$, $\beta\mathfrak{f}$, $\beta\mathfrak{p}$, $\beta\eta\mathfrak{p}$, $\beta\xi\mathfrak{p}$ and $\beta\mathfrak{f}\mathfrak{p}$ (except $\mathfrak{b}$ was used instead of $\mathfrak{p}$). For each $* \in \{\beta, \beta\eta, \beta\xi, \beta\mathfrak{f}, \beta\mathfrak{p}, \beta\eta\mathfrak{p}, \beta\xi\mathfrak{p}, \beta\mathfrak{f}\mathfrak{p}\}$ an appropriate class of models $\mathfrak{M}_*$ is defined and a corresponding (sound and complete) classical natural deduction calculus $\mathfrak{NK}_*$ is given in [3]. In the context of the present paper, we can obtain such natural deduction calculi by adding some or all of the additional extensionality rules given in Figure 2. We can also define an intuitionistic natural deduction calculus for each $*$ in the same way. Each of these defines a variant of simple type theory.

- $\Gamma \vdash^*_K s$ (where Γ is a finite subset of Λ_o and $s \in \Lambda_o$) holds when derivable using all the rules in Figure 1 and the rules indicated by $*$ in Figure 2. For example, if $*$ is $\beta\eta\mathfrak{p}$, then we include the rules $\mathbf{N}_\eta$ and $\mathbf{N}_\mathfrak{p}$ from Figure 2.
- $\Gamma \vdash^{*-}_K s$ (where Γ is a finite subset of Λ^-_o and $s \in \Lambda^-_o$) holds when derivable using the rules in Figure 2 indicated by $*$ (restrict to terms in Λ^-) and all the rules in Figure 1 (restricted to terms in Λ^-) except $\mathbf{N}^I_\forall$ and $\mathbf{N}^E_\forall$.
- $\Gamma \vdash^*_J s$ (where Γ is a finite subset of Λ_o and $s \in \Lambda_o$) holds when derivable using the rules in Figure 2 indicated by $*$ and all the rules in Figure 1 except $\mathbf{N}_c$.

Definition 5.1. *Let* $* \in \{\beta, \beta\eta, \beta\xi, \beta\mathfrak{f}, \beta\mathfrak{p}, \beta\eta\mathfrak{p}, \beta\xi\mathfrak{p}, \beta\mathfrak{f}\mathfrak{p}\}$ *be given. We say* $*$ *satisfies the* Kuroda property *if for every finite* $\Gamma \subseteq \Lambda_o$ *and* $s \in \Lambda_o$*, if* $\Gamma \vdash^*_K s$*, then* $\Phi(\Gamma) \vdash^*_J \neg\neg\Phi(s)$*. We say* $*$ *satisfies the* Glivenko property *if for every finite* $\Gamma \subseteq \Lambda^-_o$ *and* $s \in \Lambda^-_o$*, if* $\Gamma \vdash^*_K s$*, then* $\Gamma \vdash^*_J \neg\neg s$*.*

We now prove that each case without ξ-extensionality satisfies the Kuroda and Glivenko properties. This implies the universal quantifier cannot be defined from the other logical constants in simple type theory without ξ-extensionality.

Theorem 5.2 (Kuroda and Glivenko without ξ-extensionality)**.** *For each* $* \in \{\beta, \beta\eta, \beta\mathfrak{p}, \beta\eta\mathfrak{p}\}$*,* $*$ *has the Kuroda property and the Glivenko property.*

Proof. Theorem 4.5 and Corollary 4.6 mean precisely that β has both properties. One can prove the Kuroda property for $*$ by proving the obvious modification of Theorem 4.2 and Lemma 4.4. In particular, one can prove that for all finite $\Gamma \subseteq \Lambda_o^-$ and $s \in \Lambda_o^-$, if $\Gamma \vdash_K^{*-} s$, then $\Gamma \vdash_J^* \neg\neg s$ by induction on the derivation of $\Gamma \vdash_K^{*-} s$. There are possibly two new cases to consider. First, the $\mathbf{N}_\eta$ case follows from the fact that if $s \sim_\eta t$, then $\neg\neg s \sim_\eta \neg\neg t$. Second, the $\mathbf{N}_\mathfrak{p}$ case follows from the derivability of

$$\vdash_J^{\beta\mathfrak{p}} \forall p : o.\forall q : o.(p \to \neg\neg q) \to (q \to \neg\neg p) \to \neg\neg(p =_o q)$$

The Glivenko property for $*$ follows from the Kuroda property. □

Finally, we prove the failure of the Kuroda and Glivenko properties in the presence of ξ-extensionality.

Theorem 5.3 (Failure of Kuroda and Glivenko with ξ-extensionality)**.** *For each* $* \in \{\beta\xi, \beta\mathfrak{f}, \beta\xi\mathfrak{p}, \beta\mathfrak{f}\mathfrak{p}\}$*,* $*$ *has neither the Kuroda nor the Glivenko property.*

Proof. It suffices to prove $*$ does not have the Glivenko property. Let f be a variable of type ιo and let $\top$ be the formula $\bot \to \bot$. Let s be the formula

$$(\lambda x.\top) \neq_{\iota o} (\lambda x.fx \vee \neg fx) \to \exists x.\top \neq (fx \vee \neg fx)$$

Note that s does not use the logical constant $\forall_\sigma$ and so $s \in \Lambda_o^-$. We prove two claims:

(1) $\vdash_K^* s$.
(2) $\vdash_J^{\beta\mathfrak{f}\mathfrak{p}} (\neg\neg s) \to \neg\neg(\forall x.fx \vee \neg fx)$.

Note that $\nvdash_J^{\beta\mathfrak{f}\mathfrak{p}} \neg\neg\forall x.fx \vee \neg fx$. The proof of this fact is the same as the proof of Lemma 4.7 since every topos is a model of intuitionistic higher-order logic with functional and propositional extensionality. Using the second claim we conclude $\nvdash_J^* \neg\neg s$. Consequently, $*$ does not have the Glivenko property.

First, we prove $\vdash_K^* s$. Using $\mathbf{N}_\to^I$ and $\mathbf{N}_c$ it is enough to prove $\Gamma \vdash_K^* \bot$ where Γ contains exactly the two formulas

$$(\lambda x.\top) \neq_{\iota o} (\lambda x.fx \vee \neg fx) \text{ and } \neg\exists x.\top \neq (fx \vee \neg fx).$$

Clearly we have

$$\Gamma, \top \neq (fx \vee \neg fx) \vdash_K^* \exists x.\top \neq (fx \vee \neg fx)$$

using $\mathbf{N}_h$ and $\mathbf{N}_\exists^I$. By $\mathbf{N}_\to^E$ and $\mathbf{N}_h$ we have

$$\Gamma, \top \neq (fx \vee \neg fx) \vdash_K^* \bot$$

and so by $\mathbf{N}_c$ we have $\Gamma \vdash_K^* \top = (fx \vee \neg fx)$. By $\mathbf{N}_\xi$ (or $\mathbf{N}_\mathfrak{f}$ and $\mathbf{N}_\beta$) we have

$$\Gamma \vdash_K^* (\lambda x.\top) =_{\iota o} (\lambda x.fx \vee \neg fx).$$

Using $\mathbf{N}_\to^E$ and $\mathbf{N}_h$ we have $\Gamma \vdash_K^* \bot$ as desired.

Finally, we prove $\vdash_J^{\beta\mathfrak{f}\mathfrak{p}} (\neg\neg s) \to \neg\neg(\forall x.fx \vee \neg fx)$. It suffices to prove $\Gamma \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} \bot$ where Γ contains exactly the two formulas $\neg\neg s$ and $\neg\forall x.fx \vee \neg fx$. Using $\mathbf{N}_h$, $\mathbf{N}_\to^E$ and $\mathbf{N}_\to^I$ it is enough to prove $\Gamma, s \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} \bot$. We argue that $\Gamma, s \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} (\lambda x.\top) \neq (\lambda x.fx \vee \neg fx)$ in the following steps.

(1) $\Gamma, s, (\lambda x.\top) = (\lambda x.fx \vee \neg fx) \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} \top$ by $\mathbf{N}_\to^I$ and $\mathbf{N}_h$.

(2) $\Gamma, s, (\lambda x.\top) = (\lambda x.fx \vee \neg fx) \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} (\lambda x.\top)x$ by $\mathbf{N}_\beta$ and (1).

(3) $\Gamma, s, (\lambda x.\top) = (\lambda x.fx \vee \neg fx) \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} (\lambda x.fx \vee \neg fx)x$ by $\mathbf{N}_h$, $\mathbf{N}_=^E$ and (2).

(4) $\Gamma, s, (\lambda x.\top) = (\lambda x.fx \vee \neg fx) \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} fx \vee \neg fx$ by $\mathbf{N}_\beta$ and (3).

(5) $\Gamma, s, (\lambda x.\top) = (\lambda x.fx \vee \neg fx) \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} \forall x.fx \vee \neg fx$ by $\mathbf{N}_\forall^I$ and (4).

(6) $\Gamma, s, (\lambda x.\top) = (\lambda x.fx \vee \neg fx) \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} \bot$ by $\mathbf{N}_h$, $\mathbf{N}_\to^E$ and (5).

(7) $\Gamma, s \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} (\lambda x.\top) \neq (\lambda x.fx \vee \neg fx)$ by $\mathbf{N}_\to^I$ and (6).

Since s is $(\lambda x.\top) \neq_{\iota o} (\lambda x.fx) \to \exists x.\top \neq fx$ we know

$$\Gamma, s \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} \exists x.\top \neq (fx \vee \neg fx)$$

by $\mathbf{N}_h$ and $\mathbf{N}_\to^E$. By $\mathbf{N}_\exists^E$ is now suffices to prove

$$\Gamma, s, \top \neq (fx \vee \neg fx) \vdash_J^{\beta\mathfrak{f}\mathfrak{p}} \bot.$$

By $\mathbf{N}_h$, $\mathbf{N}_w$ and $\mathbf{N}^E_\to$ it suffices to prove

$$\top \neq (fx \vee \neg fx) \vdash^{\beta\mathfrak{f}\mathfrak{p}}_J \top = (fx \vee \neg fx).$$

We prove this in the following steps.

(1) $\vdash^{\beta\mathfrak{f}\mathfrak{p}}_J \top$ by $\mathbf{N}^I_\to$ and $\mathbf{N}_h$.

(2) $\top \neq (fx \vee \neg fx), fx \vdash^{\beta\mathfrak{f}\mathfrak{p}}_J fx \vee \neg fx$ by $\mathbf{N}^{IL}_\vee$.

(3) $\top \neq (fx \vee \neg fx), fx \vdash^{\beta\mathfrak{f}\mathfrak{p}}_J \top = (fx \vee \neg fx)$ by $\mathbf{N}_\mathfrak{p}$, $\mathbf{N}_w$, (1) and (2).

(4) $\top \neq (fx \vee \neg fx), fx \vdash^{\beta\mathfrak{f}\mathfrak{p}}_J \bot$ by $\mathbf{N}_h$, $\mathbf{N}^E_\to$ and (3).

(5) $\top \neq (fx \vee \neg fx) \vdash^{\beta\mathfrak{f}\mathfrak{p}}_J \neg fx$ by $\mathbf{N}^I_\to$ and (4).

(6) $\top \neq (fx \vee \neg fx) \vdash^{\beta\mathfrak{f}\mathfrak{p}}_J fx \vee \neg fx$ by $\mathbf{N}^{IR}_\vee$ and (5).

(7) $\top \neq (fx \vee \neg fx) \vdash^{\beta\mathfrak{f}\mathfrak{p}}_J \top = (fx \vee \neg fx)$ by $\mathbf{N}_\mathfrak{p}$, $\mathbf{N}_w$, (1) and (6).

□

Since the versions of simple type theory with ξ-extensionality do not satisfy the Glivenko property, it is possible that universal quantifiers can be defined from the other logical constants in the intuitionistic setting. Andrews [2] defines $\forall_\sigma$ as $\lambda f.f = \lambda x.\top$ (where $\top$ is a provable formula) in a classical simple type theory with full extensionality. It is easy to check that the term provides an appropriate definition of universal quantification in an intuitionistic version of simple type theory with functional and propositional extensionality ($\beta\mathfrak{f}\mathfrak{p}$). In fact, a similar definition is given in an intuitionistic simple type theory with functional and propositional extensionality in [12]. One can modify the definition to be $\lambda f.(\lambda x.fx) = \lambda x.\top$ and obtain a definition of $\forall_\sigma$ in the intuitionistic simple type theory for $\beta\xi\mathfrak{p}$.[1] This leaves open the possibility that universal quantifiers are definable from the other logical constants in the two remaining intuitionistic simple type theories for $\beta\xi$ and $\beta\mathfrak{f}$. We leave the resolution of this question for future work.

[1]Florian Rabe recently pointed out that η is not needed to define $\forall_\sigma$ in a private communication.

6. Conclusion

We have proven that the Glivenko result (for formulas without $\forall$) and the Kuroda translation extend to simple type theory without ξ-extensionality. As a consequence, $\forall$ is not definable from the other logical constants in the absence of ξ-extensionality. Also, we have proven that the results do not extend to simple type theory once ξ-extensionality is included.

References

[1] Peter B. Andrews. Resolution in type theory. *Journal of Symbolic Logic*, 36:414–432, 1971.

[2] Peter B. Andrews. *An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof*. Kluwer Academic Publishers, 2nd edition, 2002.

[3] Christoph Benzmüller, Chad E. Brown, and Michael Kohlhase. Higher-order semantics and extensionality. *Journal of Symbolic Logic*, 69:1027–1088, 2004.

[4] Alonzo Church. A formulation of the simple theory of types. *Journal of Symbolic Logic*, 5:56–68, 1940.

[5] Gilda Ferreira and Paulo Oliva. On various negative translations. In Steffen van Bakel, Stefano Berardi, and Ulrich Berger, editors, *Proceedings of CL&C 2010*, volume 47 of *EPTCS*, pages 21–33, 2010.

[6] Melvin Fitting. *Intuitionistic Logic, Model Theory and Forcing*. North-Holland Publishing Co., Amsterdam, 1969.

[7] Gerhard Gentzen. Über das Verhältnis zwischen intuitionistischer und klassischer Arithmetik. *Archiv für mathematische Logik und Grundlagenforschung*, 16:119–132, 1974. received in 1933.

[8] Valery Glivenko. Sur quelques points de la logique de M. Brouwer. *Bulletins de la classe des sciences*, 15:183–188, 1929.

[9] Kurt Gödel. Zur intuitionistischen Arithmetik und Zahlentheorie. *Ergebnisse eines mathematischen Kolloquiums*, 4:34–38, 1933.

[10] Andrey Kolmogorov. O principe tertium non datur. *Matematiceskij Sbornik*, 32:646–667, 1925. Translated as "On the principle of excluded middle" in [14], pp. 416–437.

[11] Sigekatu Kuroda. Intuitionistische Untersuchungen der formalistischen Logik. *Nagoya Mathematical Journal*, 2:35–47, 1951.

[12] Joachim Lambek and Philip Scott. *Introduction to higher order categorical logic*. Cambridge University Press, Cambridge, UK, 1986.

[13] Christine Rizkallah. Proof representations for higher-order logic. Master's thesis, Universität des Saarlandes, December 2009.

[14] Jean van Heijenoort. *From Frege to Gödel. A Source Book in Mathematical Logic 1879–1931*. Harvard University Press, Cambridge, Massachusetts, 1967.

[15] Konrad Zdanowski. On second order intuitionistic propositional logic without a universal quantifier. *Journal of Symbolic Logic*, 74:157–167, 2009.