Noname manuscript No.
(will be inserted by the editor)

A herbrandized functional interpretation of classical first-order logic

Fernando Ferreira · Gilda Ferreira

Received: date / Accepted: date

Abstract We introduce a new typed combinatory calculus with a type constructor that, to each type σ, associates the star type σ^* of the nonempty finite subsets of elements of type σ. We prove that this calculus enjoys the properties of strong normalization and confluence. With the aid of this star combinatory calculus, we define a functional interpretation of first-order predicate logic and prove a corresponding soundness theorem. It is seen that each theorem of classical first-order logic is connected with certain formulas which are *tautological in character*. As a corollary, we reprove Herbrand's theorem on the extraction of terms from classically provable existential statements.

Keywords Functional interpretations · first-order logic · star combinatory calculus · finite sets · tautologies · Herbrand's theorem

Mathematics Subject Classification (2000) 03F10 · 03B10 · 03B40 · 03B15

Both authors acknowledge Centro de Matemática, Aplicações Fundamentais e Investigação Operacional (Universidade de Lisboa) and the associated support of Fundação para a Ciência e a Tecnologia (FCT) [UID/MAT/04561/2013]. The second author is also grateful to FCT [UID/CEC/00408/2013 and grant SFRH/BPD/93278/2013] and to Large-Scale Informatics Systems Laboratory (Universidade de Lisboa).

F. Ferreira
Departamento de Matemática, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, Ed. C6, 1749-016, Lisboa, Portugal
E-mail: fjferreira@fc.ul.pt

G. Ferreira
Departamento de Matemática, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, Ed. C6, 1749-016, Lisboa, Portugal
E-mail: gmferreira@fc.ul.pt

1 Introduction

Benno van den Berg, Eyvind Briseid and Pavol Safarik introduced and studied in [3] a cluster of herbrandized functional interpretations of intuitionistic theories of nonstandard arithmetic. The functional interpretations are classified as *herbrandized* because the witnesses of the existential quantifiers of the interpreting formulas are *accumulated into* finite *sets*. Despite the nonstandard setting of [3], the interpretations make perfect sense for theories of standard arithmetic and even, as we shall see in this paper, for pure logic.

There are also other interpretations in the literature that accumulate existential witnesses. The earlier one seems to be the bounded functional interpretation of Paulo Oliva and the first author of this paper [8], where witnesses are accumulated below certain majorants. A realizability version was studied in [7]. Another one is the functional interpretation of Jeremy Avigad and Henry Towsner [2] of the arithmetical theory ID_1 of non-iterated inductive definitions. In this case, witnesses are accumulated into denumerable sets. A feature common to these so-called *cumulative interpretations* is a monotonicity property to the effect that the existential quantifications of the interpreting formulas are upward closed in the sense that if a certain existential functional is fit to work then a "bigger" functional is also fit (the notion of "bigger" varies from interpretation to interpretation).

In this paper, we consider *classical* first-order predicate logic without equality. This is also the calculus analyzed by Philipp Gerhardy and Ulrich Kohlenbach in [9]. Their paper presents an interpretation of classical logic based on the original (functional) *dialectica* interpretation of Kurt Gödel (cf. [10]). It actually follows a *direct* interpretation of classical first-order arithmetic (Peano arithmetic) due to Joseph Shoenfield in [12]. The Gerhardy-Kohlenbach interpretation is able to cope with the well-known contraction problem of functional interpretations (see [1] and [5] for a discussion of this issue) by means of conditional functionals. In 1974, Justus Diller and Werner Nahm had already found a way around the contraction problem. Their solution in [6] was proposed in the framework of intuitionistic arithmetic and, instead of conditional functionals, it uses finite sets in the interpretation of implication. Cumulative interpretations use the same blueprint in the interpretation of implication and, as a consequence, they are also able to cope with the contraction problem. The Diller-Nahm interpretation is not, however, a cumulative interpretation in our sense because it does not accumulate existential witnesses. In fact, accumulation is not needed for analyzing intuitionistic logic. However, if one goes to certain semi-intuitionistic systems (or to classical theories), we are led naturally to cumulative interpretations.

At first appearance there is an obvious difficulty with herbrandized functional interpretations in the context of logic (as opposed to arithmetic) because we have to speak of finite sets, and finiteness is not a logical notion. The way out of this situation is to separate the interpreted theory (classical first-order logic) from the interpreting language. The verification of the interpreting formulas can be done semantically, i.e., seen to be true in a certain structure

or, alternatively, verified in a certain theory (see the comments in Section 4 below). It can also be done, as we do here, in a purely logical way (modulo a certain combinatory calculus). In Section 2, we present and study the combinatory calculus that will be used in the paper. We prove that this calculus is strongly normalizable and satisfies the Church-Rosser property, and establish some consequences of these facts. These consequences are instrumental in defining the notion of a formula (of the interpreting language) being *tautological in character*. This is discussed in Section 3. The functional interpretation is defined in Section 4, and a soundness theorem is formulated and proved. Some examples are discussed. We finish the paper with a new proof of Herbrand's theorem concerning existential first-order sentences provable in classical first-order logic.

2 The star combinatory calculus

Let us fix a language $\mathcal{L}$ of pure first-order logic (equality is not present) with at least one constant symbol. In this section we study a combinatory calculus based on $\mathcal{L}$: the *star combinatory calculus* of $\mathcal{L}$.

Definition 1 The *types* of the star combinatory calculus of $\mathcal{L}$ are constructed from an atomic type (the ground type G) by means of two type-forming operations, $\to$ and *, in the following way:

(i) The ground type is a type.
(ii) If σ and τ are types then $\sigma \to \tau$ is a type.
(iii) If σ is a type, then σ^* is a type (a so-called *star type*).

The novelty are the star types. The intended meaning of σ^* is to give the type of all nonempty finite subsets of elements of type σ. It is clear that types have the form $\sigma_1 \to \ldots \to \sigma_n \to \rho$ ($n \geq 0$) with ρ either the ground type or a star type (i.e., ρ is G or ρ is σ^*, for a certain type σ).

The star calculus has three kinds of constants:

$\mathcal{L}$-constants. For each constant c of $\mathcal{L}$, we have a ground constant of the type calculus, also denoted by c. For each function symbol f of $\mathcal{L}$, we have a constant of the type calculus, also denoted by f, of type $G \to \ldots \to G \to G$ (the number of arrows is the arity of f).

Logical constants or combinators. A combinator $\Pi_{\sigma,\tau}$ of type $\sigma \to \tau \to \sigma$ for each pair of types σ, τ. A combinator $\Sigma_{\rho,\sigma,\tau}$ of type $(\rho \to \sigma \to \tau) \to (\rho \to \sigma) \to \rho \to \tau$, for each triple of types ρ, σ, τ.

Star constants. A constant $\mathfrak{s}_\sigma$ of type $\sigma \to \sigma^*$ for each type σ. A constant $\cup_\sigma$ of type $\sigma^* \to (\sigma^* \to \sigma^*)$ for each type σ. A constant $\bigcup_{\sigma,\tau}$ of type $\sigma^* \to (\sigma \to \tau^*) \to \tau^*$ for each pair of types σ, τ.

The combinators above are as usual, and the $\mathcal{L}$-constants are a natural thing to consider. The intended meaning of $\mathfrak{s}_\sigma$ is to map each element t^σ to the singleton set constituted by t. The intended meaning of $\cup_\sigma$ is to map

elements c and d of type σ^* to their union $c \cup d$. Finally, the intended meaning of $\bigcup_{\sigma,\tau}$ is to map $c : \sigma^*$ and $f : \sigma \to \tau^*$ to the indexed union $\bigcup_{w \in c} fw$.

Definition 2 The *terms* of the star combinatory calculus of $\mathcal{L}$ are generated by the following clauses:

(i) Constants are terms.
(ii) For each type σ there are denumerably many variables of type σ: $x^\sigma, y^\sigma, z^\sigma$, etc. Variables are terms.
(iii) If $t^{\sigma\to\tau}$ and q^σ are terms, then tq is a term of type τ.

The calculus has the following conversions:

$\Sigma tqr \rightsquigarrow tr(qr)$ and $\Pi tq \rightsquigarrow t$. These are as usual.

$\bigcup(\mathfrak{s}t)q \rightsquigarrow qt$. This conversion corresponds to the set-theoretical equality $\bigcup_{w\in\{t\}} qw = qt$.

$\bigcup(\cup tq)r \rightsquigarrow \cup(\bigcup tr)(\bigcup qr)$. This conversion corresponds to the set-theoretical equality $\bigcup_{w\in t\cup q} rw = (\bigcup_{w\in t} rw) \cup (\bigcup_{w\in q} rw)$.

We use the standard terminology of *redexes* and *contracta*, and have the usual definitions of *reduction in one step* and *reduction* (from one term t to another term q of the same type). We write $t \succ_1 q$ and $t \succeq q$ for these reductions, respectively. A term is *normal* if it has no redexes and so we can no longer apply a conversion. A term t is *strongly normalizable* if all the reduction sequences starting with t have finite length, i.e., t has a finite reduction tree. If a term t is strongly normalizable, we denote by $\nu(t)$ the maximum of the lengths of the reduction sequences starting with t.

The next order of business is to show that every term of the star calculus is strongly normalizable. We use the technique of *reducibility* introduced by William Tait in [14] (see, also, [16]). We need the following technical definitions:

Definition 3 Given a term t of type σ^*, we define a finite set of terms of type σ, the *surface members* of t, denoted by $SM(t)$. The definition is by induction on the complexity of t according to the following specification: unless t is of the form $\mathfrak{s}r$ or $\cup qr$, $SM(t)$ is the empty set; otherwise, $SM(\mathfrak{s}r)$ is $\{r\}$ and $SM(\cup qr)$ is $SM(q) \cup SM(r)$.

Notice that the surface members of a term are *not* invariant with respect to reductions. I.e., if $t \succeq q$ then $SM(t)$ and $SM(q)$ need not be the same set.

Definition 4 We say that a term t^σ is *reducible* if $t \in \mathsf{Red}_\sigma$, where Red_σ is defined by recursion on the complexity of the type σ as follows:

$t \in \mathsf{Red}_G :\equiv t$ is strongly normalizable.

$t \in \mathsf{Red}_{\sigma\to\tau} :\equiv$ for all q, if $q \in \mathsf{Red}_\sigma$ then $tq \in \mathsf{Red}_\tau$.

$t \in \mathsf{Red}_{\sigma^*} :\equiv t$ is strongly normalizable and, given any term in the reduction tree of t, its surface members are reducible.

The third clause in the definition above is new. Note that every term in the reduction tree of t has type σ^* and so its surface members have type σ. Therefore, the clause is well-defined.

Lemma 1 *Let x be a variable of type $\sigma_1 \to \ldots \to \sigma_n \to \rho$, with ρ the ground type or a star type. If $t_1, \ldots, t_n$ are strongly normalizable terms of types $\sigma_1, \ldots, \sigma_n$ (respectively), then $xt_1 \ldots t_n \in \mathsf{Red}_\rho$.*

Proof Since $t_1, \ldots, t_n$ are strongly normalizable, it is clear that $xt_1 \ldots t_n$ is strongly normalizable. Thus, when ρ is G we are done. For the case when ρ is σ^*, we need to check that the surface members of the terms in the reduction tree of $xt_1 \ldots t_n$ are reducible. Let q be a term in the reduction tree of $xt_1 \ldots t_n$. Then q has the form $xt'_1 \ldots t'_n$ with $t_i \succeq t'_i$ $(1 \leq i \leq n)$ and, therefore, $SM(q) = \emptyset$. Thus, $xt_1 \ldots t_n \in \mathsf{Red}_{\sigma^*}$. □

Lemma 2 *We have the following:*

(a) If $t^\sigma \in \mathsf{Red}_\sigma$ then t^σ is strongly normalizable.
(b) $x^\sigma \in \mathsf{Red}_\sigma$.

Proof We prove (a) and (b), simultaneously, by induction on the complexity of the type σ. The case when σ is G is immediate. Let us analyse the case when σ is of the form $\tau \to \theta$. Take $t^{\tau\to\theta} \in \mathsf{Red}_{\tau\to\theta}$. We want to prove that t is strongly normalizable. By induction hypothesis (b) for type τ, we know that $x^\tau \in \mathsf{Red}_\tau$. Thus $tx \in \mathsf{Red}_\theta$. By induction hypothesis (a) for θ, we have that tx is strongly normalizable. Therefore, t is strongly normalizable. For (b), we need to show that $x^{\tau\to\theta} \in \mathsf{Red}_{\tau\to\theta}$. Let $t \in \mathsf{Red}_\tau$. We need to prove that $xt \in \mathsf{Red}_\theta$. Suppose that θ is $\sigma_1 \to \ldots \to \sigma_n \to \rho$, with ρ either the ground type or a star type. Let $t_1, \ldots, t_n$ be reducible terms of types $\sigma_1, \ldots, \sigma_n$ respectively. By induction hypothesis (a) for types $\tau, \sigma_1, \ldots, \sigma_n$, we know that $t, t_1, \ldots, t_n$ are strongly normalizable. Thus, by Lemma 1, we have that $xtt_1 \ldots t_n \in \mathsf{Red}_\rho$. Therefore $xt \in \mathsf{Red}_\theta$.

Finally, let us turn to the case when σ is of the form ρ^*. Property (a) is immediate. Regarding (b), note that the reduction tree of x^{ρ^*} has a single term, the variable itself, and that $SM(x^{\rho^*}) = \emptyset$. Thus $x^{\rho^*} \in \mathsf{Red}_{\rho^*}$. □

Lemma 3 *If $t^\sigma \in \mathsf{Red}_\sigma$ and $t^\sigma \succeq q^\sigma$ then $q^\sigma \in \mathsf{Red}_\sigma$.*

Proof The proof is done by induction on the complexity of the type σ. The base type G is clear. The arrow case is immediate by definition of reducibility. Let us study the case when σ is τ^*. So, assume that $t \in \mathsf{Red}_{\tau^*}$ and $t \succeq q$. We want to prove that $q \in \mathsf{Red}_{\tau^*}$. By definition, t is strongly normalizable. Therefore, q is strong normalizable as well. Note also that any term in the reduction tree of q is a term in the reduction tree of t. Thus (because $t \in \mathsf{Red}_{\tau^*}$) the surface members of any term in the reduction tree of q are reducible. Hence, $q \in \mathsf{Red}_{\tau^*}$. □

Lemma 4 *We have the following:*

(a) *Let t be a term of ground type. If all terms that result from t via a one step reduction are reducible then t is also reducible.*
(b) *Let t be a term of star type. If all the surface members of t are reducible and all terms that result from t via a one step reduction are reducible, then t is also reducible.*

Proof Part (a) is obvious. To prove part (b), take a term t of type τ^* in the conditions of the hypothesis. Clearly, t is strongly normalizable. A term in the reduction tree of t is either the term t itself or it is a term in the reduction tree of t', with t' such that $t \succ_1 t'$. Since, by hypothesis, $t' \in \mathsf{Red}_{\tau^*}$ we know that given any term in the reduction tree of t' its surface members are reducible. Also, by hypothesis, the surface members of t are reducible. Therefore $t \in \mathsf{Red}_{\tau^*}$. □

Proposition 1 *All the terms of the star combinatory calculus are reducible.*

Proof The proof is by induction on the complexity of the term. By Definition 4 we know that if $t \in \mathsf{Red}_{\sigma\to\tau}$ and $q \in \mathsf{Red}_\sigma$ then $tq \in \mathsf{Red}_\tau$. Thus it suffices to prove that the constants of the calculus are reducible. (Note that we already know, by Lemma 2(b), that the variables are reducible.)

It is clear that $c^G \in \mathsf{Red}_G$, where c^G is a constant of the calculus that results from a constant c in the language $\mathcal{L}$. Let f be a constant of type $G \to \ldots \to G \to G$ that results from a n-ary function symbol f in the language $\mathcal{L}$. In order to prove that f is reducible, we need to show that if $s_1, \ldots, s_n$ are reducible terms of type G (i.e., strongly normalizable) then $fs_1 \ldots s_n$ is also reducible of type Red_G (i.e., strongly normalizable). This is obvious.

After having discussed the $\mathcal{L}$-constants, we study the logical constants (combinators). Let us argue that $\Pi_{\sigma,\tau} \in \mathsf{Red}_{\sigma\to\tau\to\sigma}$. Take $t \in \mathsf{Red}_\sigma$ and $q \in \mathsf{Red}_\tau$ in order to show that $\Pi_{\sigma,\tau}tq \in \mathsf{Red}_\sigma$. Suppose that σ has the form $\sigma_1 \to \ldots \to \sigma_n \to \rho$ ($n \geq 0$ and ρ is either G or a star type). Let $s_1, \ldots, s_n$ be reducible terms of types $\sigma_1, \ldots, \sigma_n$. We need to prove that $\Pi tqs_1 \ldots s_n \in \mathsf{Red}_\rho$. The reasoning is as in p. 107 of [16], and we do not repeat it here. We just notice that it relies on Lemma 4 and that the new case when ρ is of star type does not present any problem. The case of the combinators $\Sigma_{\rho,\sigma,\tau}$ is similar.

The really new cases to consider are the star constants. Let us start by showing that $\mathfrak{s}_\sigma \in \mathsf{Red}_{\sigma\to\sigma^*}$. Take $t \in \mathsf{Red}_\sigma$ in order to show that $\mathfrak{s}t \in \mathsf{Red}_{\sigma^*}$. Since $t \in \mathsf{Red}_\sigma$, by Lemma 2$(a)$ we know that t is strongly normalizable. Therefore, $\mathfrak{s}t$ is strongly normalizable because any reduction in $\mathfrak{s}t$ results from a reduction in t. In fact, any term in the reduction tree of $\mathfrak{s}t$ has the form $\mathfrak{s}t'$ with $t \succeq t'$. Well, $SM(\mathfrak{s}t') = \{t'\}$. So, the surface members of every term in the reduction tree of $\mathfrak{s}t$ are reducible (use Lemma 3). We have proved that $\mathfrak{s}t \in \mathsf{Red}_{\sigma^*}$.

Let us prove that $\cup_\sigma \in \mathsf{Red}_{\sigma^*\to(\sigma^*\to\sigma^*)}$. Take $t, q \in \mathsf{Red}_{\sigma^*}$ in order to show that $\cup tq \in \mathsf{Red}_{\sigma^*}$. The proof is done by induction on $\nu(t)+\nu(q)$ (this is possible because t and q are strongly normalizable). The one step reductions from $\cup tq$ are of the form $\cup t'q$ or $\cup tq'$ with $t \succ_1 t'$ and $q \succ_1 q'$. Since $\nu(t') + \nu(q) < \nu(t) + \nu(q)$, $\nu(t) + \nu(q') < \nu(t) + \nu(q)$ and t' and q' are reducible (by Lemma

3), by induction hypothesis we have that $\cup t'q$ and $\cup tq'$ are reducible. On the other hand, $SM(\cup tq) = SM(t) \cup SM(q)$. Well, the elements of $SM(t)$ and of $SM(q)$ are reducible because $t, q \in \mathsf{Red}_{\sigma^*}$. Hence, the surface members of $\cup tq$ are reducible. By Lemma 4(b), we conclude that $\cup tq \in \mathsf{Red}_{\sigma^*}$.

Let us finally prove that $\bigcup_{\sigma,\tau} \in \mathsf{Red}_{\sigma^* \to (\sigma \to \tau^*) \to \tau^*}$. We need to show that $\bigcup tq \in \mathsf{Red}_{\tau^*}$, for $t \in \mathsf{Red}_{\sigma^*}$ and $q \in \mathsf{Red}_{\sigma \to \tau^*}$. Since $SM(\bigcup tq) = \emptyset$, by Lemma 4($b$) it is enough to show that every one step reduction of $\bigcup tq$ leads to a reducible term. The proof is done by main induction on $\nu(t) + \nu(q)$ and sub-induction on $c(t)$, where $c(t)$ measures the complexity of the term t in the following way: $c(\text{constant}) = c(\text{variable}) = 1$; $c(tq) = c(t) + c(q)$. The one step reductions from $\bigcup tq$ are of the form:

(i) $\bigcup t'q$ with $t \succ_1 t'$
(ii) $\bigcup tq'$ with $q \succ_1 q'$
(iii) qr, if t is $\mathfrak{s}r$
(iv) $\cup(\bigcup rq)(\bigcup sq)$, if t is $\cup rs$.

In cases (i) and (ii), $\nu(t') + \nu(q)$ and $\nu(t) + \nu(q')$ are stricty less than $\nu(t) + \nu(q)$. We can apply the induction hypothesis (using Lemma 3) and conclude that $\bigcup t'q$ and $\bigcup tq'$ are reducible. Let us analyse case (iii). By hypothesis, $\mathfrak{s}r \in \mathsf{Red}_{\sigma^*}$ (note that t is $\mathfrak{s}r$). By Definition 4, $r \in \mathsf{Red}_{\sigma}$. Since we also have $q \in \mathsf{Red}_{\sigma \to \tau^*}$, we get $qr \in \mathsf{Red}_{\tau^*}$, as wished. In case (iv), by hypothesis, $\cup rs \in \mathsf{Red}_{\sigma^*}$ (note that t is $\cup rs$). We first argue that $r \in \mathsf{Red}_{\sigma^*}$ and $s \in \mathsf{Red}_{\sigma^*}$. It is clear that both r and s are strongly normalizable (an infinite reduction sequence starting in one of them would entail the existence of an infinite reduction sequence in $\cup rs$). Let r' be a term in the reduction tree of r. Then $\cup r's$ is a term in the reduction tree of t. Since $SM(\cup r's) = SM(r') \cup SM(s)$ and $t \in \mathsf{Red}_{\sigma^*}$, we can conclude that the surface members of r' are reducible. Thus, given a term in the reduction tree of r (similarly for s), its surface members are reducible. Therefore $r, s \in \mathsf{Red}_{\sigma^*}$. Since $\nu(r) \leq \nu(t)$ and $c(r) < c(t)$, we have by induction hypothesis that $\bigcup rq \in \mathsf{Red}_{\tau^*}$. Similarly $\bigcup sq \in \mathsf{Red}_{\tau^*}$. By definition of reducibility, both $\bigcup rq$ and $\bigcup sq$ are strongly normalizable, and this easily entails that $\cup(\bigcup rq)(\bigcup sq)$ is strongly normalizable. It remains to show that the surface members of the terms in the reduction tree of $\cup(\bigcup rq)(\bigcup sq)$ are reducible. The root of the reduction tree is $\cup(\bigcup rq)(\bigcup sq)$. Well, $SM(\cup(\bigcup rq)(\bigcup sq)) = SM(\bigcup rq) \cup SM(\bigcup rs) = \emptyset \cup \emptyset = \emptyset$. Any other term of the reduction tree of $\cup(\bigcup rq)(\bigcup sq)$ is of the form $\cup uv$, where $\bigcup rq \succeq u$ and $\bigcup sq \succeq v$. Since $SM(\cup uv)$ is $SM(u) \cup SM(v)$, the elements of $SM(\cup uv)$ are reducible (because $\bigcup rq, \bigcup sq \in \mathsf{Red}_{\tau^*}$). Therefore, $\cup(\bigcup rq)(\bigcup sq) \in \mathsf{Red}_{\tau^*}$. □

By the above proposition and Lemma 2(a), we have the following important result:

Theorem 1 (Strong normalization) *The star combinatory calculus enjoys the property of strong normalization.*

We also have the Church-Rosser property:

Theorem 2 (Church-Rosser property) *The star combinatory calculus is confluent, i.e. it enjoys the Church-Rosser property.*

Proof We have already shown that the star calculus enjoys the property of strong normalization. So, by Newman's Lemma (see [15] for a proof and the reference to the original work of Maxwell Newman), it is enough to prove that the calculus is weakly confluent, i.e., if $t \succ_1 t'$ and $t \succ_1 t''$ then there is a term t''' such that $t' \succeq t'''$ and $t'' \succeq t'''$.

Of course, if the conversions leading from t to t' and from t to t'' concern disjoint redexes, then t''' is simply obtained by converting both redexes. Let us analyse the situation where the redexes are nested. There are various cases to consider, depending on the conversions. The cases regarding the combinators are dealt as usual (cf. [16]). So, let us focus on the cases that concern the new conversions.

If $t \equiv \ldots \bigcup(\mathfrak{s}r)q \ldots$, $t' \equiv \ldots qr \ldots$ and $t'' \equiv \ldots \bigcup(\mathfrak{s}r')q \ldots$ with $r \succ_1 r'$ then $t''' \equiv \ldots qr' \ldots$ and $t' \succeq t'''$ in a single step via $r \succ_1 r'$ and $t'' \succeq t'''$ in a single step via $\bigcup(\mathfrak{s}r')q \succ_1 qr'$.

If $t \equiv \ldots \bigcup(\mathfrak{s}r)q \ldots$, $t' \equiv \ldots qr \ldots$ and $t'' \equiv \ldots \bigcup(\mathfrak{s}r)q' \ldots$ with $q \succ_1 q'$ then $t''' \equiv \ldots q'r \ldots$ and $t' \succeq t'''$ in a single step via $q \succ_1 q'$ and $t'' \succeq t'''$ in a single step via $\bigcup(\mathfrak{s}r)q' \succ_1 q'r$.

If $t \equiv \ldots \bigcup(\cup sq)r \ldots$, $t' \equiv \ldots \cup (\bigcup sr)(\bigcup qr) \ldots$ and $t'' \equiv \ldots \bigcup(\cup s'q)r \ldots$ with $s \succ_1 s'$ then $t''' \equiv \ldots \cup (\bigcup s'r)(\bigcup qr) \ldots$ and $t' \succeq t'''$ in a single step via $s \succ_1 s'$ and $t'' \succeq t'''$ in a single step via $\bigcup(\cup s'q)r \succ_1 \cup(\bigcup s'r)(\bigcup qr)$.

If $t \equiv \ldots \bigcup(\cup sq)r \ldots$, $t' \equiv \ldots \cup (\bigcup sr)(\bigcup qr) \ldots$ and $t'' \equiv \ldots \bigcup(\cup sq')r \ldots$ with $q \succ_1 q'$ then $t''' \equiv \ldots \cup (\bigcup sr)(\bigcup q'r) \ldots$ and $t' \succeq t'''$ in a single step via $q \succ_1 q'$ and $t'' \succeq t'''$ in a single step via $\bigcup(\cup sq')r \succ_1 \cup(\bigcup sr)(\bigcup q'r)$.

If $t \equiv \ldots \bigcup(\cup sq)r \ldots$, $t' \equiv \ldots \cup (\bigcup sr)(\bigcup qr) \ldots$ and $t'' \equiv \ldots \bigcup(\cup sq)r' \ldots$ with $r \succ_1 r'$ then $t''' \equiv \ldots \cup (\bigcup sr')(\bigcup qr') \ldots$ and $t' \succeq t'''$ with two applications of $r \succ_1 r'$ and $t'' \succeq t'''$ in a single step via $\bigcup(\cup sq)r' \succ_1 \cup(\bigcup sr')(\bigcup qr')$. □

It is clear that first-order terms of the language $\mathcal{L}$ have natural analogues in the star combinatory calculus. We identify these pairs of terms.

Proposition 2 (Ground normal form) *If t is a closed normal term of ground type, then t is a (closed) first-order term of the language $\mathcal{L}$.*

Proof It is clear that closed terms are of the form $at_1 \ldots t_m$ with $m \geq 0$, where a is a constant and $t_1, \ldots, t_m$ are closed terms. Therefore, a *normal* closed term must be of one of the forms: c, f, fr_1, fr_1r_2, $\ldots$, $fr_1 \ldots r_n$, Σ, Σt, Σtv, Π, Πq, $\mathfrak{s}$, $\mathfrak{s}q$, $\cup$, $\cup u_1$, $\cup u_1u_2$, $\bigcup$, $\bigcup s$ or $\bigcup sp$, for c a constant in the language $\mathcal{L}$, f a n-ary function symbol in the language $\mathcal{L}$ and $r_1, r_2, \ldots, r_n, t, v, q, u_1, u_2, s, p$ closed normal terms of appropriate types. If, further, the closed normal term is of ground type then it must be c or of the form $fr_1 \ldots r_n$, as above. By an inductive argument, we may suppose that $r_1, \ldots, r_n$ are first-order terms. We are done. □

We finish this section with a simple, but important, proposition.

Definition 5 A term t of star type is called *set-like* if it is built from terms of the form $\mathfrak{s}q^{\sigma}$ and the binary union operator $\cup_{\sigma}$.

Proposition 3 (Star normal form) *If t is a closed normal term of star type σ^*, then t is set-like and $SM(t)$ is a* nonempty *finite set of closed normal terms of type σ.*

Proof Looking at the proof of Proposition 2 and taking into account that this time the closed normal term t is of star type σ^* (and not of ground type), then t must fall into one of the following cases: (i) t is $\mathfrak{s}^{\sigma\to\sigma^*}q^{\sigma}$, (ii) t is $\cup^{\sigma^*\to(\sigma^*\to\sigma^*)}u_1^{\sigma^*}u_2^{\sigma^*}$ or (iii) t is $\bigcup^{\theta^*\to(\theta\to\sigma^*)\to\sigma^*}s^{\theta^*}p^{\theta\to\sigma^*}$.

Let us prove the result by induction on the complexity of t. Case (i) is immediate, because it is obviously set-like with $SM(t)=\{q\}$. For case (ii) we have, by induction hypothesis, that u_1 and u_2 are set-like. So, t is also set-like.

We finish the proof by arguing that case (iii) cannot hold. Take $\bigcup sp$ a closed normal term of type σ^* as above. Since s is a closed normal term of star type, by induction hypothesis it is set-like. Hence s is either of the form $\mathfrak{s}r$ or of the form $\cup s_1 s_2$. In either case, $\bigcup sp$ would not be normal. □

3 Bounded mixed formulas and the propositional calculus

In this section we introduce a language $\mathcal{L}^{\mathrm{mix}}$ which is a combination of the propositional part of the first-order language $\mathcal{L}$ and the terms of the combinatory star calculus based on $\mathcal{L}$. Moreover, this mixed language has also a new sort of quantification: a particular kind of *bounded* quantification. The atomic formulas of the *bounded mixed language* $\mathcal{L}^{\mathrm{mix}}$ are formulas of the form $P(t_1,\ldots,t_n)$, where P is a n-ary predicate symbol of the language $\mathcal{L}$ and $t_1,\ldots,t_n$ are ground terms of the combinatory star calculus. Note that these ground terms include the first-order terms of $\mathcal{L}$ but go beyond them (for instance, it also includes terms like $\Pi x^G y^{\rho}$). Therefore, the atomic formulas of the bounded mixed language include, but are not restricted to, the atomic formulas of $\mathcal{L}$. The formulas of $\mathcal{L}^{\mathrm{mix}}$ (the bounded mixed formulas) are obtained from the atomic formulas by means of propositional connectives (we use negation, conjunction and disjunction as primitives; sometimes we also use implication, seen as the usual abbreviation) and *bounded quantifiers* of the form $\forall x^{\sigma}\in t^{\sigma^*}(\ldots)$ and $\exists x^{\sigma}\in t^{\sigma^*}(\ldots)$, where t is a term (of star type σ^*) in which the variable x (of type σ) does not occur. There is certainly an underlying intended meaning in our use of the membership sign in the quantifications but in this work the bounded set quantifiers should be seen as a mere syntactic matter. We will use below the membership sign both as a syntactic sign of the mixed language $\mathcal{L}^{\mathrm{mix}}$ and naively as it is usual in mathematics, but the intended use will always be clear from the context.

As it is well-known, there is a propositional calculus associated with the first-order language $\mathcal{L}$, where the atomic propositional formulas are the closed atomic formulas of $\mathcal{L}$. In the next definition we are going to associate to each *sentence* of $\mathcal{L}^{\mathrm{mix}}$ a formula of the propositional calculus of $\mathcal{L}$. In order to do

this, we need the results of the previous section. In particular, if t is a term of the star combinatory calculus, we denote by $\mathrm{nf}(t)$ the (unique) normal form of t. If t is a closed term of type σ^*, we denote by t^{s} the set $SM(\mathrm{nf}(t))$. By Proposition 3, this is a nonempty finite set of closed normal terms of type σ. Clearly, if t and q have the same normal form, then $t^{\mathrm{s}} = q^{\mathrm{s}}$.

The following lemma is crucial:

Lemma 5 *Let t be a closed normal term of type σ.*

(a) $t \in (\mathfrak{s}t)^{\mathrm{s}}$.
(b) Let q and r be closed terms of type σ^*. If $t \in q^{\mathrm{s}}$ or $t \in r^{\mathrm{s}}$, then $t \in (\cup qr)^{\mathrm{s}}$.
(c) Let q be a closed term of type τ^* and r be a closed term of type $\tau \to \sigma^*$. Suppose that u is a closed term of type τ, and assume that $u \in q^{\mathrm{s}}$ and $t \in (ru)^{\mathrm{s}}$. Then $t \in (\bigcup qr)^{\mathrm{s}}$.

Proof For (a), note that $(\mathfrak{s}t)^{\mathrm{s}} = SM(\mathrm{nf}(\mathfrak{s}t)) = SM(\mathfrak{s}(\mathrm{nf}(t))) = SM(\mathfrak{s}t) = \{t\}$. For (b), just observe that

$$(\cup qr)^{\mathrm{s}} = SM(\mathrm{nf}(\cup qr)) = SM(\cup \mathrm{nf}(q)\mathrm{nf}(r)) = SM(\mathrm{nf}(q)) \cup SM(\mathrm{nf}(r)) = q^{\mathrm{s}} \cup r^{\mathrm{s}}.$$

In order to show (c) we may assume, without loss of generality, that q is normal. This is clear because $q^s = (\mathrm{nf}(q))^{\mathrm{s}}$ and $(\bigcup qr)^{\mathrm{s}} = (\bigcup \mathrm{nf}(q)r)^{\mathrm{s}}$. The proof is by induction on the complexity of the closed normal term q. By Proposition 3, q is set-like. Suppose that q is $\mathfrak{s}v$ for a given term v of type τ. We have $q^{\mathrm{s}} = SM(\mathrm{nf}(q)) = SM(q) = SM(\mathfrak{s}v) = \{v\}$. Since, by hypothesis, $u \in q^{\mathrm{s}}$, we conclude that u and v are the same term. But,

$$\left(\bigcup qr\right)^{\mathrm{s}} = \left(\bigcup (\mathfrak{s}v)r\right)^{\mathrm{s}} = SM\left(\mathrm{nf}(\bigcup (\mathfrak{s}v)r)\right) = SM(\mathrm{nf}(rv)) = (rv)^{\mathrm{s}}.$$

Also by hypothesis, $t \in (ru)^{\mathrm{s}}$. Hence, $t \in (rv)^{\mathrm{s}}$ and, therefore, $t \in (\bigcup qr)^{\mathrm{s}}$.

It remains to see the case when $q = \cup q_1 q_2$, for given terms q_1 and q_2 (necessarily closed and normal) of type τ^*. It is easy to see that $q^{\mathrm{s}} = q_1^{\mathrm{s}} \cup q_2^{\mathrm{s}}$. Moreover

$$\left(\bigcup qr\right)^{s} = SM\left(\mathrm{nf}(\bigcup qr)\right) = SM\left(\mathrm{nf}(\cup(\bigcup q_1 r)(\bigcup q_2 r))\right) =$$

$$SM\left(\cup \mathrm{nf}(\bigcup q_1 r)\mathrm{nf}(\bigcup q_2 r)\right) = SM\left(\mathrm{nf}(\bigcup q_1 r)\right) \cup SM\left(\mathrm{nf}(\bigcup q_2 r)\right) =$$

$$\left(\bigcup q_1 r\right)^{\mathrm{s}} \cup \left(\bigcup q_2 r\right)^{\mathrm{s}}$$

By hypothesis, $u \in q^{\mathrm{s}}$. Without loss of generality, we assume that $u \in q_1^{\mathrm{s}}$. Since $t \in (ru)^{\mathrm{s}}$ we get, by induction hypothesis, that $t \in (\bigcup q_1 r)^s$. By the above equality, $t \in (\bigcup qr)^s$. □

Let us finish this section with the definitions of two important notions.

Definition 6 We define a map $\phi \mapsto \phi^{\mathrm{p}}$ from *sentences* ϕ of the bounded mixed language $\mathcal{L}^{\mathrm{mix}}$ to formulas ϕ^{p} of the propositional calculus of $\mathcal{L}$ according to the following clauses:

(a) $(P(t_1,\ldots,t_n))^{\mathrm{p}}$ is $P(\mathrm{nf}(t_1),\ldots,\mathrm{nf}(t_n))$, where P is a n-ary predicate symbol of $\mathcal{L}$ and $t_1,\ldots,t_n$ are *closed* terms of ground type of the star combinatory calculus.
(b) $(\neg\phi)^{\mathrm{p}}$ is $\neg(\phi^{\mathrm{p}})$, $(\phi\wedge\psi)^{\mathrm{p}}$ is $\phi^{\mathrm{p}}\wedge\psi^{\mathrm{p}}$, and $(\phi\vee\psi)^{\mathrm{p}}$ is $\phi^{\mathrm{p}}\vee\psi^{\mathrm{p}}$.
(c) $(\forall x\in t\,\phi(x))^{\mathrm{p}}$ is $\bigwedge_{q\in t^s}\phi(q)^{\mathrm{p}}$ and $(\exists x\in t\,\phi(x))^{\mathrm{p}}$ is $\bigvee_{q\in t^s}\phi(q)^{\mathrm{p}}$, where t is a closed term of star type.

The following lemma is straightforward:

Lemma 6 *Let $\phi(x_1,\ldots,x_n)$ be a formula of $\mathcal{L}^{\mathrm{mix}}$ with its free variables among $x_1,\ldots,x_n$. Let $t_1,\ldots,t_n$ and $q_1,\ldots,q_n$ be closed terms of appropriate types. If, for each i with $1\leq i\leq n$, the terms t_i and q_i have the same normal form, then $\phi(t_1,\ldots,t_n)^{\mathrm{p}}$ is the* same *propositional formula as $\phi(q_1,\ldots,q_n)^{\mathrm{p}}$.*

In the next section, we will need to work with open formulas of $\mathcal{L}^{\mathrm{mix}}$. The following definition is important:

Definition 7 Let $\phi(x_1,\ldots,x_n)$ be a formula of $\mathcal{L}^{\mathrm{mix}}$ with its free variables among $x_1,\ldots,x_n$. We say that $\phi(x_1,\ldots,x_n)$ is *tautological in character* if, for all closed normal terms $t_1,\ldots,t_n$ (of appropriate types) of the combinatory star calculus, $\phi(t_1,\ldots,t_n)^{\mathrm{p}}$ is a tautology.

4 The herbrandized interpretation

We present in this section a herbrandized functional interpretation of pure classical logic. We work with the primitives $\neg$, $\vee$ and $\forall$. The other connectives are defined (classically) as usual.

Definition 8 To each formula ϕ of the first-order language $\mathcal{L}$ we assign formulas ϕ^{SH} and ϕ_{SH} so that ϕ^{SH} is of the form $\forall a\exists b\,\phi_{\mathrm{SH}}(a,b)$, with $\phi_{\mathrm{SH}}(a,b)$ a bounded mixed formula, according to the following clauses:

1. ϕ^{SH} and ϕ_{SH} are ϕ, for quantifier-free formulas ϕ of $\mathcal{L}$.

For the *remaining* cases, if we have already interpretations for ϕ and ψ given (respectively) by $\forall a\exists b\,\phi_{\mathrm{SH}}(a,b)$ and $\forall d\exists e\,\psi_{\mathrm{SH}}(d,e)$ then we define:

3. $(\phi\vee\psi)^{\mathrm{SH}}$ is $\forall a,d\,\exists b,e\,[\phi_{\mathrm{SH}}(a,b)\vee\psi_{\mathrm{SH}}(d,e)]$,
4. $(\neg\phi)^{\mathrm{SH}}$ is $\forall f\exists a'\,[\exists a\in a'\,\neg\phi_{\mathrm{SH}}(a,fa)]$,
5. $(\forall x\,\phi(x))^{\mathrm{SH}}$ is $\forall x,a\exists b\,[\phi_{\mathrm{SH}}(x,a,b)]$,

where the bounded mixed formulas between square brackets are the corresponding *lower* SH-formulas.

In the above, the letters a, b, d and e stand for (possibly empty) tuples of variables of the star combinatory calculus, but we speak of them as if they were only a single variable. The formulas $\phi_{\mathrm{SH}}(a,b)$ also have the free first-order variables of the interpreted formula ϕ (note that first-order variables can be

considered ground type variables of the star combinatory calculus). We omit showing them unless it is convenient or necessary. In these cases, we often use a single (first-order) variable to stand for a tuple.

We have used unbounded quantifications with respect to variables of the star combinatory calculus. Our language $\mathcal{L}^{\mathrm{mix}}$ does not have such quantifications, but the use is incidental. It is customary in functional interpretations to use such expressions because they help convey the meaning of the lower SH-formulas. These bounded mixed formulas are the ones that have real meaning (together with the information saying which variables are in universal or existential positions). When $\phi(x)$ is a quantifier-free first-order formula of $\mathcal{L}$, it is worth noting that $(\exists x \phi(x))^{\mathrm{SH}}$ is $\exists a^{G^*} \exists x \in a\, \phi(x)$, modulo classical logic in the lower SH-formulas. This is seen by computing the interpretation of $\neg\forall x \neg\phi(x)$ according to the clauses (1), (4) and (5) of Definition 8.

We now turn to the soundness theorem. The original plan of the paper was to state soundness with a semantical verification of the interpreting formulas. The verification would say that certain formulas of $\mathcal{L}^{\mathrm{mix}}$ are true in the finite-order set-theoretical structure naturally arising from a given first-order structure for $\mathcal{L}$ (see definition 6 of [9]). Actually, in [9], the verifications are done within a finite-order logical theory (of which the set-theoretical finite-order stuctures are models). When writing this paper, it dawned on the first author that the verification can be done in a particularly nice way, via the notion of tautology in character of Section 3. This is how we will state the soundness theorem below.

As it is usual with functional interpretations which accumulate witnesses, we have a monotonicity property. For the next lemma, the reader should notice that the types of the existential variables b in $\phi_{\mathrm{SH}}(a, b)$ are necessarily of star type (the universal variables a need not be of star type).

Lemma 7 (Monotonicity) *Let t and q be closed terms of star type σ^*, and let $\phi(x)$ be a first-order formula of $\mathcal{L}$. Suppose that $t^{\mathrm{s}} \subseteq q^{\mathrm{s}}$. Then the implication $\phi_{\mathrm{SH}}(x, a, t) \to \phi_{\mathrm{SH}}(x, a, q)$ is tautological in character.*

Proof Under the hypothesis of the lemma, we must show that the bounded mixed formula $\neg\phi_{\mathrm{SH}}(x, a, t) \vee \phi_{\mathrm{SH}}(x, a, q)$ is tautological in character. The proof is straightforward by induction on the complexity of ϕ, but the negation case is proved directly. In the negation case, we must show that the bounded mixed formula $\neg\exists a \in t\, \neg\phi_{\mathrm{SH}}(x, a, fa) \vee \exists a \in q\, \neg\phi_{\mathrm{SH}}(x, a, fa)$ is tautological in character. Let s be a closed first-order term and r closed normal terms of appropriate types. We must argue that

$$(\neg\exists a \in t\, \neg\phi_{\mathrm{SH}}(s, a, ra) \vee \exists a \in q\, \neg\phi_{\mathrm{SH}}(s, a, ra))^{\mathrm{p}}$$

is a tautology. By definition, this propositional formula of $\mathcal{L}$ is

$$\neg \bigvee_{u \in t^{\mathrm{s}}} \neg\phi_{\mathrm{SH}}(s, u, ru)^{\mathrm{p}} \vee \bigvee_{u \in q^{s}} \neg\phi_{\mathrm{SH}}(s, u, ru)^{\mathrm{p}}$$

The above is clearly a tautology because the last disjunction includes the disjuncts of the first disjunction. □

It is this monotonicity property that distinguishes our interpretation as a cumulative interpretation in the sense discussed in the introduction. The accumulation is, of course, done into finite sets. The *intuitionistic* interpretations of nonstandard arithmetic of Berg *et al.* in [3] also enjoy a similar monotonicity property (curiously, their classical interpretation is not cumulative). Of course, there is an extra ingredient in [3] that is totally absent in our interpretation. Since Berg *et al.* have to deal with nonstandardness, they have a "compactness" feature incorporated in their interpretation in order to deal with the unrestricted quantifiers (by the way, this feature is also present in the bounded functional interpretation [8] in the interpretation of the bounded quantifiers). With respect to the logical part, the manner in which the existential witnesses are accumulated in our interpretation and in the functional interpretations of [3] differs slightly. In the treatment of [3] there is an identification of elements of $(\sigma \to \rho^*)^*$ with elements of $\sigma \to \rho^*$, whereas in our treatment no such identification takes place. The reader can consult [4] for the modification of the interpretations of [3] along the lines underlying the interpretation of the present paper.

Theorem 3 (Soundness) *Let ϕ be a sentence of $\mathcal{L}$ and suppose that ϕ is provable in classical first-order predicate logic without equality. Then there are closed terms t (of appropriate types) of the star combinatory language such that the bounded mixed formula $\phi_{\mathrm{SH}}(a, ta)$ is tautological in character.*

Proof Let $\phi(x)$ be a first-order formula of $\mathcal{L}$ whose free variables are among x (this letter may stand for a tuple of variables). We show that if $\phi(x)$ is provable in classical first-order predicate logic, then there are terms t, whose free variables are among x, such that the bounded mixed formula $\phi_{\mathrm{SH}}(x, a, ta)$ is tautological in character. We consider a suitable logical calculus and prove the result by induction on the number of inferences of the derivation. We use a calculus due to Shoenfield, as described in sections 2.6 and 8.3 of [12]. The calculus consists of two axiom schemas

- Excluded middle: $\neg\phi \vee \phi$
- Substitution: $\forall x \phi(x) \to \phi(t)$

and five rules

- Expansion: from ϕ infer $\psi \vee \phi$
- Contraction: from $\phi \vee \phi$ infer ϕ
- Associativity: from $\phi \vee (\psi \vee \gamma)$ infer $(\phi \vee \psi) \vee \gamma$
- Cut: from $\phi \vee \psi$ and $\neg\phi \vee \gamma$ infer $\psi \vee \gamma$
- $\forall$-introduction: from $\phi(x) \vee \psi$ infer $\forall x \phi(x) \vee \psi$, provided that x does not occur free in ψ

We will present the witnessing terms that do the job. The verification that these terms work (i.e., that they yield formulas tautological in character) is straightforward using Lemma 5, Lemma 6 and the monotonicity lemma (Lemma 7). We will, nevertheless, present the details of this verification for the cut rule (this is the most involved case).

Suppose ϕ is a first-order formula and that ϕ^{SH} is $\forall a \exists b\, \phi_{\mathrm{SH}}(a,b)$. It is easy to see that $(\neg\phi \vee \phi)^{\mathrm{SH}}$ is $\forall f, d \exists a', e[\exists a \in a' \neg\phi_{\mathrm{SH}}(a, fa) \vee \phi_{\mathrm{SH}}(d,e)]$. We must find terms t and q such that

$$\exists a \in tfd\, \neg\phi_{\mathrm{SH}}(a, fa) \ \vee\ \phi_{\mathrm{SH}}(d, qfd)$$

is tautological in character. It is easy to see that $t :\equiv \lambda f, d.\mathfrak{s}d$ and $q :\equiv \lambda f, d.fd$ work.

The SH-interpretation of the substitution axiom for the formula ϕ is (modulo classical logic in the lower SH-formulas)

$$\forall f, \tilde{a} \exists c^{G^*} \exists a', b\, [\forall x \in c \forall a \in a'\, \phi_{\mathrm{SH}}(x, a, fxa) \to \phi_{\mathrm{SH}}(t, \tilde{a}, b)].$$

We must present terms p, q and r such that

$$\forall x \in pf\tilde{a}\, \forall a \in qf\tilde{a}\, \phi_{\mathrm{SH}}(x, a, fxa) \to \phi_{\mathrm{SH}}(t, \tilde{a}, rf\tilde{a})$$

is tautological in character. The terms $p :\equiv \lambda f, \tilde{a}.\mathfrak{s}t$, $q :\equiv \lambda f, \tilde{a}.\mathfrak{s}\tilde{a}$ and $r :\equiv \lambda f, \tilde{a}.ft\tilde{a}$ do the job.

Expansion is immediate provided that we show that every type is inhabited by a closed term. Given a starting constant symbol c of $\mathcal{L}$ (by hypothesis, $\mathcal{L}$ has at least a constant symbol), we can build systematically an inhabitant c^σ for every type σ: c^G is just c; $c^{\tau\to\rho}$ is $\lambda x^\tau.c^\rho$; and c^{τ^*} is $\mathfrak{s}c^\tau$. Associativity is trivial.

Let us look at the contraction rule. Suppose that, by induction hypothesis, there are terms t and q such that

$$\phi_{\mathrm{SH}}(a, tad) \vee \phi_{\mathrm{SH}}(d, qad)$$

is tautological in character. Then, $\phi_{\mathrm{SH}}(a, \cup(taa)(qaa))$ is tautological in character. For the verification, we use the monotonicity lemma (and (b) of Lemma 5).

The $\forall$-introduction rule does not present any trouble (the terms that witness the consequent are essentially the same that witness the antecedent).

We now discuss the cut rule. Suppose that ψ^{SH} is $\forall d \exists e\, \psi_{\mathrm{SH}}(d,e)$ and that γ^{SH} is $\forall u \exists v\, \gamma_{\mathrm{SH}}(u,v)$. By induction hypothesis, there are terms t, q, r and s such that

$$\text{(I)} \qquad \phi_{\mathrm{SH}}(a, tad) \ \vee\ \psi_{\mathrm{SH}}(d, qad)$$

and

$$\text{(II)} \qquad \exists a \in rfu\, \neg\phi_{\mathrm{SH}}(a, fa) \ \vee\ \gamma_{\mathrm{SH}}(u, sfu)$$

are tautological in character. Take the terms $k :\equiv \lambda d, u. \bigcup(r(\lambda a.tad)u)(\lambda w.qwd)$ and $l :\equiv \lambda d, u.s(\lambda a.tad)u$. We claim that the bounded mixed formula

$$\psi_{\mathrm{SH}}(d, kdu) \vee \gamma_{\mathrm{SH}}(u, ldu)$$

is tautological in character. To see this, we must show that, for all closed normal terms $\tilde{d}$ and $\tilde{u}$ of appropriate types, $(\psi_{\mathrm{SH}}(\tilde{d}, k\tilde{d}\tilde{u}) \vee \gamma_{\mathrm{SH}}(\tilde{u}, l\tilde{d}\tilde{u}))^{\mathrm{p}}$ is a

tautology (we are ignoring first-order parameters). By (II), the propositional formula

$$\left(\bigvee_{\tilde{a}\in(r(\lambda a.ta\tilde{d})\tilde{u})^{\mathrm{s}}} \neg\, \phi_{\mathrm{SH}}(\tilde{a}, t\tilde{a}\tilde{d})^{\mathrm{p}} \right) \vee\ \gamma_{\mathrm{SH}}(\tilde{u}, s(\lambda a.ta\tilde{d})\tilde{u})^{\mathrm{p}}$$

is a tautology. By (I), we know that for each of the finitely-many closed (normal) terms $\tilde{a}$ such that $\tilde{a} \in (r(\lambda a.ta\tilde{d})\tilde{u})^{\mathrm{s}}$,

$$\phi_{\mathrm{SH}}(\tilde{a}, t\tilde{a}\tilde{d})^{\mathrm{p}} \vee \psi_{\mathrm{SH}}(\tilde{d}, q\tilde{a}\tilde{d})^{\mathrm{p}}$$

is a tautology. It is clear that all these tautologies (tautologically) imply the propositional formula

$$\left(\bigvee_{\tilde{a}\in(r(\lambda a.ta\tilde{d})\tilde{u})^{\mathrm{s}}} \psi_{\mathrm{SH}}(\tilde{d}, q\tilde{a}\tilde{d})^{\mathrm{p}} \right) \vee\ \gamma_{\mathrm{SH}}(\tilde{u}, s(\lambda a.ta\tilde{d})\tilde{u})^{\mathrm{p}}$$

Therefore, the above formula is a tautology. Let us study each one of the disjuncts above. The last one is (by Lemma 6) the formula $\gamma_{\mathrm{SH}}(\tilde{u}, l\tilde{d}\tilde{u})^{\mathrm{p}}$. For each of the finitely-many terms $\tilde{a}$ in $(r(\lambda a.ta\tilde{d})\tilde{u})^{\mathrm{s}}$, we have $(q\tilde{a}\tilde{d})^{\mathrm{s}} \subseteq (\bigcup(r(\lambda a.ta\tilde{d})\tilde{u})(\lambda w.qw\tilde{d}))^{\mathrm{s}}$. This is a consequence of (c) of Lemma 5: just notice that if a closed normal term p is in $(q\tilde{a}\tilde{d})^{\mathrm{s}}$ then $p \in ((\lambda w.qw\tilde{d})\tilde{a})^{\mathrm{s}}$. Hence, by the monotonicity lemma,

$$\psi_{\mathrm{SH}}(\tilde{d}, q\tilde{a}\tilde{d})^{\mathrm{p}} \to \psi_{\mathrm{SH}}(\tilde{d}, k\tilde{d}\tilde{u})^{\mathrm{p}}$$

is a tautology. It now follows that $\psi_{\mathrm{SH}}(\tilde{d}, k\tilde{d}\tilde{u})^{\mathrm{p}} \vee \gamma_{\mathrm{SH}}(\tilde{u}, l\tilde{d}\tilde{u})^{\mathrm{p}}$ is a tautology, as wanted. □

Example 1 Consider the first-order validity $\phi :\equiv \exists x \forall y\, (Px \vee \neg Py)$, where P is a unary predicate. The SH-translation of ϕ is $\forall f \exists a\, [\exists x \in a \forall y \in fx\, (Px \vee \neg Py)]$, modulo classical logic in the lower SH-formulas. By the above theorem, it is possible to extract from a proof of ϕ a closed term $\Phi : (G \to G^*) \to G^*$ that "realizes" a in f . The closed term $\Phi :\equiv \lambda f. \cup (\mathfrak{s}c)(fc)$ does the job (here c is some constant of the language). So, we are claiming that

$$\exists x \in \cup(\mathfrak{s}c)(fc)\, \forall y \in fx\, (Px \vee \neg Py)$$

is tautological in character. This means that, for every closed normal term q of type $G \to G^*$,

$$\bigvee_{t\in(\cup(\mathfrak{s}c)(qc))^s} \bigwedge_{r\in(qt)^s} (Pt \vee \neg Pr)$$

is a tautology. If $(qc)^s$ is the finite set $\{t_1, \ldots, t_k\}$ of closed first-order terms, then $(\cup(\mathfrak{s}c)(qc))^s = \{c, t_1, \ldots, t_k\}$ and the above formula is

$$\bigwedge_{i=1}^{k} (Pc \vee \neg Pt_i) \vee \bigvee_{j=1}^{k} \bigwedge_{r\in(qt_j)^s} (Pt_j \vee \neg Pr)$$

This is classically equivalent to

$$Pc \vee \bigwedge_{i=1}^{k} \neg Pt_i \vee \bigvee_{j=1}^{k} Pt_j \vee \bigvee_{j=1}^{k} \bigwedge_{r \in (qt_j)^s} \neg Pr$$

which is obviously a tautology.

The way our interpretation works for this example is somewhat different from the Gerhardy-Kohlenbach analysis in [9]. Compare also with Example 3 below.

Example 2 Let us now consider the first-order validity

$$\phi := \exists x \forall w \exists y \exists z \, ((Q(c,y) \vee Q(d,z)) \to Q(x,w))$$

where c and d are contants and Q is a binary predicate. The SH-translation is $\forall \Phi \exists X, F, H \, \phi_{\text{SH}}(\Phi, X, F, H)$, where $\phi_{\text{SH}}(\Phi, X, F, H)$ is (modulo classical logic) the following bounded mixed formula:

$$\exists x \in X, f \in F, h \in H \, \forall w \in \Phi x f h \, \exists y \in f w, z \in h w \, ((Q(c,y) \vee Q(d,z)) \to Q(x,w))$$

Let σ be the type of Φ, i.e., σ is $G \to (G \to G^*) \to (G \to G^*) \to G^*$. By the above theorem, there are closed terms $t : \sigma \to G^*$ and q and r of type $\sigma \to (G \to G^*)^*$ which "realize" X, F and H in Φ (respectively). The following terms do the job: $t = \lambda \Phi . \cup (\mathfrak{s} c)(\mathfrak{s} d)$, $q = \lambda \Phi . \cup (\mathfrak{s} p)(\mathfrak{s}(\lambda w^G . \Phi c p p))$ and $r = \lambda \Phi . \cup (\mathfrak{s} p)(\mathfrak{s}(\lambda w^G . \Phi d p p))$. Here p can be any closed term of appropriate type. For definiteness, we can take $p = \lambda w^G . (\mathfrak{s} w)$.

The verification that the above terms do the job hinges (informally) on a discussion by cases, by considering whether (or not) there are $y_0 \in \Phi(cpp)$ and $z_0 \in \Phi(dpp)$ such that $\neg Q(c, y_0)$ and $\neg Q(d, z_0)$.

We finish the paper with a new proof of Herbrand's theorem:

Corollary 1 (Herbrand's theorem) *Let $\phi(x)$ be a quantifier-free formula of the first-order language $\mathcal{L}$, with x as the only free variable. Suppose that $\exists x \, \phi(x)$ is a theorem of classical logic. Then there are closed first-order terms $r_1, \ldots, r_n$ such that $\phi(r_1) \vee \ldots \vee \phi(r_n)$ is a tautology.*

Proof We remarked that $(\exists x \, \phi(x))^{\text{SH}}$ is $\exists a \exists x \in a \, \phi(x)$. By the soundness theorem, there is a closed term t of type G^* such that $(\exists x \in t \, \phi(x))^{\text{P}}$ is a tautology. This means that

$$\bigvee_{r \in t^s} \phi(r)$$

is a tautology. We are done. □

Example 3 Let ϕ be the first-order validity $\exists x \, (P(x) \vee \neg P(f(x)))$, where P is a unary predicate symbol and f is a unary function symbol. From a proof of ϕ it is possible to extract a closed term t of type G^* that "realizes" the bounded mixed formula $\exists x \in t \, (P(x) \vee \neg P(f(x)))$. It is clear that we can take t to be $\cup(\mathfrak{s} c)(\mathfrak{s}(f c))$, where c is a constant of the language.

The procedure underlying the construction of the terms in the proof of the soundness theorem is of low complexity. The terms extracted need not be in normal form, though. The well-known super-exponential feature on the number of terms in Herbrand's theorem (cf. [13]) is due to the fact that they are obtained by carrying out the *normalization* of the extracted term t of type G^*.

References

1. J. Avigad and S. Feferman. Gödel's functional ("Dialectica") interpretation. In S. R. Buss, editor, *Handbook of Proof Theory*, volume 137 of *Studies in Logic and the Foundations of Mathematics*, pages 337–405. North Holland, Amsterdam, 1998.
2. J. Avigad and H. Towsner. Functional interpretation and inductive definitions. *The Journal of Symbolic Logic*, 74(4):1100–1120, 2009.
3. B. van den Berg, E. Briseid, and P. Safarik. A functional interpretation for nonstandard arithmetic. *Annals of Pure and Applied Logic*, 163(12):1962–1994, 2012.
4. A. Borges. On the herbrandised interpretation for nonstandard arithmetic. Master's thesis, Universidade de Lisboa, 2016.
5. J. Diller. Logical problems of functional interpretations. *Annals of Pure and Applied Logic*, 114:27–42, 2002.
6. J. Diller and W. Nahm. Eine Variante zur Dialectica-Interpretation der Heyting-Arithmetik endlicher Typen. *Archive für mathematische Logik und Grundlagenforschung*, 16:49–66, 1974.
7. F. Ferreira and A. Nunes. Bounded modified realizability. *The Journal of Symbolic Logic*, 71:329–346, 2006.
8. F. Ferreira and P. Oliva. Bounded functional interpretation. *Annals of Pure and Applied Logic*, 135:73–112, 2005.
9. P. Gerhardy and U. Kohlenbach. Extracting Herbrand disjunctions by functional interpretation. *Archive for Mathematical Logic*, 44:633–644, 2005.
10. K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *dialectica*, 12:280–287, 1958. Reprinted with an English translation in [11], pp. 240-251.
11. K. Gödel. *Collected Works, Vol. II.* S. Feferman *et al.*, eds. Oxford University Press, Oxford, 1990.
12. J. R. Shoenfield. *Mathematical Logic.* Addison-Wesley Publishing Company, 1967. Republished in 2001 by AK Peters.
13. R. Statman. Lower bounds on Herbrand's theorem. *Proceedings of the American Mathematical Society*, 75(1):104–107, 1979.
14. W. Tait. Intentional interpretations of functionals of finite type I. *The Journal of Symbolic Logic*, 32:198–212, 1967.
15. A. S. Troelstra and H. Schwichtenberg. *Basic Proof Theory.* Cambridge University Press, Cambridge, 1996.
16. A. S. Troelstra (ed.). *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1973.