Construction of a canonical model for a first-order non-Fregean logic with a connective for reference and a total truth predicate

STEFFEN LEWITZKA*, *Universidade Federal da Bahia – UFBA, Instituto de Matemática, Departamento de Ciência da Computação, Campus de Ondina, 40170-110 Salvador – BA, Brazil.*

Abstract

Logics with quantifiers that range over a model-theoretic universe of propositions are interesting for several applications. For example, in the context of epistemic logic the knowledge axioms can be expressed by the single sentences $\forall x.(K_i x \rightarrow x)$, and in a truth-theoretical context an analogue to Tarski's T-scheme can be expressed by the single axiom $\forall x.(x : true \leftrightarrow x)$. In this article, we consider a first-order non-Fregean logic, originally developed by Sträter, which has a total truth predicate and is able to model propositional self-reference. We extend this logic by a connective '$<$' for propositional reference and study semantic aspects. $\varphi < \psi$ expresses that the proposition denoted by formula ψ says something about (refers to) the proposition denoted by φ. This connective is related to a syntactical reference relation on formulas and to a semantical reference relation on the propositional universe of a given model. Our goal is to construct a *canonical model*, i.e. a model that establishes an order-isomorphism from the set of sentences (modulo alpha-congruence) to the universe of propositions, where syntactical and semantical reference are the respective orderings. The construction is not trivial because of the impredicativity of quantifiers: the bound variable in $\exists x.\varphi$ ranges over all propositions, in particular over the proposition denoted by $\exists x.\varphi$ itself. Our construction combines ideas coming from Sträter's dissertation with the algebraic concept of a *canonical domain*, which is introduced and studied in this article.

Keywords: non-Fregean logic, propositional quantifiers, impredicativity, propositional (self-) reference, truth theory

1 Introduction

$\in_T$-Logic was originally designed by Werner Sträter [7] as a theory of propositional self-reference and truth. The object language contains the classical propositional connectives, which are complemented by a connective for propositional identity $\equiv$, operators ($:true$ and $:false$) for a truth predicate of the object language, and a first-order quantifier for existence that ranges over a model-theoretic universe of propositions. An essential feature is the term character of the language — there is no distinction between terms and formulas. A model consists of a propositional universe M, which is divided into two disjoint subsets *TRUE* and *FALSE* (the true and the false propositions, respectively), and a semantic function Γ (called Gamma-function) which for any given assignment γ of propositions to variables maps a formula φ to its denotation $\Gamma(\varphi,\gamma) \in M$. In this way, a proposition is not the equivalence class of a formula modulo logical equivalence, but is explicitly given as the

*E-mail: steffen@dcc.ufba.br

doi:10.1093/jigpal/jzr050 Advance Access published 13 January 2012

denotation of a formula. Formulas having the same truth value, or even logically equivalent formulas, may denote different propositions in the ambient model.[1] Propositional self-reference can be managed by means of the identity connective $\equiv$.[2] Let us look at an example. An equation $\varphi \equiv \psi$ is true iff φ and ψ denote the same proposition. The formula $c : true$ denotes the proposition 'c is true'. Thus, if the equation $c \equiv (c : true)$ is satisfied, then c denotes the proposition 'This proposition is true'. That is, the equation defines a truth-teller. Although the formulas c and $c : true$ are logically equivalent, there are models which satisfy the above equation and others which do not. Paradoxical self-referential statements, such as the liar, can also be asserted by equations, e.g. $c \equiv (c : false)$. However, the (classical) truth conditions of a model ensure that such assertions are always false, i.e. the respective equations are unsatisfiable and therefore contradictory formulas in $\in_T$-Logic. Propositions that correspond to semantic antinomies, such as the liar, cannot exist (as elements of a model-theoretic universe). Note that propositions, not sentences, are the primary bearers of truth values. A sentence is said to be true (false) in a given model if it denotes a true (a false) proposition. Since every proposition is either true or false, the operators $: true$ and $: false$ of the object language represent a total truth predicate: in every model, and for every formula φ, either $\varphi : true$ or $\varphi : false$ is true. Moreover, the truth conditions imply that $\varphi : true$ is true iff φ is true, i.e. $\varphi : true \leftrightarrow \varphi$ is valid for any formula φ. That is, an analogue of the Tarski-biconditionals (Tarski's T-scheme) holds and can be expressed in the object language.[3]

The Gamma-function of an $\in_T$-model is required to satisfy certain structural conditions. For instance, it must be ensured that $\Gamma(\varphi[x := \psi], \gamma) = \Gamma(\varphi[x := \chi], \gamma)$, whenever $\Gamma(\psi, \gamma) = \Gamma(\chi, \gamma)$, where $\varphi[x := \psi]$ is the result of substituting all occurrences of the free variable x with ψ. We call this the Substitution Principle. One also expects that equations between alpha-congruent formulas are valid (i.e. are true in all models under all assignments), where formulas φ and ψ are said to be alpha-congruent, notation: $\varphi =_\alpha \psi$, if they differ at most on their bound variables. Moreover, the Gamma-function should satisfy truth conditions, which reflect the intended meaning of the connectives and operators and of the quantifier. The existence of models is not obvious and must be proved. A main result of Sträter's dissertation [7] is the construction of a so-called intensional model where any two sentences, which are not alpha-congruent, denote distinct propositions. That is, an equation $\varphi \equiv \psi$ is true iff φ and ψ are alpha-congruent. Although the existence of such a model seems to be evident, its construction is far from being trivial as we will see below. Sträter also presents an extensional model containing only two propositions: the true proposition and the false proposition. Essential technical improvements and simplifications of Sträter's original definition of $\in_T$-Logic were introduced by Philip Zeitz who studied the framework as an extension of a given underlying (classical) logic in abstract form (see [10]).

In the present article, we consider semantic aspects of $\in_T$-Logic and concentrate on the construction of a canonical model. We also improve some aspects of preceding versions of $\in_T$-Logic. Our presentation is self-contained and does not require any previous knowledge on $\in_T$-Logic. We extend the language by a new connective for reference: $\varphi < \psi$ can be read as 'the proposition denoted by

[1]The proposition denoted by a formula φ in an ambient model $(\mathcal{M}, \gamma)$ (γ is an assignment of propositions to variables) can be represented by the equivalence class $\{\psi \mid (\mathcal{M}, \gamma) \vDash \psi \equiv \varphi\}$. The underlying equivalence relation is, in general, neither finer nor coarser than logically equivalence.

[2]$\varphi \equiv \psi$ expresses that φ and ψ have the same *Bedeutung* (denotation). Each formula of the form $(\varphi \equiv \psi) \rightarrow (\varphi \leftrightarrow \psi)$ is valid (true in all models), but the so-called Fregean Axiom $(\varphi \leftrightarrow \psi) \rightarrow (\varphi \equiv \psi)$ is not. Logics having this feature are called non-Fregean. Non-Fregean logics were introduced by Roman Suszko (see, e.g. [1, 8, 9]). $\in_T$-Logic can be seen as an extension of Suszko's basic non-Fregean logic SCI (the Sentential Calculus with Identity) by operators for the truth predicate and propositional quantifiers. In fact, the axioms of SCI (see [1]) derive from the axiomatization of $\in_T$-Logic given by Zeitz [10]. They also derive from the set of axioms of the epistemic, quantifier-free $\in_T$-style logic presented in [5].

[3]This analogue of the Tarski-biconditionals can also be expressed by the single, valid sentence $\forall x.(x : true \leftrightarrow x)$.

ψ says something about (refers to) the proposition denoted by φ'.[4] In a model $\mathcal{M}$ with universe M, the connective $<$ is interpreted by a semantical reference, i.e. a transitive relation $<^M$ on M. The reference connective was introduced in [3] and has also been studied in a quantifier-free context [4–6]. The semantics of $<$ depends on a suitable syntactical reference relation $\prec$ defined on the set of formulas. We require that $\varphi \prec \psi$ implies the validity (i.e. truth in all models) of formula $\varphi < \psi$. The syntactical reference $\varphi \prec \psi$ stands for the intuitition that 'ψ says something about (refers to) φ'. For example, $x : true$ says that x is true, and $\chi_1 \to \chi_2$ says about χ_1 that it implies χ_2. So the definition of $\prec$ must ensure that $x \prec (x : true)$ and $\chi_1 \prec (\chi_1 \to \chi_2)$. At a first glance, $\varphi \prec \psi$ seems to coincide with the relation 'φ is a proper subformula of ψ'. However, we will see that $\prec$ strictly refines that subformula relation. For instance, the variable x is a proper subformula of $\exists x.(x : true)$. But $\exists x.(x : true)$ does not say anything about x. In order to capture this intuition, the definition of $\prec$ must also ensure that $x \not\prec \exists x.(x : true)$. Our definition of $\prec$ can be informally described as follows: $\varphi \prec \psi$ holds iff there is φ' such that $\varphi' =_\alpha \varphi$ and φ' is a proper subformula of ψ (in particular, $\varphi' \neq \psi$) and every occurrence of a free variable in φ' remains free in ψ. Of course, $\varphi \prec \varphi$ is impossible since φ cannot be alpha-congruent to a proper subformula of itself. Self-reference must be shifted to the semantic level. The existence of a self-referential *proposition* can be forced by an equation of the form $\varphi \equiv \psi$ with $\varphi \prec \psi$. For instance, there is a model $\mathcal{M}$ where $c \equiv (c : true)$ is true, i.e. the sentences c and $(c : true)$ have the same denotation. Since $c \prec (c : true)$, the formula $c < (c : true)$ is true, too. By the Substitution Principle, we can replace $(c : true)$ by a formula with the same denotation, in particular by c. Thus, $c < c$ is true in $\mathcal{M}$. By the semantics of the reference connective, we get $m <^M m$, where $m \in M$ is the propostion denoted by c, and $<^M$ is the semantical reference relation on the universe M of $\mathcal{M}$. That is, c denotes the self-referential proposition m, which in this specific example is a truth-teller. Notice that the class of all well-founded models, i.e. models with no self-referential propositions, can be axiomatized by a single sentence, namely $\neg\exists x.(x < x)$. This would be impossible without reference connective[5].

We saw that syntactical reference implies semantical reference: $\varphi \prec \psi$ implies the validity of formula $\varphi < \psi$, i.e. $(\mathcal{M}, \gamma) \vDash \varphi < \psi$, for all models $\mathcal{M}$ and all assignments γ. Should *every* semantical reference be based on a certain syntactical reference? We could easily impose the following semantic constraint on every model $\mathcal{M}$: For any two formulas φ, ψ and any assignment γ, if $(\mathcal{M}, \gamma) \vDash \varphi < \psi$, then there are formulas φ' and ψ' such that $(\mathcal{M}, \gamma) \vDash (\varphi' \equiv \varphi) \wedge (\psi' \equiv \psi)$ and $\varphi' \prec \psi'$. A model $\mathcal{M}$ satisfying this additional condition is called $<$-intensional (see also [5] for a similar definition). If M is the universe of an $<$-intensional model $\mathcal{M}$, and $<^M$ is the reference relation on M, then $m <^M m'$ implies the existence of an appropriate assignment γ and formulas φ and ψ such that $\varphi \prec \psi$ and $\Gamma(\varphi, \gamma) = m$ and $\Gamma(\psi, \gamma) = m'$. That is, $(\mathcal{M}, \gamma) \vDash \varphi < \psi$. Thus, the syntatical reference relation is mirrored by its semantical counterpart in the interpretation $(\mathcal{M}, \gamma)$, and vice versa; that is, the reference connective is interpreted in accordance with its intended meaning. From an intended model, we require additionally that every proposition is denoted by a sentence, i.e. there are no non-standard elements.[6] Then the $<$-intensional models without non-standard elements are precisely the intended

[4]The noun 'reference' is often used in the literature as a translation of Frege's 'Bedeutung'. We do not adopt this translation here because of its ambiguity. For instance, 'reference' is also used in the sense of 'self-reference' (in german: Selbstreferenz) where it has clearly a different meaning. We prefer to translate 'Bedeutung' as 'denotation' (in some other contexts also as 'meaning'), and we will apply the term 'reference' in contexts where also the term 'self-reference' is meaningful (e.g. reference between sentences, reference between propositions).

[5]One may argue that a formula of the form $\forall x.\varphi$ involves a kind of implicit propositional (self-) reference: the formula asserts that all propositions — including the proposition denoted by $\forall x.\varphi$ itself — have property φ. However, we do not consider such implicit references here. Our reference connective expresses explicit propositional reference in the sense explained above.

[6]A non-standard element is an element of the universe which is not denoted by any sentence.

models. For these intended models we suggest the term *standard model* (see also [5]). Nearly all models of any interest are standard models. For instance, a canonical model as well as a two-element extensional model are standard models. So why we do not require the condition of $<$-intensionality as an additional constraint in the definition of semantics? The reason is a pragmatic one: the condition seems to be too strong for the existence of a complete calculus. We refer the reader to the deductive systems and respective completeness theorems given in [5] and [6], where quantifier-free $\in_T$-style logics with a reference connective are studied. The reference connective is characterized by the following two (schemes of) axioms: $\varphi < \psi$ whenever $\varphi \prec \psi$; and $(\varphi < \psi) \to ((\psi < \chi) \to (\varphi < \chi))$.[7] Note that we cannot add the axiom '$\neg(\varphi < \psi)$ whenever $\varphi \not\prec \psi$' as the counter-example $c < c$ shows (recall that the formula $c < c$ is satisfiable, but $c \not\prec c$). In fact, we have no axiom at hand that negates the existence of models which are not $<$-intensional. In other words, the existence of such non-standard models is the price that we have to pay for the existence of a complete calculus.[8] However, this limitation can be avoided by imposing another semantic constraint. If we require that all models satisfy exactly the same set of equations Φ, then we get a more restrictive logic where an equation is satisfiable iff it is valid (i.e. true in all models under all assignments).[9] In such a logic, we may consider the equivalence classes $\overline{\varphi} = \{\varphi' \mid \varphi \equiv \varphi' \in \Phi\}$ of formulas and define $\overline{\varphi} \prec \overline{\psi} :\Leftrightarrow$ there are $\varphi' \in \overline{\varphi}$ and $\psi' \in \overline{\psi}$ such that $\varphi' \prec \psi'$. Then the following are axioms: $\varphi \equiv \psi$ whenever $\overline{\varphi} = \overline{\psi}$; and $\neg(\varphi \equiv \psi)$ whenever $\overline{\varphi} \neq \overline{\psi}$. The two axioms above concerning reference can now be complemented by the following axiom: $\neg(\varphi < \psi)$ whenever $\overline{\varphi} \not\prec \overline{\psi}$. This corresponds to the following stronger semantic condition: $\overline{\varphi} \prec \overline{\psi}$ iff $\varphi < \psi$ is true in a given model (iff $\varphi < \psi$ is valid). It is obvious that in such a logic all models are $<$-intensional.[10]

The main goal of the present article is to construct a *canonical model* for the $\in_T$-language extended by our reference connective.[11] The new connective not only enriches the expressive power of the language but is also an essential tool for the model construction. We call a model $\mathcal{M}$ canonical if it does not contain non-standard elements, and $\mathcal{M} \vDash \varphi < \psi \Leftrightarrow \varphi \prec \psi$ for all sentences φ, ψ. It follows that a model is canonical iff its Gamma-function establishes an order-isomorphism from the set of sentences modulo alpha-congruence to the propositional universe, where syntactical and semantical reference are the respective orderings. It turns out that a canonical model is in particular intensional in Sträter's sense, i.e. $\mathcal{M} \vDash \varphi \equiv \psi \Leftrightarrow \varphi =_\alpha \psi$, for all sentences φ, ψ. The intensional models constructed by Sträter [7] and Zeitz [10], however, contain non-standard elements and are therefore not canonical.[12]

Why is the construction of a canonical model of interest? In a canonical model, the Gamma-function is a bijection from the set of equivalence classes of sentences modulo alpha-congruence (or equivalently, from the set of normalized sentences[13]) onto the universe

[7]The second axiom, expressing transitivity, is invalid if transitivity of the semantic reference relation of a model is not explicitly required. One can show that in a standard model the transitivity of the semantic reference relation follows from the transitivity of $\prec$ (see the proof of Lemma 3.11 in [5]). In this article as well as in [5] we require transitivity of the semantical reference relation in *all* models.

[8]A complete calculus for our logic will not be given in the present article. A complete sequent calculus and a Hilbert-style calculus for original $\in_T$-Logic are presented in [7] and [10], respectively.

[9]The logic developed in [6] has this property.

[10]These additional constraints concerning reference are applicable to the logic developed in [6].

[11]The term 'canonical model' in logics with Kripke semantics usually refers to a specific model employed for a completeness proof of an underlying deductive system. Note that this is not the intended meaning of that notion in the present article.

[12]Sträter sketches out ideas how to extend his original construction in order to get a model without non-standard elements. The suggested extension, however, is not trivial and turns the construction even more complex.

[13]We will see that in each equivalence class of sentences modulo alpha-congruence there is exactly one sentence having a certain *normal form*. We call such a sentence *normalized*.

of propositions. That is, each proposition can be identified with — up to alpha-congruence — exactly one sentence; the semantic content of a proposition is given by the syntactical form (intension, sense) of the corresponding sentence together with a truth value. A canonical model gives rise to a universe of sentences which is closed under the truth condition of the quantifier: the truth of a quantified sentence $\exists x.\varphi$ can be witnessed by a *sentence* ψ of the universe such that $\varphi[x:=\psi]$ is true. This sentential universe can be seen as a term model which satisfies only the trivial equations (given by alpha-congruence) and is uniquely determined by the truth values of the constant symbols.[14] Such a term model may serve as the starting point for the construction of further standard models that satisfy specific non-trivial equations.[15]

As already pointed out and discussed in [7, 10], the essential difficulty of the construction of intensional models is the impredicativity of quantifiers. In a first attempt, the following strategy for a 'construction' seems to be appealing. For the propositional universe M, we choose the set of normalized sentences which are partially ordered by the semantical reference relation $<^M := \prec$. The Gamma-function is now given by the map that sends each sentence to its normal form — this establishes the desired order-isomorphism. Now, in the second step of the 'construction', we aim to determine the truth values of the elements of M, i.e. of the (normalized) sentences. We try to do this stepwise by induction over a suitable defined rank, which gives rise to a well-founded ordering on formulas. There are several possibilities to define such a rank. A major problem, however, is to apply it to the following truth condition of the quantifier: '$\exists x.\varphi$ is true iff there is a sentence $\psi \in M$ such that $\varphi[x:=\psi]$ is true'. That is, in order to determine the truth value of $\exists x.\varphi \in M$ we already have to know the truth values of all $\psi \in M$ and of $\varphi[x:=\psi]$. Note that the ranks of ψ and of $\varphi[x:=\psi]$ are possibly equal to or greater than the rank of $\exists x.\varphi$. Because of this impredicativity of the quantifier, the above truth condition cannot be applied to determine the truth value of $\exists x.\varphi$. One might ask whether it is sufficient to check the condition '$\varphi[x:=\psi]$ is true' only for those sentences ψ having the property that the ranks of ψ and of $\varphi[x:=\psi]$ are smaller than the rank of $\exists x.\varphi$. More generally, one might ask for the existence of a well-founded ordering on formulas that leads to a suitable notion of 'small sentence' and to a proof of the following kind of assertion: there is some $\psi \in M$ such that $\varphi[x:=\psi]$ is true iff there is some *small* $\psi \in M$ such that $\varphi[x:=\psi]$ is true. Unfortunately, until now all attempts to find such an ordering have failed. Therefore, two different approaches are proposed in [7, 10], each of them leading to the construction of an intensional model. Although both constructions differ essentially from each other they follow the same strategy: the two steps of the above outlined 'construction' are carried out in inverse order. That is, the propositional universe and the truth values of its elements are fixed *before* the Gamma-function is defined. This strategy avoids the problem of the impredicativity of quantifiers. On the other hand, it requires much more sophisticated ideas and machinery to define the Gamma-function. In our construction, which is partially based on ideas developed in [7], we adopt this general strategy as well. For this we introduce the new concept of a *canonical domain* as a partially ordered structure of abstract objects, which have a truth value (either true or false). If we choose a canonical domain as the propositional universe of the model wanted, then the truth values of propositions and the semantical reference relation are already given. The definition of the Gamma-function as an order-isomorphism from the set of normalized sentences (ordered by syntactical reference) to the canonical domain (ordered by semantical reference) then derives from the algebraic properties of the canonical domain and the truth conditions of the connectives, operators and the quantifier.

[14] See the construction of a canonical term model outlined below, after Theorem 5.35.

[15] It would be interesting to study a category of standard models in which the term model would be the initial object.

2 Syntax

The alphabet consists of a (possibly empty) set C of constant symbols, a countable infinite set of variables $V=\{v_0,v_1,v_2,...\}$ which is well-ordered by the given enumeration, connectives for classical negation and disjunction $\neg$, $\vee$, the existential quantifier $\exists$, predicates for truth and falsity $:true$, $:false$, respectively (we use postfix notation), the identity conective $\equiv$, the reference connective $<$, and auxiliary symbols:), (and dot. The set C of constant symbols is viewed as a set of parameters, i.e. for each set C we can define a set of formulas and a respective logic over C.

For our purposes it is desirable that a string $\exists z.\varphi$ is a formula iff the variable z occurs free in the formula φ. For example, $\exists x.c$ and $\exists x.\exists y.(y\vee y)$ should not be formulas. This is ensured by the following definitions.

Definition 2.1
Let C be a set of constant symbols. The set $Expr(C)_0$ is the smallest set X that contains $V\cup C$ and is closed under the following condition (a):
(a) If φ, $\psi\in X$, then $(\varphi:true)$, $(\varphi:false)$, $(\neg\varphi)$, $(\varphi\vee\psi)$, $(\varphi\equiv\psi)$, $(\varphi<\psi)\in X$.
We denote the set of variables occurring in $\chi\in Expr(C)_0$ by $fvar(\chi)$.
The set $Expr(C)_1$ is the smallest set X that contains $Expr(C)_0$ and is closed under condition (a) and under the following condition (b), where $n=0$:
(b) If $\varphi\in Expr(C)_n$ and $x\in fvar(\varphi)$, then $(\exists x.\varphi)\in X$.

Now we suppose that for some $n\geq 1$ the set $Expr(C)_n$ is already defined.

Definition 2.2
The set $fvar(\psi)$ of free variables of an expression $\chi\in Expr(C)_n$ is inductively defined as follows:

$$
\begin{aligned}
&fvar(x)=\{x\},\ \text{for } x\in V\\
&fvar(c)=\varnothing,\ \text{for } c\in C\\
&fvar(\varphi\vee\psi)=fvar(\varphi\equiv\psi)=fvar(\varphi<\psi)=fvar(\varphi)\cup fvar(\psi)\\
&fvar(\neg\varphi)=fvar(\varphi:true)=fvar(\varphi:false)=fvar(\varphi)\\
&fvar(\exists x.\varphi)=fvar(\varphi)\smallsetminus\{x\}
\end{aligned}
$$

Definition 2.3
The set $Expr(C)_{n+1}$ is the smallest set X that contains $Expr(C)_n$ and is closed under the conditions (a) and (b) of Definition 2.1. Finally, the set of all expressions is defined as $Expr(C)=\bigcup_{n<\omega}Expr(C)_n$.

The set of variables (constant symbols) occurring in $\chi\in Expr(C)$ is denoted by $var(\chi)$ (by $con(\chi)$). We put $fcon(\chi):=fvar(\chi)\cup con(\chi)$. The notion of subexpression (or subformula) is defined as usual. We write $\varphi\in subex(\psi)$ in order to express that φ is a subexpression of ψ. φ is a proper subexpression of ψ if $\varphi\in subex(\psi)\smallsetminus\{\psi\}$. Usually, we omit outermost parentheses. We may also omit parentheses respecting the following descending priority of the connectives and the quantifier: $\neg$, $:true$, $:false$, $\vee$, $\equiv$, $<$, $\exists$. For instance, $(\exists x.x)\vee y$ and $\exists x.x\vee y$ are different expressions. We may introduce further connectives and the quantifier 'for all' by means of the following usual abbreviations: $\varphi\wedge\psi:=\neg(\neg\varphi\vee\neg\psi)$, $\forall x.\varphi:=\neg\exists x.\neg\varphi$, $\varphi\rightarrow\psi:=\neg\varphi\vee\psi$, $\varphi\leftrightarrow\psi:=(\varphi\rightarrow\psi)\wedge(\psi\rightarrow\varphi)$.

Definition 2.4
The quantifier rank $R_q(\varphi)$ of a formula $\varphi\in Expr(C)$ is the smallest number n such that $\varphi\in Expr(C)_n$.

Definition 2.5
Let $U \subset V = \{v_0, v_1, \ldots\}$ be a finite subset of the well-ordered set of variables. We define $lub(U)$ (the least upper bound of U) as the smallest variable of the set of all variables which are greater than all elements of U.

Note that $lub(U) \notin U$ and $lub(\varnothing) = v_0$.

2.1 Substitutions

Definition 2.6
A substitution is a function $\sigma : V \cup C \to Expr(C)$. If $A \subseteq V \cup C$ and $\sigma(u) = u$ for all $u \in (V \cup C) \setminus A$, then we write $\sigma : A \to Expr(C)$. If σ is a substitution, $u_0, \ldots, u_n \in V \cup C$ and $\varphi_0, \ldots, \varphi_n \in Expr(C)$, then the substitution $\sigma[u_0 := \varphi_0, \ldots, u_n := \varphi_n]$ is defined by:

$$\sigma[u_0 := \varphi_0, \ldots, u_n := \varphi_n](v) = \begin{cases} \varphi_i & \text{if } v = u_i, \text{for some } i \leq n \\ \sigma(v) & \text{else} \end{cases}$$

The identity substitution $u \mapsto u$ is denoted by ε. Instead of $\varepsilon[u_0 := \varphi_0, \ldots, u_n := \varphi_n]$ we also write $[u_0 := \varphi_0, \ldots, u_n := \varphi_n]$. A substitution σ extends in the following way to a function $[\sigma] : Expr(C) \to Expr(C)$ (we use postfix notation for $[\sigma]$):

$$\begin{aligned}
&u[\sigma] := \sigma(u), \text{ for } u \in V \cup C \\
&(\varphi : true)[\sigma] := \varphi[\sigma] : true \\
&(\varphi : false)[\sigma] := \varphi[\sigma] : false \\
&(\neg\varphi)[\sigma] := \neg\varphi[\sigma] \\
&(\varphi \vee \psi)[\sigma] := \varphi[\sigma] \vee \psi[\sigma] \\
&(\varphi \equiv \psi)[\sigma] := \varphi[\sigma] \equiv \psi[\sigma] \\
&(\varphi < \psi)[\sigma] := \varphi[\sigma] < \psi[\sigma] \\
&(\exists x.\varphi)[\sigma] := \exists y.\varphi[\sigma[x := y]],
\end{aligned}$$

where y is the variable $y := lub(\bigcup\{fvar(\sigma(u)) \mid u \in fcon(\exists x.\varphi)\})$. We say that y is forced by σ w.r.t. $\exists x.\varphi$. For two substitutions σ and τ the composition is the substitution $\sigma \circ \tau$ defined by $(\sigma \circ \tau)(u) = \sigma(u)[\tau]$, for $u \in V \cup C$.

The following properties of substitutions are useful and not hard to show. We omit the proofs.

Lemma 2.7
Let $\varphi \in Expr(C)$ and let σ, τ be substitutions. Then

(i) $fcon(\varphi[\sigma]) = \bigcup\{fcon(\sigma(y)) \mid y \in fcon(\varphi)\}$.
(ii) If $\sigma(u) = \tau(u)$ for all $u \in fcon(\varphi)$, then $\varphi[\sigma] = \varphi[\tau]$.

Lemma 2.8
If $\exists x.\psi$ is an expression and σ is a substitution, then the variable y forced by σ w.r.t. $\exists x.\psi$ is $y = lub(fvar((\exists x.\psi)[\sigma]))$.

Lemma 2.9
Suppose σ, δ are substitutions and $x, y, z \in V$ such that for all $u \in C \cup V \setminus \{x\}$, $y \notin fvar(\sigma(u))$. Then $\sigma[x := y] \circ \delta[y := z] = (\sigma \circ \delta)[x := z]$.

Corollary 2.10
Let $\varphi \in Expr(C)$, $x, y, z \in V$ and let σ, δ be substitutions with $y \notin fvar(\sigma(u))$, for all $u \in fcon(\varphi) \setminus \{x\}$. Then $\varphi[\sigma[x := y] \circ \delta[y := z]] = \varphi[(\sigma \circ \delta)[x := z]]$.

Lemma 2.11
Let $\varphi \in Expr(C)$ and let σ, τ be substitutions. Then $\varphi[\sigma \circ \tau] = \varphi[\sigma][\tau]$.

Corollary 2.12
$\sigma \circ (\tau \circ \delta) = (\sigma \circ \tau) \circ \delta$, for all substitutions σ, τ, δ.

2.2 The alpha-congruence

Definition 2.13
The relation $=_\alpha$ (α-congruence or alpha-congruence) is the smallest equivalence relation on $Expr(C)$ satisfying the following conditions:

- If $\varphi =_\alpha \psi$, then $\neg\varphi =_\alpha \neg\psi$, $\varphi : true =_\alpha \psi : true$, $\varphi : false =_\alpha \psi : false$
- If $\varphi_1 =_\alpha \psi_1$ and $\varphi_2 =_\alpha \psi_2$, then $(\varphi_1 \vee \varphi_2) =_\alpha (\psi_1 \vee \psi_2)$, $(\varphi_1 \equiv \varphi_2) =_\alpha (\psi_1 \equiv \psi_2)$, $(\varphi_1 < \varphi_2) =_\alpha (\psi_1 < \psi_2)$
- If $\exists x.\varphi$ and $\exists y.\psi$ are expressions such that $\varphi =_\alpha \psi[y := x]$, and $x \neq y$ implies $x \notin fvar(\psi)$, then $\exists x.\varphi =_\alpha \exists y.\psi$.

For $\varphi \in Expr(C)$ we denote the equivalence class of φ modulo alpha-congruence by $\overline{\varphi}$.

Two expressions are alpha-congruent iff they differ at most on their bound variables. For instance, $(\exists x.z \equiv x) =_\alpha (\exists y.z \equiv y)$.

Lemma 2.14
If $\varphi =_\alpha \psi$, then $fcon(\varphi) = fcon(\psi)$, $R_q(\varphi) = R_q(\psi)$ and $\varphi[\sigma] = \psi[\sigma]$, for any substitution σ.

Note that applying the identity substitution ε to a formula with quantifiers generally results in a different formula because of renaming of bound variables. However, we may show the following.

Lemma 2.15
For any expression φ and any substitution σ, $\varphi[\varepsilon][\sigma] = \varphi[\sigma]$. In particular, $\varphi[\varepsilon][\varepsilon] = \varphi[\varepsilon]$.

Proof. For all $u \in fcon(\varphi)$ we have $u[\varepsilon][\sigma] = u[\varepsilon \circ \sigma] = \varepsilon(u)[\sigma] = u[\sigma]$. Now the assertion follows from Lemma 2.7. ∎

Definition 2.16
We say that an expression φ is normalized if $\varphi[\varepsilon] = \varphi$. The function $norm : Expr(C) \to Expr(C)$ is defined by $\varphi \mapsto \varphi[\varepsilon]$.

By Lemma 2.15, $norm(\varphi) = \varphi[\varepsilon]$ is normalized, for any expression φ.

Lemma 2.17
$\varphi \in Expr(C)$ is normalized iff for all $\psi \in subex(\varphi)$ of the form $\psi = \exists x.\psi'$ it holds that $x = lub(fvar(\exists x.\psi'))$.

Proof. The direction from left to right can be shown by induction on $R_q(\varphi)$. The direction from right to left can be shown by induction on φ. ∎

Corollary 2.18
An expression is normalized iff all its subexpressions are normalized. Two normalized alpha-congruent expressions are equal.

Lemma 2.19
If $\varphi \in Expr(C)$ and σ is a substitution such that $\sigma(u)$ is normalized for all $u \in fcon(\varphi)$, then $\varphi[\sigma]$ is normalized.

Proof. By hypothesis, $(\sigma \circ \varepsilon)(u) = \sigma(u)[\varepsilon] = \sigma(u)$, for any $u \in fcon(\varphi)$. By Lemma 2.7, $\varphi[\sigma][\varepsilon] = \varphi[\sigma \circ \varepsilon] = \varphi[\sigma]$. Thus, $\varphi[\sigma]$ is normalized. ∎

Lemma 2.20
For any expression φ, $\varphi[\varepsilon] =_\alpha \varphi$.

Corollary 2.21
Let $\varphi \in Expr(C)$. Then $norm(\varphi)$ is the unique formula which is normalized and alpha-congruent with φ.

Proof. The assertion concerning existence follows from Lemma 2.15 and Lemma 2.20. Uniqueness follows from Corollary 2.18. ∎

Lemma 2.22
For all $\varphi, \psi \in Expr(C)$: $norm(\varphi) = norm(\psi) \Leftrightarrow \varphi =_\alpha \psi$.

Proof. The implication from right to left follows from Lemma 2.14. The implication from left to right is a consequence of the equations $\varphi =_\alpha norm(\varphi) = norm(\psi) =_\alpha \psi$. ∎

Lemma 2.23
Suppose $\varphi =_\alpha \psi$ and let σ, τ be substitutions such that $\sigma(u) =_\alpha \tau(u)$, for all $u \in fcon(\varphi)$. Then:

(i) $\varphi[\sigma] =_\alpha \varphi[\tau]$ and $\psi[\sigma] =_\alpha \psi[\tau]$.
(ii) $\varphi[\sigma] =_\alpha \psi[\tau]$.

2.3 The syntactical reference

We define a binary relation $\prec$ on the set of expressions which refines the subexpression relation. $\varphi \prec \psi$ will capture the intuitive notion of 'expression ψ says something about expression φ' or 'expression ψ refers to expression φ'.

Definition 2.24
For $\varphi, \psi \in Expr(C)$ we define $\varphi \prec \psi :\Leftrightarrow$ there are $x \in V$ and $\psi' \in Expr(C) \setminus \{x\}$ such that $x \in fvar(\psi')$ and $\psi'[x := \varphi] =_\alpha \psi$. The relation $\prec$ is called syntactical reference.

Remark 2.25
Note that in the above definition ψ' and x are not unique. If $y \neq x$ is a variable not occurring in ψ', then $\psi'' := \psi'[x := y]$ is an expression containing the free variable y and $\psi''[y := \varphi] = \psi'[x := \varphi] =_\alpha \psi$. Thus, y and ψ'' witness $\varphi \prec \psi$, too.

Example 2.26
$\chi \prec (\chi : true)$. $x \in subex(\exists x.x)$, but $x \not\prec \exists x.x$. Let $\psi = \forall y.(y \to (y \equiv x))$ and $\varphi = \exists x.(x \wedge \psi)$. Then $\psi \in subex(\varphi)$, but $\psi \not\prec \varphi$ since $x \in fvar(\psi)$. Let $\chi := (x \wedge c) \vee \exists x.x$. The variable x occures free and bound in χ. Observe that the free occurrence of x in χ remains free in $((\forall x.x \equiv x) \wedge \chi)$. Thus, $\chi \prec ((\forall x.x \equiv x) \wedge \chi)$.

The syntactical reference can be characterized by the following Lemma (choose $\sigma := \varepsilon$ and consider a *proper* subexpression φ' of ψ) which is shown by induction on formulas.

LEMMA 2.27
Suppose that φ and ψ are expressions. The following two conditions are equivalent:

(i) There is some expression ψ', a variable $x \in fvar(\psi')$ and a substitution σ such that $\psi'[\sigma[x:=\varphi]] =_\alpha \psi$.
(ii) φ is alpha-congruent to a subexpression φ' of ψ, and every free occurrence of a variable in φ' remains free in ψ.

Recall that $\overline{\varphi}$ denotes the equivalence class of φ modulo alpha-congruence.

DEFINITION 2.28
Let φ, ψ be expressions. We define: $\overline{\varphi} \prec \overline{\psi} :\Leftrightarrow \varphi \prec \psi$.

REMARK 2.29
Let us show that $\overline{\varphi} \prec \overline{\psi}$ is well-defined. Suppose $\varphi \prec \psi$, and let $\varphi_1 =_\alpha \varphi$ and $\psi_1 =_\alpha \psi$. Since $\varphi \prec \psi$, there is a variable x and an expression $\psi' \neq x$ such that $x \in fvar(\psi')$ and $\psi'[x:=\varphi] =_\alpha \psi$. Then $\psi_1 =_\alpha \psi =_\alpha \psi'[x:=\varphi] =_\alpha \psi'[x:=\varphi_1]$, by Lemma 2.23. Hence, ψ' and φ_1 witness $\varphi_1 \prec \psi_1$.

The proofs of the next two results are not difficult but technically complex, so we will omit them.

LEMMA 2.30
If $\varphi \prec \psi$ and σ is a substitution, then $\varphi[\sigma] \prec \psi[\sigma]$.

LEMMA 2.31
The syntactical reference $\prec$ is a transitive relation on $Expr(C)$.

3 Semantics

DEFINITION 3.1
A model $\mathcal{M} = (M, TRUE, FALSE, <^{\mathcal{M}}, \Gamma)$ is given by the following ingredients:

- a set M of propositions (which are generally given as abstract entities without any inner structure)
- a set $TRUE \subseteq M$ of true propositions and a set $FALSE \subseteq M$ of false propositions such that $M = TRUE \cup FALSE$ and $TRUE \cap FALSE = \varnothing$
- a transitive relation $<^{\mathcal{M}} \subseteq M \times M$ for semantical reference
- a semantic function, called Gamma-function, $\Gamma : Expr(C) \times M^V \to M$ that maps an expression φ to its denotation $\Gamma(\varphi, \gamma) \in M$. Γ depends on assignments $\gamma : V \to M$ of propositions to variables. If $\gamma \in M^V$ is an assignment and σ is a substitution, then $\gamma\sigma \in M^V$ denotes the assignment defined by $x \mapsto \Gamma(\sigma(x), \gamma)$. If $x \in V, m \in M$, then γ_x^m is the assignment defined by

$$\gamma_x^m(y) := \begin{cases} m, & \text{if } x = y \\ \gamma(y), & \text{else.} \end{cases}$$

The Gamma-function satisfies the following structure conditions:

(EP) For all $x \in V$ and all assignments $\gamma \in M^V$, $\Gamma(x, \gamma) = \gamma(x)$. (Extension Property)
(CP) If $\varphi \in Expr(C)$, $\gamma, \gamma' \in M^V$, and $\gamma(x) = \gamma'(x)$ for all $x \in fvar(\varphi)$, then $\Gamma(\varphi, \gamma) = \Gamma(\varphi, \gamma')$. (Coincidence Property)[16]
(SP) If $\varphi \in Expr(C)$, $\gamma \in M^V$ and $\sigma : V \to Expr(C)$ is a substitution, then $\Gamma(\varphi[\sigma], \gamma) = \Gamma(\varphi, \gamma\sigma)$. (Substitution Property)

[16] If $fvar(\varphi) = \varnothing$, then (CP) justifies to write $\Gamma(\varphi)$ instead of $\Gamma(\varphi, \gamma)$.

(RP) If $\varphi \prec \psi$, then $\Gamma(\varphi,\gamma) <^{\mathcal{M}} \Gamma(\psi,\gamma)$, for all $\varphi,\psi \in Expr(C)$ and all assignments γ. (Reference Property)

The Gamma-function satisfies the following truth conditions. For all $\varphi,\psi \in Expr(C)$ and all assignments $\gamma \in M^V$:

(i) $\Gamma(\varphi:true,\gamma) \in TRUE \Leftrightarrow \Gamma(\varphi,\gamma) \in TRUE$
(ii) $\Gamma(\varphi:false,\gamma) \in TRUE \Leftrightarrow \Gamma(\varphi,\gamma) \in FALSE$
(iii) $\Gamma(\varphi \vee \psi,\gamma) \in TRUE \Leftrightarrow \Gamma(\varphi,\gamma) \in TRUE$ or $\Gamma(\psi,\beta) \in TRUE$
(iv) $\Gamma(\neg\varphi,\gamma) \in TRUE \Leftrightarrow \Gamma(\varphi,\gamma) \notin TRUE$
(v) $\Gamma(\varphi \equiv \psi,\gamma) \in TRUE \Leftrightarrow \Gamma(\varphi,\gamma) = \Gamma(\psi,\gamma)$
(vi) $\Gamma(\varphi < \psi,\gamma) \in TRUE \Leftrightarrow \Gamma(\varphi,\gamma) <^{\mathcal{M}} \Gamma(\psi,\gamma)$
(vii) $\Gamma(\exists x.\varphi,\gamma) \in TRUE \Leftrightarrow$ there is some $m \in M$ such that $\Gamma(\varphi,\gamma_x^m) \in TRUE$

Definition 3.2
Let $\mathcal{M} = (M, TRUE, FALSE, <^{\mathcal{M}}, \Gamma)$ be a model, $\gamma \in M^V$ and $\varphi \in Expr(C)$. The satisfaction relation $\vDash$ is defined by:

$$(\mathcal{M},\gamma) \vDash \varphi :\Leftrightarrow \Gamma(\varphi,\gamma) \in TRUE.$$

The tupel $(\mathcal{M},\gamma)$ is called an interpretation. If $(\mathcal{M},\gamma) \vDash \varphi$, then we say that $(\mathcal{M},\gamma)$ is a model of φ. If φ is a sentence, then we may omit assignments writing $\mathcal{M} \vDash \varphi$. Analogously for sets Φ of expressions. The consequence relation $\Vdash$ is given in the usual model-theoretical way:

$$\Phi \Vdash \varphi :\Leftrightarrow \text{ every model of } \Phi \text{ is a model of } \varphi.$$

The Substitution Property (SP) implies the following Substitution Principle.

Lemma 3.3 (Substitution Principle)
For all formulas $\varphi,\psi,\chi \in Expr(C)$ and $x \in V$,

$$\Vdash \psi \equiv \chi \rightarrow \varphi[x:=\psi] \equiv \varphi[x:=\chi]$$

Proof. Let $\sigma_1 = [x:=\psi]$ and $\sigma_2 = [x:=\chi]$. Suppose $\Gamma(\psi,\gamma) = \Gamma(\chi,\gamma)$. Then $\gamma\sigma_1 = \gamma\sigma_2$. (SP) implies: $\Gamma(\varphi[\sigma_1],\gamma) = \Gamma(\varphi,\gamma\sigma_1) = \Gamma(\varphi,\gamma\sigma_2) = \Gamma(\varphi[\sigma_2],\gamma)$. ∎

Substitution Lemmata in various forms have been proved in [7, 10]. The proof of the following Lemma can be adopted from [4] (Lemma 3.14). It is a simplified version of a proof due to Zeitz [10].

Lemma 3.4 (Substitution Lemma)
Let $\mathcal{M} = (M, TRUE, FALSE, <^{\mathcal{M}}, \Gamma)$ be a model and $\varphi \in Expr(C)$.

(i) If σ,σ' are substitutions and $\gamma,\gamma' \in M^V$ are assignments with $\Gamma(\sigma(u),\gamma) = \Gamma(\sigma'(u),\gamma')$ for all $u \in fcon(\varphi)$, then $\Gamma(\varphi[\sigma],\gamma) = \Gamma(\varphi[\sigma'],\gamma')$.
(ii) If $\gamma \in M^V$ is an assignment and σ is a substitution such that $\Gamma(u) = \Gamma(\sigma(u),\gamma)$ for every $u \in con(\varphi)$, then $\Gamma(\varphi[\sigma],\gamma) = \Gamma(\varphi,\gamma\sigma)$.

Item (ii) of the Substitution Lemma says that the SP holds essentially for all substitutions $\sigma : V \cup C \rightarrow Expr(C)$ and not only for those of the form $\sigma : V \rightarrow Expr(C)$. A further consequence of SP is the Alpha Property which says that alpha-congruent formulas have the same denotations.

This property must be required as an additional semantic condition in [7] where substitutions are defined less restrictively. Sträter's original notion of substitutions also implies the Alpha Property. His notion, however, which is based on a ternary relation and a further function, is very complex.

Corollary 3.5 (Alpha Property (αP))
Let $\mathcal{M}$ be a model. For all expressions φ, ψ and all assignments $\gamma : V \to M$, if $\varphi =_\alpha \psi$, then $\Gamma(\varphi,\gamma) = \Gamma(\psi,\gamma)$.

Proof. Suppose $\varphi =_\alpha \psi$. By Lemma 2.22, this is equivalent with the condition $\varphi[\varepsilon] = \psi[\varepsilon]$. Note that $\gamma = \gamma\varepsilon$, for any assignment $\gamma : V \to M$. By item (ii) of the Substitution Lemma we get the following: $\Gamma(\varphi,\gamma) = \Gamma(\varphi,\gamma\varepsilon) = \Gamma(\varphi[\varepsilon],\gamma) = \Gamma(\psi[\varepsilon],\gamma) = \Gamma(\psi,\gamma\varepsilon) = \Gamma(\psi,\gamma)$. ∎

4 An extensional model

A model is extensional if any two formulas with the same truth value have the same denotation. In such two-element models, the sentences $\forall x.\forall y.((x \leftrightarrow y) \leftrightarrow (x \equiv y))$ and $\forall x.\forall y.((x \leftrightarrow y) \equiv (x \equiv y))$ are true. We call such models also Fregean.[17] We will see that also $\forall x.(x < x)$ is necessarily satisfied. Extensional $\in_T$-models (without reference connective) were already constructed in [2, 7, 10]. We adopt the construction given in [2].

Theorem 4.1 (Existence of models)
For any set C of constant symbols there exists an extensional model with respect to the language $Expr(C)$.

Proof. Let $TRUE = \{\top\}$, $FALSE = \{\bot\}$, $M = TRUE \cup FALSE$. Notice that in a two-element model the reference connective must be interpreted necessarily by the universal relation $M \times M$. This follows from the condition RP: $(x \vee \neg x) \prec ((x \vee \neg x) \vee y)$ shows $\top <^{\mathcal{M}} \top$, $(x \vee \neg x) \prec \neg(x \vee \neg x)$ shows $\top <^{\mathcal{M}} \bot$, etc. So we put $<^{\mathcal{M}} := M \times M$. We assume that there is a partition $C = C_t \cup C_f$ on the set of constant symbols. For any given $\gamma : V \to M$, the Γ-function is defined as follows: $\Gamma(x,\gamma) = \gamma(x)$ for $x \in V$, $\Gamma(c) = \top$ if $c \in C_t$, and $\Gamma(c) = \bot$ if $c \in C_f$. For any formulas φ, ψ, $\Gamma(\varphi \equiv \psi, \gamma) = \top$ if $\Gamma(\varphi,\gamma) = \Gamma(\psi,\gamma)$, and $\Gamma(\varphi \equiv \psi,\gamma) = \bot$ otherwise. For all φ, ψ, $\Gamma(\varphi < \psi, \gamma) = \top$. For the remaining cases, the definition proceeds in the obvious way inductively over the construction of formulas, in accordance with the respective truth conditions. It follows immediately from the definition that the Gamma-function satisfies the truth conditions. It is also clear that EP and RP are satisfied. CP and SP follow by induction on the expressions. Let us look at the quantifier case of SP. Suppose $\varphi = \exists x.\psi$, and let $\sigma : V \to Expr(C)$ be a substitution. First, one shows that for all $m \in M$ and all $y \in fvar(\psi)$ the following holds: $(\gamma\sigma)^m_x(y) = (\gamma^m_z \sigma[x := z])(y)$, where z is the variable forced by σ w.r.t. $\exists x.\psi$. That is, $z := lub(fvar((\exists x.\psi)[\sigma]))$ and $\varphi[\sigma] = (\exists x.\psi)[\sigma] = \exists z.\psi[\sigma[x := z]]$. Using this and the induction hypothesis, one shows the equivalence $\Gamma(\varphi,\gamma\sigma) = \top \Leftrightarrow \Gamma(\varphi[\sigma],\gamma) = \top$. Since the universe contains only two elements, it follows that $\Gamma(\varphi,\gamma\sigma) = \Gamma(\varphi[\sigma],\gamma)$. Thus, $\mathcal{M} = (M, TRUE, FALSE, <^{\mathcal{M}}, \Gamma)$ is an extensional model. ∎

5 A canonical model

The essential tool for our model construction is the notion of a canonical domain, which we introduce and study in the following. Such an algebraic structure is given by an infinite set M, a transitive

[17] If there were only Fregean models, then the logic would be Fregean, i.e. $(\varphi \leftrightarrow \psi) \to (\varphi \equiv \psi)$ would be valid.

well-founded relation $<^M$ and two unary predicates T, F on M. If C is a given set of constant symbols, then we may find a canonical domain $(M, T, F, <^M)$ such that $C \subseteq M$ is a set of $<^M$-minimal elements in M. Then we consider the whole set M as a set of constant symbols and construct a model $\mathcal{M} = (M, TRUE, FALSE, <^M, \Gamma)$ with respect to the language $Expr(M)$ such that the Gamma-function is a certain homomorphism between canonical domains. We will show that $\mathcal{M}$ is a canonical model with respect to the reduced language $Expr(C) \subseteq Expr(M)$ in the sense of the following definition.

Definition 5.1
Let C be a set of constant symbols. A model $\mathcal{M}$ is said to be canonical (with respect to the language $Expr(C)$) if there are no non-standard elements (i.e., every proposition $m \in M$ is denoted by some sentence $\chi \in Sent(C)$, $\Gamma(\chi) = m$), and for all sentences $\varphi, \psi \in Sent(C)$ the following holds:

$$\mathcal{M} \vDash \varphi < \psi \Leftrightarrow \varphi \prec \psi.$$

Lemma 5.2
If $\mathcal{M}$ is canonical, then for all sentences $\varphi, \psi \in Sent(C)$:

$$\mathcal{M} \vDash \varphi \equiv \psi \Leftrightarrow \varphi =_\alpha \psi.$$

Proof. Let $\mathcal{M}$ be canonical and let $\varphi \neq_\alpha \psi$, for two sentences φ, ψ. Suppose $\mathcal{M} \vDash \varphi \equiv \psi$. By (RP), $\mathcal{M} \vDash \varphi < (\varphi : true)$. The Substitution Principle yields $\mathcal{M} \vDash \psi < (\varphi : true)$. Since the model is canonical, $\psi \prec \varphi : true$. Then $\psi \prec \varphi$, since $\psi \neq_\alpha \varphi$. Thus, $\mathcal{M} \vDash \psi < \varphi$. Again by the Substitution Principle, $\mathcal{M} \vDash \varphi < \varphi$. It follows that $\varphi \prec \varphi$, a contradiction. Thus, $\mathcal{M} \nvDash \varphi \equiv \psi$. ∎

Note that the above proof does not need the condition that there are no non-standard elements.

Theorem 5.3
A model $\mathcal{M}$ is canonical w.r.t. $Expr(C)$ iff its Gamma-function induces an order-isomorphism $\Gamma : (Sent(C) /_{=_\alpha}, \prec) \to (M, <^M)$.

Proof. Suppose $\mathcal{M}$ is a canonical model. By the Alpha Property (Corollary 3.5), alpha-congruent sentences denote the same proposition. So we may factorize $Sent(C)$ modulo alpha-congruence. From Lemma 5.2 it follows that Γ is injective on $Sent(C) /_{=_\alpha}$. Since for every proposition there is a sentence that denote it, the Gamma-function is also surjective. It follows that Γ preserves the respective orderings. Thus, Γ induces the desired order-isomorphism. Now suppose that the Gamma-function of a model $\mathcal{M}$ induces an order-isomorphism such as given in the Theorem. Then it is easy to see that the conditions of Definition 5.1 are satisfied. Thus, $\mathcal{M}$ is canonical. ∎

Sträter calls a model $\mathcal{M}$ intensional if $\mathcal{M} \vDash \varphi \equiv \psi \Leftrightarrow \varphi =_\alpha \psi$, for all sentences φ, ψ. We saw that canonical models are in particular intensional in Sträter's sense. Is it possible to formulate conditions such that the converse is true, too? Recall that a model $\mathcal{M}$ is said to be $<$-intensional if for any formulas φ, ψ, and any assignment $\gamma : V \to M$, $(\mathcal{M}, \gamma) \vDash \varphi < \psi$ implies the existence of formulas φ', ψ' such that $(\mathcal{M}, \gamma) \vDash (\varphi \equiv \varphi') \wedge (\psi \equiv \psi')$ and $\varphi' \prec \psi'$ (see the introductory part of this article). Also recall that $\mathcal{M}$ is said to be a standard model if it is $<$-intensional and does not contain non-standard elements. Standard models are the intended ones. Let us say that a model is intensional if it satisfies Sträter's condition of intensionality above and additionally the condition of $<$-intensionality. Suppose now that $\mathcal{M}$ is an intensional standard model and $\mathcal{M} \vDash \varphi < \psi$, for two sentences φ and ψ. By definition, there are formulas φ' and ψ' such that $(\mathcal{M}, \gamma) \vDash (\varphi \equiv \varphi') \wedge (\psi \equiv \psi')$ and $\varphi' \prec \psi'$, where γ is any fixed assignment. For every $x \in fvar(\varphi') \cup fvar(\psi')$ let χ_x be a sentence such that $\Gamma(\chi_x) = \gamma(x)$.

Such sentences exist, since there are no non-standard elements. Let σ be the substitution defined by $x \mapsto \chi_x$, for $x \in fvar(\varphi') \cup fvar(\psi')$, and let $\varphi'' := \varphi'[\sigma]$, $\psi'' := \psi'[\sigma]$. By Lemma 2.30, $\varphi'' \prec \psi''$. By the Substitution Principle, $\Gamma(\varphi'') = \Gamma(\varphi', \gamma)$ and $\Gamma(\psi'') = \Gamma(\psi', \gamma)$. Since the model is intensional in Sträter's sense, it follows that $\varphi'' =_\alpha \varphi$ and $\psi'' =_\alpha \psi$. Thus, $\varphi \prec \psi$, and $\mathcal{M}$ is canonical. This, together with Lemma 5.2, yields the following characterization of canonical models.

Corollary 5.4
$\mathcal{M}$ is canonical iff $\mathcal{M}$ is an intensional standard model (i.e. $\mathcal{M}$ is intensional in Sträter's sense, $<$-intensional and without non-standard elements).

5.1 Canonical domains

In this section, we define canonical domains by a set of axioms which characterize them up to isomorphisms. Let us first study a useful tool, namely a rank that assigns to every element of a set M an ordinal (or ∞) in accordance with a given binary relation on M.[18]

Definition 5.5
Let M be a set and let E be a binary relation on M. We define a rank $R_E^M : M \to On \cup \{\infty\}$ in the following way (On denotes the class of all ordinals):

- $R_E^M(b) \geq 0$, if $b \in M$
- $R_E^M(b) \geq \alpha + 1$, if there is some $b' \in M$ such that $b'Eb$ and $R_E^M(b') \geq \alpha$
- $R_E^M(b) \geq \lambda$, if $R_E^M(b) \geq \beta$ for all $\beta < \lambda$, where λ is a limit ordinal
- $R_E^M(b) = \alpha$, if $R_E^M(b) \geq \alpha$ and $R_E^M(b) \ngeq \alpha + 1$
- $R_E^M(b) = \infty$, if $R_E^M(b) \geq \alpha$ for all $\alpha < On$

If the context is clear, then we may omit the superscript M and write $R_E(b)$ instead of $R_E^M(b)$. Let α be an ordinal. An E-chain of length α is a sequence $(b_i \mid i < \alpha)$ of (not necessarily pairwise distinct) elements of A such that $b_{\beta+1} E b_\beta$ whenever $\beta < \beta + 1 < \alpha$.

For example, an infinite E-chain starts as follows: $...b_3Eb_2Eb_1Eb_0$. In this sense, E-chains are always descending chains. The proof of the following facts relies on well-known set-theoretic standard arguments, so we will omit it.

Proposition 5.6
Let M be any set and $E \subseteq M \times M$.

(i) If $R_E^M(b) = \alpha$ and $b'Eb$, then $R_E^M(b') < \alpha$.
(ii) If $R_E^M(b) = \alpha + 1$, then there exists some $b'Eb$ such that $R_E^A(b') = \alpha$.
(iii) Suppose $|M| \geq \omega$. The following conditions are equivalent for $b \in M$:
  (a) $R_E^M(b) \geq |M|^+$.[19]
  (b) There exists some $a \in M$ such that aEb and $R_E^M(a) \geq |M|^+$.
  (c) There is an E-chain of length ω in M starting with b.
  (d) $R_E^M(b) = \infty$.
(iv) For $n < \omega$ and $b \in M$ the following two conditions are equivalent:
  (a) $R_E^M(b) \geq n$.
  (b) There is an E-chain of length $n+1$ starting with b.

[18] Our definition of rank is inspired by a well-known concept from classical model theory (namely rank of a type, where a type is, roughly speaking, a consistent set $\Phi(\bar{x})$ of formulas having all their free variables among the sequence $\bar{x}$).
[19] $|M|^+$ is the successor cardinal of $|M|$.

Note that $R_E^M(b) \geq \omega$ implies the existence of E-chains in M of arbitrary finite length. But in general it does not imply the existence of an infinite E-chain.

DEFINITION 5.7
Let M be a set and let E be a binary relation on M. We say that E is well-founded on M if $R_E^M(b)$ is ordinal valued for every $b \in M$.

As expected, and as we have seen in Proposition 5.6, E is well-founded on M if and only if there is no infinite E-chain in M.

The idea behind the following definition is that M represents the set of (normalized) sentences together with a truth predicate, and $<^M$ stands for the syntactic reference on the set of (normalized) sentences. Then the axioms of a canonical domain reflect the relevant properties of the reference relation.

DEFINITION 5.8
Let $\mathcal{M} = (M, T, F, <^M)$ be a structure with an infinite set M, unary predicates T and F on M, and a transitive relation $<^M \subseteq M \times M$. A subset $A \subseteq M$ is said to be closed if $b \in A$ and $a <^M b$ imply $a \in A$. For $b \in M$ we call the set $ext(b) := \{a \in M \mid a <^M b\}$ the extension of b. For a closed $A \subseteq M$ we define the set of successors of A by $succ(A) := \{b \in M \mid A = ext(b)\}$, and we put

$$succ(A)^T := \{b \in M \mid b \in succ(A) \text{ and } T(b)\},$$
$$succ(A)^F := \{b \in M \mid b \in succ(A) \text{ and } F(b)\}.$$

$\mathcal{M}$ is called a (classical) canonical domain if the following axioms are satisfied.

(i) Every element has finite extension.
(ii) For every non-empty finite and closed $A \subseteq M$, the sets $succ(A)^T$, $succ(A)^F$ have cardinality ω, and the sets $succ(\varnothing)^T$, $succ(\varnothing)^F$ have cardinality $|M|$.
(iii) For all $a \in M$, either $T(a)$ or $F(a)$.
(iv) $<^M$ is well-founded on M.

REMARK 5.9
Let $\mathcal{M} = (M, T, F, <^M)$ be a canonical domain. By axiom (iv), $a \not<^M a$ for all $a \in M$, hence $<^M$ is a partial order on M. $succ(\varnothing)$ is the infinite set of all $<^M$-minimal elements of M, i.e. the set of all elements of $R^M_{<^M}$-rank 0. If $A \neq B$ are closed subsets of M, then $succ(A) \cap succ(B) = \varnothing$. Every finite and closed set A is the extension of some element $a \in M$: $A = ext(a)$. On the other hand, for every $a \in M$ the set $ext(a)$ is finite and closed. That is, the finite and closed sets are precisely the sets $ext(a)$, $a \in M$. $ext(a) = \varnothing$ iff a has rank 0. If $ext(a) \neq \varnothing$, then this set has an element with maximal rank, say α. It follows that a has rank $\alpha + 1$. Finite closed sets are closed under the operations of union and intersection. Furthermore, $b \in succ(ext(a))$ iff $ext(a) = ext(b)$. Since $succ(ext(a))$ is an infinite set, $ext(a) = ext(b)$ does not imply $a = b$. In this sense, the principle of extensionality, which is fundamental in classical set theory, is here violated.

In the previous section, we proved the existence of (extensional) models. This is needed in the next result which states the existence of canonical domains. The proof of the theorem may also serve to improve the intuition behind the notion of canonical domain.

THEOREM 5.10 (Existence of canonical domains)
Let C be a set of constant symbols and let $\mathcal{M}$ be any model w.r.t. $Expr(C)$. If $|C| = \kappa \geq \omega$, then we assume that κ many constant symbols denote true propositions and κ many constant symbols denote

false propositions. We define two predicates T,F on the set of sentences $Sent(C)$ by

$$T(\varphi):\Leftrightarrow \mathcal{M}\vDash\varphi:true$$
$$F(\varphi):\Leftrightarrow \mathcal{M}\vDash\varphi:false$$

Then $\mathcal{A}=(Sent(C)/_\alpha,T,F,\prec)$ is a canonical domain, where $Sent(C)/_\alpha$ is the set of equivalence classes of sentences modulo alpha-congruence.[20]

Proof. In order to show the assertion, it is sufficient to work with normalized sentences instead of equivalence classes $\overline{\varphi}$ modulo alpha-congruence. So let us assume that all sentences considered in this proof are normalized. By Lemma 2.31, we know that $\prec$ is transitive. $\varphi\prec\psi$ implies that φ is alpha-congruent to a proper subformula of ψ. The number of subformulas is finite. Since alpha-congruent normalized sentences are equal, axiom (i) of a canonical domain follows. Let $A=\{\varphi_1,\dots,\varphi_n\}\subseteq Sent(C)$ be a finite and closed set of normalized sentences, i.e. $\varphi\in A$ and $\chi\prec\varphi$ imply $\chi\in A$. We suppose that $n=0$ iff $A=\varnothing$. A successor of this set is a sentence ψ such that $A=\{\varphi\mid\varphi\prec\psi\}$. If A is non-empty, then it is clear that at most ω many such sentences ψ can be constructed. If A is empty, then all constant symbols are among the successors of A. Thus, $|succ(A)|=|C|+\omega=|Sent(C)/_\alpha|$. For example, the following ω many normalized sentences are among the successors of A:

$$\exists x.(((x\equiv\varphi_1)\vee\dots\vee(x\equiv\varphi_n))\wedge(x\equiv x):true),$$
$$\exists x.(((x\equiv\varphi_1)\vee\dots\vee(x\equiv\varphi_n))\wedge(x\equiv x):true:true),$$
$$\dots$$

These sentences are valid and are therefore satisfied by $\mathcal{M}$. Moreover, for each sentence ψ of the list it holds that $\varphi\prec\psi$ iff $\varphi\in A$. Hence, each ψ of the list is a successor of A with truth value $T(\psi)$, i.e. $\psi\in succ(A)^T$. Similarly, one can find such a list of sentences which are elements of $succ(A)^F$. This shows that axiom (ii) of a canonical domain is satisfied. Since every sentence is either true or false w.r.t. $\mathcal{M}$, axiom (iii) holds. Of course, there is no infinite descending $\prec$-chain of sentences. Thus, $\prec$ is well-founded in the sense of Definition 5.7 and axiom (iv) is satisfied. ∎

Proposition 5.11
Let $\mathcal{M}=(M,T,F,<^M)$ be a canonical domain. Then every $a\in M$ has finite $R^M_{<^M}$-rank.

Proof. Towards a contradiction, we assume that $R^M_{<^M}(a)\geq\omega$ for some $a\in M$. Then by definition, $R^M_{<^M}(a)\geq\alpha+1$ for all $\alpha<\omega$. Hence, for all $\alpha<\omega$ there is some $a'\in M$ such that $a'<^M a$ and $R^M_{<^M}(a')\geq\alpha$. Since the extension of a is finite, there are only finitely many elements a' such that $a'<^M a$. This implies the existence of some $a'\in M$ with $a'<^M a$ and $R^M_{<^M}(a')\geq\omega$. Repeating the argument ω times we obtain an infinite chain $\dots a'''<^M a''<^M a'<^M a$, thus $<^M$ cannot be well-founded — a contradiction. Hence, $R^M_{<^M}(a)<\omega$. ∎

Theorem 5.12
Canonical domains of the same cardinality are isomorphic. That is, if $\mathcal{M}=(M,T,F,<^M)$ and $\mathcal{M}'=(M',T',F',<^{M'})$ are canonical domains of cardinality $\kappa\geq\omega$, then there exists a bijection $f:M\to M'$ such that $a<^M b\Leftrightarrow f(a)<^{M'}f(b)$, $T(a)\Leftrightarrow T'(f(a))$ and $F(a)\Leftrightarrow F'(f(a))$, for all $a,b\in M$.

[20]Recall that we have defined $\overline{\varphi}\prec\overline{\psi}:\Leftrightarrow\varphi\prec\psi$. Because of the Alpha-Property, we may also define $T(\overline{\varphi}):\Leftrightarrow T(\varphi)$ and $F(\overline{\varphi}):\Leftrightarrow F(\varphi)$.

PROOF. Let $\mathcal{M}=(M,T,F,<^M)$ and $\mathcal{M}'=(M',T',F',<^{M'})$ be canonical domains of cardinality $\kappa\geq\omega$. For $n<\omega$ let $M_n=\{a\in M \mid R^M_{<^M}(a)=n\}$ and $M'_n=\{a\in M' \mid R^{M'}_{<^{M'}}(a)=n\}$. By Proposition 5.11, each element has finite rank. Thus, $M=\bigcup_{n<\omega}M_n$ and $M'=\bigcup_{n<\omega}M'_n$. For all $n<\omega$, we define inductively isomorphisms $f_n:\bigcup_{i\leq n}M_i\to\bigcup_{i\leq n}M'_i$ between substructures of $\mathcal{M}$ and $\mathcal{M}'$ (where 'substructure' means that all relations are restricted to the respective subset) such that $f:=\bigcup_{i<\omega}f_i$ eventually will be an isomorphism from $\mathcal{M}$ onto $\mathcal{M}'$. By axiom (ii) of a canonical domain, there are κ many elements $a\in M_0$ with $T(a)$ and κ many elements $b\in M_0$ with $F(b)$. Similarly for M'_0. So we may find a bijection $f_0:M_0\to M'_0$ with the property: $T(a)$ iff $T'(f_0(a))$, and $F(a)$ iff $F'(f_0(a))$, for all $a\in M_0$. This is an isomorphism between the substructures with universes M_0 and M'_0, respectively. Now suppose that $f_n:\bigcup_{i\leq n}M_i\to\bigcup_{i\leq n}M'_i$ is already defined for some $n<\omega$. Put $\mathcal{B}_n:=\{B\subseteq\bigcup_{i\leq n}M_i \mid B$ is finite and closed and contains an element of $R^M_<$-rank $n\}$. Since f_n is an isomorphism, we get $f_n(\mathcal{B}_n)=\{B'\subseteq\bigcup_{i\leq n}M'_i \mid B'$ is finite and closed and contains an element of $R^{M'}_{<'}$-rank $n\}$. $B,B'\in\mathcal{B}_n$ and $B\neq B'$ implies $succ(B)^T\cap succ(B')^T=\varnothing$ and $succ(B)^F\cap succ(B')^F=\varnothing$. The elements of these successor sets are all of rank $n+1$. We show the following:

$$\bigcup_{B\in\mathcal{B}_n}(succ(B)^T\cup succ(B)^F)=M_{n+1}. \tag{5.1}$$

Suppose $a\in M_{n+1}$ and let $B=ext(a)$. Then $B\in\mathcal{B}_n$ and $a\in succ(B)^T\cup succ(B)^F$. Now it is clear that the Equation (5.1) holds. For every $B\in\mathcal{B}_n$, let $g_B:succ(B)^T\to succ(f_n(B))^{T'}$ and $h_B:succ(B)^F\to succ(f_n(B))^{F'}$ be bijections. We define:

$$f_{n+1}:=f_n\cup\bigcup_{B\in\mathcal{B}_n}(g_B\cup h_B).$$

From Equation (5.1) and from the fact that f_n is an isomorphism it follows that $\bigcup_{B\in\mathcal{B}_n}(g_B\cup h_B)$ is a bijection from M_{n+1} onto M'_{n+1}. Now one verifies that f_{n+1} is an isomorphism between substructures given by the universes $\bigcup_{i\leq n+1}M_i$ and $\bigcup_{i\leq n+1}M'_i$, respectively. Finally, we put $f:=\bigcup_{n<\omega}f_n$. Then f is an isomorphism from $\mathcal{M}$ onto $\mathcal{M}'$. ∎

5.2 The model construction

Let C be a set of constant symbols, and let $(M,T,F,<^M)$ be a canonical domain of cardinality $\kappa=|C|+\omega$. Then the set $M_0=succ(\varnothing)=succ(\varnothing)^T\cup succ(\varnothing)^F$ of elements with $R_{<^M}$-rank 0 has cardinality κ, by axiom (ii) of a canonical domain. One can find an embedding $h:C\to M_0$ such that the sets $succ(\varnothing)^T\smallsetminus h(C)$ and $succ(\varnothing)^F\smallsetminus h(C)$ have cardinality ω, respectively. We will identify C with the subset $h(C)\subseteq M_0\subseteq M$ and will consider the whole set $M\supseteq C$ as a set of constant symbols. In the following, we will work with the language $Expr(M)$. We put $K:=M\smallsetminus C$. From Equation (5.1) in the proof of Theorem 5.12 and axiom (ii) of a canonical domain it follows that each set M_n of elements of rank $n<\omega$ has cardinality κ. Thus, $|K|=\kappa$. All $c\in C$ have $R_{<^M}$-rank 0 in the canonical domain M. Besides the elements $c\in C$ there are further sentences $\varphi\in Sent(M)$ with $R_\prec$-rank 0, e.g. $\exists x.\exists y.\neg(x\equiv y)$. For these sentences, we have reserved the elements of the set $M_0\smallsetminus C$ as images. Recall that we have ensured that $|M_0\smallsetminus C|=\omega$. Sentences of $R_\prec$-rank $n>0$ then will be mapped to elements of M of $R_{<^M}$-rank n. As the universe of our model we choose the canonical domain M. Then the sets $TRUE:=\{m\in M \mid T(m)\}$ and $FALSE:=\{m\in M \mid F(m)\}$ as well as the reference relation

$<^M$ on M are already given. The rest of this section is dedicated to the construction of the Gamma-function. For this we will define predicates t and f on $Sent(M)$, and a function $red : Sent(M) \to M$ such that $red(m) = m$ and

- $red : (Sent(M), t, f, \prec) \to (M, T, F, <^M)$ is a strong homomorphism (i.e. red is an homomorphism and $red(\varphi) <^M red(\psi)$ implies the existence of sentences φ' and ψ' such that $\varphi' \prec \psi'$ and $red(\varphi') = red(\varphi)$ and $red(\psi') = red(\psi)$) which gives rise to
- an isomorphism from the factor structure $(Sent(M)/_{Ker}, t', f', \prec')$ to the canonical domain $(M, T, F, <^M)$ (where Ker is the kernel of red), and finally gives rise to
- an isomorphism between the canonical domains $(Sent(C)/_{=_\alpha}, t, f, \prec)$ and $(M, T, F, <^M)$, if we consider the reduced language $Expr(C)$.

The Gamma-function then is defined by means of the function red. If we reduce the language to $Expr(C) \subseteq Expr(M)$, then the Gamma-function coincides on $Sent(C)/_{=_\alpha}$ with the above-mentioned isomorphism between canonical domains. From this it will follow that the model is canonical with respect to the reduced language $Expr(C) \subseteq Expr(M)$.

The following definitions are inspired by similar concepts developed by Sträter in [7]. In particular, we adopt the function red from [7] and extend it to our richer language (which contains a reference connective).

Definition 5.13
An expression $\varphi \in Expr(M)$ is said to be reduced if $\psi \prec \varphi$ implies $\psi \in M$ or $fvar(\psi) \neq \varnothing$.

An expression φ is reduced if and only if all ψ with $\psi \prec \varphi$ are reduced.

Example 5.14
The sentence $\varphi = \exists x.((x < x) \wedge (\forall y.(y : true) \to (y \equiv x)))$ is reduced since there is no ψ with $\psi \prec \varphi$. The sentence $\chi = \exists x.((x < x) \vee (x < \exists y.y))$ is not reduced since $\exists y.y \prec \chi$. Variables and constant symbols are reduced.

Definition 5.15
We define the following sets. (Recall that $R_q(\varphi)$ is the quantifier-rank of φ, see Definition 2.4.)

- $A_0 := \{\neg m, m : true, m : false, m \equiv m', m \vee m', m < m' \mid m, m' \in M\}$
- $A_n := \{\exists x.\varphi \in Sent(M) \mid \exists x.\varphi$ is normalized, reduced and $R_q(\exists x.\varphi) = n\}$, for $n \geq 1$

Definition 5.16
Suppose that for each $n < \omega$, $g_n : A_n \to K$ is an injective function. The function $red : Expr(M) \to Expr(M)$ is defined as follows:

$$red(x) := x, \text{ for } x \in V$$

$$red(m) := m, \text{ for } m \in M = C \cup K$$

$$red(\neg\varphi) := \begin{cases} g_0(\neg red(\varphi)), & \text{if } red(\varphi) \in M \\ \neg(red(\varphi)), & \text{else} \end{cases}$$

$$red(\varphi : true) := \begin{cases} g_0(red(\varphi) : true), & \text{if } red(\varphi) \in M \\ red(\varphi) : true, & \text{else} \end{cases}$$

$$red(\varphi : false) := \begin{cases} g_0(red(\varphi) : false), & \text{if } red(\varphi) \in M \\ red(\varphi) : false, & \text{else} \end{cases}$$

$$red(\varphi \Box \psi) := \begin{cases} g_0(red(\varphi) \Box red(\psi)), & \text{if } red(\varphi) \in M \text{ and } red(\psi) \in M \\ red(\varphi) \Box red(\psi), & \text{else} \end{cases}$$

whenever $\Box \in \{\vee, \equiv, <\}$,

$$red(\exists x.\varphi) := \begin{cases} g_n(norm(\exists x.red(\varphi))), & \text{if } norm(\exists x.red(\varphi)) \in A_n \\ \exists x.red(\varphi), & \text{else} \end{cases}$$

red_n is the restriction of the function red on the set of expressions of quantifier-rank at most $n < \omega$ (we will need the fact that red_n already can be defined so far the functions $g_0, \ldots, g_n$ are defined):

$$red_n(\varphi) := \begin{cases} red(\varphi), \text{ if } R_q(\varphi) \leq n \\ \text{not defined, else} \end{cases}$$

The following results can be shown by induction on the construction of expressions. We will omit the proofs. Item (i) of the next result shows that for any expression $\exists x.\varphi$ we have $x \in fvar(red(\varphi))$. Thus, $\exists x.red(\varphi)$ is in fact an expression, and the last item of the definition of red is well defined.

Lemma 5.17
Let $\varphi \in Expr(M)$. Then:

(i) $fvar(\varphi) = fvar(red(\varphi))$.
(ii) $fvar(\varphi) = \varnothing \Longrightarrow red(\varphi) \in M$.
(iii) If $\psi \in subex(red(\varphi))$, then $\psi \in M$ or $fvar(\psi) \neq \varnothing$. Thus, $red(\varphi)$ is reduced.
(iv) If φ is normalized, then $red(\varphi)$ is normalized.
(v) If φ is reduced and $fvar(\varphi) \neq \varnothing$, then $red(\varphi) = \varphi$.

Lemma 5.18
Let $\varphi \in Expr(M)$ and let $\sigma : V \to V$ be a substitution. Then $red(\varphi[\sigma]) = red(\varphi)[\sigma]$.

Lemma 5.19
For all $\varphi \in Expr(M)$, $red(\varphi) = red(red(\varphi))$.

Lemma 5.20
Let $\varphi, \psi \in Expr(M)$. Then $\varphi =_\alpha \psi$ implies $red(\varphi) =_\alpha red(\psi)$.

Lemma 5.21
Let $\varphi \in Expr(M)$ and let σ, τ, ϱ be substitutions such that for all $u \in fcon(\varphi)$, $\varrho(u) = red((\sigma \circ \tau)(u))$. Then $red(\varphi[\sigma \circ \tau]) = red(\varphi[\varrho])$.

Corollary 5.22
Let $\varphi, \psi \in Expr(M)$ and $v \in V \cup M$. Then $red(\varphi[v := \psi]) = red(\varphi[v := red(\psi)])$.

PROOF. Choose $\sigma=\varepsilon$ and $\tau=[v:=\psi]$. Let ϱ be the substitution defined by $\varrho(u)=red(\tau(u))$, $u\in V\cup M$. Then the Corollary is an instance of Lemma 5.21. ∎

LEMMA 5.23
If $\varphi\in Expr(M)$ and $\sigma: V\to Expr(M)$, then $red(red(\varphi)[\sigma])=red(\varphi[\sigma])$.

We define two predicates t,f on $Sent(M)$. First, we put $t(m):\Leftrightarrow T(m)$ and $f(m):\Leftrightarrow F(m)$, for all $m\in M$. Now we define t and f on A_0 in accordance with the truth conditions of a model: $t(m:true):\Leftrightarrow t(m)$, $f(m:true):\Leftrightarrow f(m)$, $t(m:false):\Leftrightarrow f(m)$, $f(m:false):\Leftrightarrow t(m)$, $t(m\vee m'):\Leftrightarrow t(m)$ or $t(m')$, $f(m\vee m'):\Leftrightarrow f(m)$ and $f(m')$, $t(\neg m):\Leftrightarrow f(m)$, $f(\neg m):\Leftrightarrow t(m)$, $t(m<m'):\Leftrightarrow m<^M m'$, $f(m<m'):\Leftrightarrow$ not $m<^M m'$, $t(m\equiv m'):\Leftrightarrow m=m'$, $f(m\equiv m'):\Leftrightarrow m\neq m'$, for all $m,m'\in M$.

In the next step, we define the functions $g_n: A_n\to K$ using the algebraic properties of a canonical domain. For each finite and closed $B\subseteq M$ we choose partitions

$$(succ(B)^T\smallsetminus C)=\bigcup_{0<i<\omega} succ(B)^T_i\subseteq K,$$

$$(succ(B)^F\smallsetminus C)=\bigcup_{0<i<\omega} succ(B)^F_i\subseteq K$$

consisting of sets $succ(B)^T_i$ and $succ(B)^F_i$ of cardinality ω, respectively. Note that the set C in the above equations is relevant only in the case $B=\varnothing$, since $c\in C$ has rank 0 and cannot be the successor of a non-empty set B. Also recall that $succ(\varnothing)^T$ has cardinality κ, whereas $succ(\varnothing)^T\smallsetminus C$ has cardinality ω. Similarly for $succ(\varnothing)^F$. We suppose that the sets $succ(B)^T_i$ and $succ(B)^F_i$ are well ordered. The idea is the following. Every $\varphi\in A_0$ will be mapped by g_0 to an element of $succ(B)^T_1\cup succ(B)^F_1$, where B is the finite closed set generated by the constant symbols contained in φ.[21] Similarly, every $\varphi\in A_n$, where $n\geq 1$, will be mapped by g_n to an element of $succ(B)^T_n\cup succ(B)^F_n$, where B is generated by the constant symbols of φ. Recall that for $\varphi\in A_n$, $\psi\prec\varphi$ imlies $\psi\in M$. The construction will ensure that for a given finite closed set B, the ω many sentences of A_n whose constant symbols generate B are in one-to-one correspondence with ω many elements of $succ(B)^T_n$ or of $succ(B)^F_n$. The proof of Theorem 5.10 illustrates that infinitely many sentences with the same quantifier rank $n\geq 1$ can be constructed over the same finite set of constant symbols. On the other hand, if $\varphi\in A_0$, then there are only finitely many other sentences in A_0 with the same set of constant symbols as φ. Therefore, we map sentences of A_0 and A_1, which generate the same set B, both to elements of the infinite set $succ(B)^T_1\cup succ(B)^F_1$. But now let us define g_0.

Let $(\varphi^t_j)_{j<\kappa}$, $(\varphi^f_j)_{j<\kappa}$ be enumerations of the sets of true, false formulas of A_0, respectively. Let $j<\kappa$ and suppose that for all $l<j$, $g_0(\varphi^t_l)$ and $g_0(\varphi^f_l)$ are already defined. According to the definition of A_0, the formula $\varphi^t_j\in A_0$ contains only one or two constant symbols $m,m'\in M$ (possibly $m=m'$). Similarly, the formula $\varphi^f_j\in A_0$ contains only one or two constant symbols $p,p'\in M$ (possibly $p=p'$). Then for the finite and closed sets $B=ext(m)\cup ext(m')\cup\{m,m'\}$ and $B'=ext(p)\cup ext(p')\cup\{p,p'\}$ we define

$$g_0(\varphi^t_j):=min(succ(B)^T_1\smallsetminus\{g_0(\varphi^t_l)\mid l<j\}),$$

$$g_0(\varphi^f_j):=min(succ(B')^F_1\smallsetminus\{g_0(\varphi^f_l)\mid l<j\}).$$

where we consider the given well-orderings on $succ(B)^T_1$ and $succ(B')^F_1$. It is clear that g_0 is injective. Also note that g_0 maps only finitely many sentences to each $succ(B)^T_1$ because there are only finitely

[21] By 'B is generated by the constant symbols of φ' we mean that $b\in B$ iff there is a constant symbol m of φ such that $b=m$ or $b<^M m$. In other words, $B\subseteq M$ is the smallest down-set containing $con(\varphi)$. Note that it is enough to consider the $<^M$-maximal constant symbols of φ as the set of generating elements.

many sentences in A_0 with constant symbols that generate precisely the set B. Similarly for $succ(B)^F_1$. If $g_0,\ldots,g_n$ are already defined, and the predicates t,f are extended to A_n, then also red_n is defined and we are able to define g_{n+1}. First, we extend the definition of our predicates t,f to the set A_{n+1}:

$$t(\exists x.\varphi):\Leftrightarrow \text{ there is some } m\in M \text{ such that } red_n(\varphi[x:=m])\in TRUE$$
$$f(\exists x.\varphi):\Leftrightarrow \text{ there is no } m\in M \text{ such that } red_n(\varphi[x:=m])\in TRUE.$$

Notice that Lemma 5.17 guarantees that $red_n(\varphi[x:=m])$ in the above definition is in fact an element of M. Let $(\varphi^t_j)_{j<\kappa}$, $(\varphi^f_j)_{j<\kappa}$ be enumerations of the true, false sentences of A_{n+1}, respectively. Let $j<\kappa$ and suppose that for all $l<j$, $g_{n+1}(\varphi^t_l)$ and $g_{n+1}(\varphi^f_l)$ are already defined. Since the sentence $\varphi^t_j\in A_{n+1}$ is reduced, $\psi\prec\varphi^t_j$ implies $\psi\in M$. If $\{m_1,\ldots,m_r\}\subseteq M$ is the (possibly empty) set of all $\psi\prec\varphi^t_j$, then let $B:=ext(m_1)\cup\ldots\cup ext(m_r)\cup\{m_1,\ldots,m_r\}$. Similarly, we define $B'\subseteq M$ as the finite and closed set that corresponds to the sentence φ^f_j. Then:

$$g_{n+1}(\varphi^t_j):=min(succ(B)^T_{n+1}\smallsetminus(im(g_0)\cup\{g_{n+1}(\varphi^t_l)\mid l<j\})),$$
$$g_{n+1}(\varphi^f_j):=min(succ(B')^F_{n+1}\smallsetminus(im(g_0)\cup\{g_{n+1}(\varphi^f_l)\mid l<j\})),$$

where $im(g_0)$ denotes the image of function g_0. Since we have embedded the image of g_0 into the sets $succ(B)^T_1\cup succ(B)^F_1$, $im(g_0)$ is relevant only in the definitions of g_0 and g_1. Also recall that $im(g_0)\cap(succ(B)^T_1\cup succ(B)^F_1)$ is finite (see the definition above). g_{n+1} is obviously an injective function.

Lemma 5.24
$\bigcup_{n<\omega}g_n(A_n)=K$.

Proof. By definition, $\bigcup_{n<\omega}g_n(A_n)\subseteq K$. Let $k\in K$ and suppose $T(k)$. Let $\{m_1,\ldots,m_r\}$ be the set of maximal elements of $ext(k)$. Consider the closed set $B=ext(k)=ext(m_1)\cup\ldots\cup ext(m_r)\cup\{m_1,\ldots,m_r\}$. Then $k\in succ(B)^T$. By the above defined partition, $k\in succ(B)^T_n$, for some $n>0$. By the construction, $k\in im(g_n)\cup im(g_0)$. Note that $k\in im(g_0)$ is possible only if $r=1$ or $r=2$. If $F(k)$, then we argue similarly. ∎

The predicates t,f are now completely defined on $\bigcup_{n<\omega}A_n\cup M$. For every $n<\omega$ and every $\varphi\in A_n$, $t(\varphi)\Leftrightarrow g_n(\varphi)\in TRUE$ and $f(\varphi)\Leftrightarrow g_n(\varphi)\in FALSE$. Finally, we extend the predicates t,f to the whole set of sentences $Sent(M)$:

Definition 5.25
For $\varphi\in Sent(M)\smallsetminus(\bigcup_{n<\omega}A_n\cup M)$ we define

$$t(\varphi):\Leftrightarrow red(\varphi)\in TRUE$$
$$f(\varphi):\Leftrightarrow red(\varphi)\in FALSE.$$

Theorem 5.26
The function $red:(Sent(M),t,f,\prec)\to(M,T,F,<^M)$ is a strong homomorphism. More precisely, red is surjective and the following holds for all $\varphi,\psi\in Sent(M)$:

- $t(\varphi)\Leftrightarrow T(red(\varphi))$,
- $f(\varphi)\Leftrightarrow F(red(\varphi))$,
- $\varphi\prec\psi\Rightarrow red(\varphi)<^M red(\psi)$,

- If $red(\varphi) <^M red(\psi)$, then there are sentences φ' and ψ' such that $red(\varphi') = red(\varphi)$ and $red(\psi') = red(\psi)$ and $\varphi' \prec \psi'$.[22]

Proof. Since $red(m) = m$ for every $m \in M$, *red* is surjective. The first two items of the Theorem follow easily from the construction. Suppose $\varphi \prec \psi$. There is a chain $\varphi = \varphi_0 \prec \varphi_1 \prec \ldots \prec \varphi_k = \psi$ of maximal length. Then $R_\prec(\varphi_{i+1}) = R_\prec(\varphi_i) + 1$ for $i < k$. We show $red(\varphi_i) <^M red(\varphi_{i+1})$, $i < k$. The assertion then follows from the transitivity of $<^M$. Without lost of generality, we may assume that $\psi = \varphi_1$, i.e. $R_\prec(\psi) = R_\prec(\varphi) + 1$. Then ψ is alpha-congruent to $\varphi : true$ or to $\varphi : false$ or to $\neg\varphi$ or to a sentence of the form $\varphi \Box \chi$ or $\chi \Box \varphi$ or $\exists x.\xi$, where $\Box \in \{\vee, \equiv, <\}$ and $\varphi \prec \xi$. In all these cases, it follows from the definition of the functions *red* and g_n that *red* maps ψ to a successor of a finite closed set $B \subseteq M$ which is given by the sentence ψ (see the construction above). Moreover, one verifies that in all these cases $red(\varphi) = m \in B$. Thus, $red(\varphi) <^M red(\psi)$. Now suppose $red(\varphi) = m <^M m' = red(\psi)$. We must show the existence of two sentences φ', ψ' such that $\varphi' \prec \psi'$ and $red(\varphi') = m$ and $red(\psi') = m'$. Let $B = ext(m')$. Then $m \in B$ and $m' \in succ(B)$. In particular, $m' \in K$. Since $\bigcup_{n<\omega} g_n(A_n) = K$, there is $n < \omega$ and a sentence $\psi' \in A_n$ such that $b \in B \Leftrightarrow b \prec \psi'$, and $red(\psi') = m'$. Since $m \in B$, we get $m \prec \psi'$. Moreover, $red(m) = m$ and $red(\psi') = m'$. That is, $\varphi' = m$ and ψ' are the desired sentences. ∎

Definition 5.27
Let $Ker := \{(\varphi, \psi) \in Sent(M)^2 \mid red(\varphi) = red(\psi)\}$ and let $Sent(M)/_{Ker}$ be the corresponding partition. The equivalence class of $\varphi \in Sent(M)$ modulo *Ker* is denoted by $|\varphi|$. We define relations t', f' and $\prec'$ on $Sent(M)/_{Ker}$ by:

$$\begin{aligned}
&t'(|\varphi|) :\Leftrightarrow t(\varphi)\\
&f'(|\varphi|) :\Leftrightarrow f(\varphi)\\
&|\varphi| \prec' |\psi| :\Leftrightarrow \text{ there are } \varphi', \psi' \text{ such that } |\varphi'| = |\varphi|, |\psi'| = |\psi| \text{ and } \varphi' \prec \psi'
\end{aligned}$$

Corollary 5.28
$red : (Sent(M)/_{Ker}, t', f', \prec') \rightarrow (M, T, F, <^M)$, given by $|\varphi| \mapsto red(\varphi)$, is an isomorphism between canonical domains.

Proof. The relations $t', f', \prec'$ are well defined, this follows from Theorem 5.26. Then $(Sent(M)/_{Ker}, t', f', \prec')$ is the factor structure of $(Sent(M), t, f, \prec)$ modulo the congruence relation *Ker*. It turns out that the assertion of the Corollary is an instance of the Homomorphism Theorem of Universal Algebra. ∎

By Lemma 5.20, alpha-congruence between sentences refines the congruence relation *Ker*. Recall that we have defined the syntactical reference relation $\prec$ not only on the set of expressions, but also on the set of equivalence classes of expressions modulo alpha-congruence. Similarly, we may extend the predicates t and f on the set of equivalence classes of sentences modulo alpha-congruence: $t(\overline{\varphi}) :\Leftrightarrow t(\varphi)$, and $f(\overline{\varphi}) :\Leftrightarrow f(\varphi)$. From Lemma 5.20 and Theorem 5.26, it follows that these definitions are well defined. We now consider the reduced language $Expr(C) \subseteq Expr(M)$, and in particular $Sent(C) \subseteq Sent(M)$. By the previous remarks, we can work with the factor structure $(Sent(C)/_{=_\alpha}, t, f, \prec)$.

Theorem 5.29
Each class $|\varphi| \in Sent(M)/_{Ker}$ contains exactly one class $\overline{\psi} \in Sent(C)/_{=_\alpha}$ w.r.t. the reduced language. That is, for every $\varphi \in Sent(M)$ there is, up to alpha-congruence, exactly one $\psi \in Sent(C) \subseteq Sent(M)$ such that $red(\psi) = red(\varphi)$. Moreover, if $\varphi, \psi \in Sent(M)$ and $\varphi \prec \psi$, then there are $\varphi', \psi' \in Sent(C)$ such that $\varphi' \prec \psi'$, $red(\varphi') = red(\varphi)$ and $red(\psi') = red(\psi)$.

[22] Notice that we cannot expect that $red(\varphi) <^M red(\psi)$ implies $\varphi \prec \psi$. Consider, e.g. $\varphi = c \in C$ and $\psi = k = red(c : true) = g_0(c : true) \in succ(\{c\}) \in K$. Then $c <^M k$. However, $c \nprec k$. Nevertheless, $red(c) = c$ and $red(c : true) = red(k) = k$ and $c \prec c : true$.

PROOF. Let $\varphi \in Sent(M)$. The existence of the desired $\psi \in Sent(C)$ follows by induction on the rank $R_{\prec'}$ w.r.t. the canonical domain $(Sent(M)/_{Ker}, t', f', \prec')$. Now one shows that all such sentences are alpha-congruent. This follows from the following more general result which can be proved by induction on ψ: If $\psi, \psi' \in Expr(C)$, then $red(\psi) =_\alpha red(\psi')$ implies $\psi =_\alpha \psi'$. ∎

COROLLARY 5.30
The function $red : (Sent(C)/_{=_\alpha}, t, f, \prec) \to (M, T, F, <^M)$, given by $\overline{\varphi} \mapsto red(\varphi)$, is an isomorphism between canonical domains.

PROOF. This follows from Corollary 5.28 and Theorem 5.29. ∎

Now we reduce the language $Expr(M)$ to $Expr(C)$ and define our model with respect to the reduced language.

DEFINITION 5.31
For an assignment $\gamma : V \to M$ and $\varphi \in Expr(C)$ we define

$$\Gamma(\varphi, \gamma) := red(\varphi[\gamma]),$$

where the assignment γ is at the same time a substitution. Then our model with respect to the language $Expr(C)$ is $\mathcal{M} := (M, TRUE, FALSE, <^M, \Gamma)$.

It remains to show that $\mathcal{M}$ is actually a model, i.e. $\mathcal{M}$ must satisfy the structural properties EP, CP, SP, RP, and the truth conditions. Since $\Gamma(x, \gamma) = red(x[\gamma]) = red(\gamma(x)) = m$ if and only if $\gamma(x) = m$, EP holds. CP follows from Lemma 2.7(ii).

COROLLARY 5.32
$\mathcal{M}$ satisfies SP.

PROOF. Let $\varphi \in Expr(C)$, $\sigma : V \to Expr(C)$, $\gamma : V \to M$. We must show $\Gamma(\varphi[\sigma], \gamma) = \Gamma(\varphi, \gamma\sigma)$. Let $\gamma' : V \to M$ be the assignment $x \mapsto red((\sigma \circ \gamma)(x))$. By Lemma 5.21, $\Gamma(\varphi[\sigma], \gamma) = red(\varphi[\sigma][\gamma]) = red(\varphi[\sigma \circ \gamma]) = red(\varphi[\gamma']) = \Gamma(\varphi, \gamma')$, where $\gamma'(x) = red((\sigma \circ \gamma)(x)) = red(\sigma(x)[\gamma]) = \Gamma(\sigma(x), \gamma) = \gamma\sigma(x)$. Thus, $\Gamma(\varphi[\sigma], \gamma) = \Gamma(\varphi, \gamma') = \Gamma(\varphi, \gamma\sigma)$. ∎

LEMMA 5.33
$\mathcal{M}$ satisfies RP.

PROOF. This follows from Lemma 2.30 and Theorem 5.26. ∎

LEMMA 5.34
$\mathcal{M}$ satisfies the truth conditions.

PROOF. We show only the quantifier case.

$$\begin{aligned}
&\Gamma(\exists x.\varphi, \gamma) \in TRUE \\
&\Leftrightarrow red((\exists x.\varphi)[\gamma]) \in TRUE \\
&\overset{(i)}{\Leftrightarrow} red(\exists z.(\varphi[\gamma[x := z]])) \in TRUE \\
&\overset{(ii)}{\Leftrightarrow} g_n(\exists z.red(\varphi[\gamma[x := z]])) \in TRUE \\
&\Leftrightarrow red(red(\varphi[\gamma[x := z]])[z := m]) \in TRUE, \text{ for some } m \in M \\
&\overset{(iii)}{\Leftrightarrow} red(\varphi[\gamma[x := z]][z := m]) \in TRUE, \text{ for some } m \in M \\
&\Leftrightarrow red(\varphi[\gamma[x := m]]) \in TRUE, \text{ for some } m \in M \\
&\Leftrightarrow \Gamma(\varphi, \gamma_x^m) \in TRUE, \text{ for some } m \in M
\end{aligned}$$

(i): $z = lub(fvar((\exists x.\varphi)[\gamma])) = lub(\varnothing) = v_0 = min(V)$ is the variable forced by γ w.r.t. $\exists x.\varphi$.
(ii): $\varphi[\gamma[x:=z]]$ is normalized, by Lemma 2.19. Then $red(\varphi[\gamma[x:=z]])$ is normalized, by Lemma 5.17. Also by Lemma 5.17, $fvar(red(\varphi[\gamma[x:=z]])) = fvar(\varphi[\gamma[x:=z]])$. By Lemma 2.17 and Corollary 2.18, $\exists z.red(\varphi[\gamma[x:=z]])$ is normalized. Thus, $red(\exists z.\varphi[\gamma[x:=z]]) = g_n(\exists z.red(\varphi[\gamma[x:=z]]))$, by definition of *red*.
(iii): by Lemma 5.23. ∎

THEOREM 5.35
The model $\mathcal{M} = (M, TRUE, FALSE, <^M, \Gamma)$ is canonical with respect to the sublanguage $Expr(C) \subseteq Expr(M)$.

PROOF. This follows readily from Corollary 5.30 and Theorem 5.3. ∎

As already pointed out in the introductory part of the article, a canonical model gives rise to a unique 'term model' whose universe is given by the set of equivalence classes of sentences modulo alpha-congruence or, equivalently, by the set of all normalized sentences. In the following, we present the term model $\mathcal{M}^*$ which originates from the above constructed canonical model $\mathcal{M}$. Of course, $\mathcal{M}^*$ will be canonical, too. The universe M^* of the term model is the set of all normalized sentences, $TRUE^* = \{\varphi \in M^* \mid \Gamma(\varphi) \in TRUE\}$ and $FALSE^* = \{\varphi \in M^* \mid \Gamma(\varphi) \in FALSE\}$. $<^* := \prec \restriction (M^* \times M^*)$ is the semantic reference on M^*. Finally, the Gamma-function is defined by $\Gamma^*(\varphi, \gamma) := \varphi[\gamma]$, for any expression φ and any assignment $\gamma : V \to M^*$. Note that in this special case an assignment can be seen as a substitution. By Lemma 2.19, $\varphi[\gamma]$ is normalized. Hence, the Gamma-function is well defined. We must show that the structural and the truth conditions of a model are satisfied. It is clear that EP holds. CP follows from Lemma 2.7, RP follows from Lemma 2.30. We prove SP. Let $\sigma : V \to Expr(C)$ be a substitution and let $\gamma : V \to M^*$ be an assignment. Then for any $x \in V$, $\gamma\sigma(x) = \Gamma(\sigma(x), \gamma) = \sigma(x)[\gamma] = (\sigma \circ \gamma)(x)$. That is, $\gamma\sigma = \sigma \circ \gamma$. Using Lemma 2.7 and Lemma 2.11 it follows that $\Gamma^*(\varphi, \gamma\sigma) = \varphi[\gamma\sigma] = \varphi[\sigma \circ \gamma] = \varphi[\sigma][\gamma] = \Gamma^*(\varphi[\sigma], \gamma)$, for all expressions φ. Thus, SP is satisfied. It remains to show that the truth conditions are satisfied. Let us look at the quantifier case. We must show: $\Gamma^*(\exists x.\varphi, \gamma) \in TRUE^*$ iff there is a normalized sentence ψ such that $\Gamma^*(\varphi, \gamma_x^\psi) \in TRUE^*$. In the following equivalences, we suppose that $m \in M$ and ψ is the normalized sentence satisfying $\Gamma(\psi) = m$, and z is the forced variable w.r.t. the substitution (assignment) γ applied to $\exists x.\varphi$. If $\beta : V \to M$ is an assignment w.r.t. the model $\mathcal{M}$ and σ is the substitution $[z := \psi]$, then one checks that $\beta_z^m = \beta\sigma$. SP yields $\Gamma(\varphi[\gamma[x:=z]], \beta_z^m) = \Gamma(\varphi[\gamma[x:=\psi]], \beta)$. Consequently, if β is any assignment in model $\mathcal{M}$, then we get the following:

$$
\begin{aligned}
&\Gamma^*(\exists x.\varphi, \gamma) = \exists z.(\varphi[\gamma[x:=z]]) \in TRUE^* \\
&\Leftrightarrow \Gamma(\exists z.(\varphi[\gamma[x:=z]]), \beta) \in TRUE \\
&\Leftrightarrow \Gamma(\varphi[\gamma[x:=z]], \beta_z^m) \in TRUE \text{ for some } m \in M \\
&\Leftrightarrow \Gamma(\varphi[\gamma[x:=\psi]], \beta) \in TRUE \text{ for some normalized sentence } \psi \\
&\Leftrightarrow \Gamma^*(\varphi, \gamma_x^\psi) \in TRUE^* \text{ for some normalized sentence } \psi.
\end{aligned}
$$

We now consider the case of an equation. $\Gamma^*(\varphi \equiv \psi, \gamma) = (\varphi[\gamma] \equiv \psi[\gamma]) \in TRUE^* \Leftrightarrow \varphi[\gamma] = \psi[\gamma] \Leftrightarrow \Gamma^*(\varphi, \gamma) = \Gamma^*(\psi, \gamma)$ (recall that alpha-congruent normalized sentences are equal). The remaining cases follow similarly. Of course, the canonical term model $\mathcal{M}^* = (M^*, TRUE^*, FALSE^*, <^*, \Gamma^*)$ satisfies precisely the same set of sentences as $\mathcal{M}$. ∎

Zeitz [10] presents a completely different construction of a model $\mathcal{M}'$ (for the language without reference connective) such that $\mathcal{M}' \vDash \varphi \equiv \psi \Leftrightarrow \varphi =_\alpha \psi$, for all sentences φ, ψ. His construction is shorter and simpler than Sträter's original construction, but the resulting model has some counter-intuitive properties, in particular it contains non-standard elements. In the following, we sketch out the construction with respect to the extended language which contains the reference connective. In order to avoid the problem of the impredicativity of quantifiers, Zeitz follows the same strategy as in the construction above: first, fix the universe of the model and the truth values of its elements; only after that define the Gamma-funcion. We know that every sentence will denote either a true proposition or a false proposition. The idea of Zeitz is to assign to each sentence two objects which serve as possible denotations — one being true, the other one being false. This determines the subsets $TRUE'$ and $FALSE'$ of the propositional universe. The Gamma-function now can be defined inductively. Let us give the details. For every sentence φ we provide two representations: $norm(\varphi:true)$ and $norm(\varphi:false)$.[23] Then we put $TRUE' := \{norm(\varphi:true) \mid \varphi \in Sent(C)\}$, $FALSE' := \{norm(\varphi:false) \mid \varphi \in Sent(C)\}$, and $M' := TRUE' \cup FALSE'$. We define $||\varphi:true|| := ||\varphi:false|| := \varphi$, for $\varphi \in Sent(C)$. The semantical reference is given by $m <^{M'} m' :\Leftrightarrow ||m|| \prec ||m'||$. For every assignment $\gamma : V \to M'$ let $\tau_\gamma : V \to Sent(C)$ be defined by $\tau_\gamma(x) = ||\gamma(x)||$. Note that τ_γ is a substitution. By Lemma 2.19, $\varphi[\tau_\gamma]$ is normalized, i.e. $norm(\varphi[\tau_\gamma]) = \varphi[\tau_\gamma]$. We assume a partition $C = C_t \cup C_f$ on the set of constant symbols. The Gamma-function is defined over the inductive construction of formulas, simultaneously for all assignments $\gamma : V \to M'$: $\Gamma'(x,\gamma) := \gamma(x)$, for $x \in V$, and

$$\Gamma'(c,\gamma) := \begin{cases} c:true, \text{ if } c \in C_t \\ c:false, \text{ if } c \in C_f \end{cases}$$

$$\Gamma'(\varphi < \psi,\gamma) := \begin{cases} (\varphi < \psi)[\tau_\gamma]:true, \text{ if } \Gamma'(\varphi,\gamma) <^{M'} \Gamma'(\psi,\gamma) \\ (\varphi < \psi)[\tau_\gamma]:false, \text{ else} \end{cases}$$

$$\Gamma'(\exists x.\varphi,\gamma) := \begin{cases} (\exists x.\varphi)[\tau_\gamma]:true, \text{ if there is some } m \in M' \\ \quad \text{such that } \Gamma'(\varphi,\gamma_x^m) \in TRUE' \\ (\exists x.\varphi)[\tau_\gamma]:false, \text{ else} \end{cases}$$

and so on ... .

That is, if φ is a complex formula, then $\Gamma'(\varphi,\gamma)$ either equals $\varphi[\tau_\gamma]:true$ or it equals $\varphi[\tau_\gamma]:false$, according to the respective truth condition given by φ. It follows that $||\Gamma'(\varphi,\gamma)|| = \varphi[\tau_\gamma]$, for every $\varphi \in Expr(C)$. By Lemma 2.30, the structure $\mathcal{M}' = (M', TRUE', FALSE', \Gamma', <^{M'})$ satisfies the reference property RP. It is not hard to show that also EP, CP and SP do hold. Finally, the truth conditions follow directly from the definition of the Gamma-function. One easily checks that for all *sentences* φ, ψ, $\mathcal{M}' \vDash \varphi \equiv \psi \Leftrightarrow \varphi =_\alpha \psi$ and $\mathcal{M}' \vDash \varphi < \psi \Leftrightarrow \varphi \prec \psi$. Thus, the model would be canonical if there were no non-standard elements in the universe. Such non-standard elements can easily be identified. For instance, if $c \in C_t$, then $\Gamma'(c) = c:true$, and there is no sentence denoting the 'false proposition' $c:false \in FALSE'$. For $c \in C_t$ one verifies that the sentence $\psi := \exists x.((c < x) \wedge (x < (c:true:true)) \wedge \neg(x \equiv (c:true)))$ is true in model $\mathcal{M}'$ — consider an assignment that maps x to $c:true:false$. The truth of ψ depends on the existence of a non-standard element. ψ turns out to be false in a canonical model since in such a model there is only one proposition witnessing the truth of

[23] Zeitz works with equivalence classes of sentences modulo alpha-congruence instead of our function *norm*.

$\exists x.((c<x)\wedge(x<(c:true:true))$, namely the denotation of the sentence $c:true$. The example of the sentence ψ shows that we cannot simply remove the non-standard elements from the universe in order to get a canonical model: the Gamma-function w.r.t. the reduced universe would violate the truth condition concerning the existential quantifier.[24] It is not clear if there is any (more sophisticated) way to eliminate the non-standard elements and to transform $\mathcal{M}'$ into a canonical model.

The first construction method, which extends Sträter's approach and relies on our concept of a *canonical domain*, leads directly to a canonical model. The concept of a canonical domain also seems to be rather flexible with respect to further extensions of the language. The term model $\mathcal{M}^*$, originated from the canonical model $\mathcal{M}$, could be the starting point for the construction of further standard models that satisfy specific non-trivial equations. An example of such a standard model that contains precisely two self-referential propositions (a true truth-teller and a false truth-teller) was constructed for a quantifier-free language in [4]. Such a model satisfies precisely the equations of a set E^* which is generated by a set E of given equations and a certain congruence property (see the construction and the discussion on the properties of E^* in section 4.4 of [4]). In [5] (Section 5.2) a general and rather weak condition for a given set E of equations was found which implies the existence of a (infinite) standard model that satisfies precisely the equations of the generated set E^*. In many cases (see the examples given in [5]), the condition which E is required to satisfy can be reduced to the existence of a finite or even an extensional (two-element) model of E. Such models can easily be constructed. It would be interesting to further investigate such conditions and to extend the results to the first-order language of $\in_T$ with reference connective. Such future investigations could lead to the development of a general model theory that provides an overview of all standard models.

Acknowledgements

I would like to thank the anonymous referees for many helpful comments.

References

[1] S. L. Bloom and R. Suszko. Investigation into the sentential calculus with identity. *Notre Dame Journal of Formal Logic*, **13**, 289–308, 1972.

[2] S. Lewitzka. $\in_T(\Sigma)$-*Logik: Eine Erweiterung der Prädikatenlogik erster Stufe mit Selbstreferenz und totalem Wahrheitsprädikat*. Diplomarbeit, Technische Universität Berlin, 1998.

[3] S. Lewitzka. $\in_4$*: a* 4*-valued truth theory and meta-logic*. preprint 2007.

[4] S. Lewitzka. $\in_I$: an intuitioninistic logic without Fregean axiom and with predicates for truth and falsity. *Notre Dame Journal of Formal Logic*, **50**, 275–301, 2009.

[5] S. Lewitzka. $\in_K$: a non-Fregean logic of explicit knowledge. *Studia Logica*, **97**, 233–264, 2011.

[6] S. Lewitzka. Semantically closed intuitionistic abstract logics. *Journal of Logic and Computation*, 2011; doi: 10.1093/logcom/exq060

[7] W. Sträter. $\in_T$ Eine Logik erster Stufe mit Selbstreferenz und totalem Wahrheitsprädikat. Dissertation, KIT-Report 98. Technische Universität Berlin, 1992.

[8] R. Suszko. Non-Fregean Logic and Theories. *Analele Universitatii Bucuresti, Acta Logica*, **11**, 105–125, 1968.

[24]Note that ψ contains the $<$-connective. We were unable to find a similar example of a sentence ψ', but without reference connective, such that the satisfaction of ψ' in $\mathcal{M}'$ depends on a non-standard element as well.

[9] R. Suszko. Abolition of the Fregean Axiom. In *Logic Colloquium*. Vol. 453 of Lecture Notes in Mathematics, R. Parikh ed. Springer, 169–239, 1975.

[10] Ph. Zeitz. *Parametrisierte* $\in_T$*-Logik–eine Theorie der Erweiterung abstrakter Logiken um die Konzepte Wahrheit, Referenz und klassische Negation*. Dissertation. Logos Verlag Berlin, 2000.

Received 26 April 2011