

## EXTENSIONAL SEMANTICS FOR HIGHER-ORDER LOGIC PROGRAMS WITH NEGATION\*

PANOS RONDOGIANNIS AND IOANNA SYMEONIDOU

Department of Informatics & Telecommunications, National and Kapodistrian University of Athens,  
Greece

*e-mail address:* {prondo,i.symeonidou}@di.uoa.gr

**ABSTRACT.** We develop an extensional semantics for higher-order logic programs with negation, generalizing the technique that was introduced in [2, 3] for positive higher-order programs. In this way we provide an alternative extensional semantics for higher-order logic programs with negation to the one proposed in [6]. We define for the language we consider the notions of *stratification* and *local stratification*, which generalize the familiar such notions from classical logic programming, and we demonstrate that for stratified and locally stratified higher-order logic programs, the proposed semantics never assigns the *unknown* truth value. We conclude the paper by providing a negative result: we demonstrate that the well-known stable model semantics of classical logic programming, if extended according to the technique of [2, 3] to higher-order logic programs, does not in general lead to extensional stable models.

### 1. INTRODUCTION

Research results developed in [23, 2, 3, 16, 7] have explored the possibility of designing higher-order logic programming languages with *extensional semantics*. Extensionality implies that program predicates essentially denote sets, and therefore one can use standard set theoretic concepts in order to understand the meaning of programs and reason about them. The key idea behind this line of research is that if we appropriately restrict the syntax of higher-order logic programming, then we can achieve extensionality and obtain languages that are simple both from a semantic as-well-as from a proof-theoretic point of view. Therefore, a main difference between the extensional and the more traditional *intensional* higher-order logic programming languages [19, 9] is that the latter have richer syntax and expressive capabilities but a non-extensional semantics.

There exist at present two main extensional semantic approaches for capturing the meaning of positive (i.e., negationless) higher-order logic programs. The first approach, developed in [23, 16, 7], uses classical domain-theoretic tools. The second approach, developed in [2, 3], uses a fixed-point construction on the ground instantiation of the source program.

*Key words and phrases:* Higher-Order Logic Programming, Negation in Logic Programming, Extensional Semantics.

\* A preliminary version of this paper has appeared in the proceedings of the 15th European Conference on Logics in Artificial Intelligence (JELIA), pages 447–462, 2016.

Despite their different philosophies, these two approaches have recently been shown [8] to agree for a broad and useful class of programs. This fact suggests that the two aforementioned techniques can be employed as useful alternatives for the further development of higher-order logic programming.

A natural question that arises is whether one can still obtain an extensional semantics if negation is added to programs. This question was recently undertaken in [6], where it was demonstrated that the domain-theoretic results obtained for positive logic programs in [23, 16, 7], can be extended to apply to programs with negation. More specifically, as demonstrated in [6], every higher-order logic program with negation has a distinguished extensional model constructed over a logic with an infinite number of truth values. It is therefore natural to wonder whether the alternative extensional technique introduced in [2, 3], can also be extended to higher-order logic programs with negation. It is exactly this question that we answer affirmatively. This brings us to the following contributions of the present paper:

- We extend the technique of [2, 3] to the class of higher-order logic programs with negation. In this way we demonstrate that Bezem’s approach is more widely applicable than possibly initially anticipated. Our extension relies on the *infinite-valued semantics* [20], a technique that was developed in order to provide a purely model-theoretic semantics for negation in classical logic programming.
- The extensional semantics we propose appears to be simpler compared to [6] because it relies on the ground instantiation of the higher-order program and does not require the rather involved domain-theoretic constructions of [6]. However, each technique has its own merits and we believe that both will prove to be useful tools in the further study of higher-order logic programming.
- As a case study of the applicability of the new semantics, we define the notions of *stratification* and *local stratification* for higher-order logic programs with negation and demonstrate that for such programs the proposed semantics never assigns the *unknown* truth value. These two new notions generalize the corresponding ones from classical (first-order) logic programming. It is worth mentioning that such notions have not yet been studied under the semantics of [6].
- We demonstrate that not all semantic approaches that have been successful for classical logic programs with negation lead to extensional semantics when applied to higher-order logic programming under the framework of [2, 3]. In particular we demonstrate that the well-known stable model semantics of classical logic programming [15], if extended according to the technique of [2, 3] to higher-order logic programs with negation, does not in general lead to extensional stable models.

The rest of the paper is organized as follows. Section 2 presents in an intuitive way the semantics that will be developed in this paper. Section 3 contains background material on the infinite-valued semantics that will be the basis of our construction. Section 4 introduces the syntax and Section 5 the semantics of our source language. Section 6 demonstrates that the proposed semantics is extensional. In Section 7 the notions of *stratification* and *local stratification* for higher-order logic programs are introduced. Section 8 demonstrates that the stable model semantics do not in general lead to extensional stable models when applied to higher-order logic programs. Section 9 concludes the paper by discussing the connections of the present work with that of Zoltán Ésik, to whom the present special issue is devoted, and by providing pointers to future work.

The present paper extends and revises the conference paper [21]. More specifically, the present paper contains complete proofs of all the claimed results (which were either missing or sketched in [21]), it contains the new Section 8, and it has more detailed and polished material in most of the remaining sections.

## 2. AN INTUITIVE OVERVIEW OF THE PROPOSED APPROACH

In this paper we consider the semantic technique for positive higher-order logic programs proposed in [2, 3], and we extend it in order to apply to programs with negation in clause bodies. Given a positive higher-order logic program, the starting idea behind Bezem’s approach is to take its “ground instantiation”, in which we replace variables with well-typed terms of the Herbrand Universe of the program (i.e., terms that can be created using only predicate constants, function symbols, and individual constants that appear in the program). For example, consider the higher-order program below (for the moment we use ad-hoc Prolog-like syntax):

$$\begin{aligned} & q(a) . \\ & q(b) . \\ & p(Q) : \neg Q(a) . \\ & id(R)(X) : \neg R(X) . \end{aligned}$$

In order to obtain the ground instantiation of this program, we consider each clause and replace each variable of the clause with a ground term that has the same type as the variable under consideration (the formal definition of this procedure will be given in Definition 4.9). In the above example, the variable  $X$  represents a data object, so it could be replaced by  $a$  or  $b$ , namely the only individual constant symbols that appear in the program. The variables  $P$  and  $Q$  have the same type, as they both represent a predicate that maps data objects to truth values (in other words, a unary first-order predicate). Each of them could be replaced by any such predicate that appears in the program, such as  $q$ , or a more complex expression that also maps data objects to truth values, such as  $id(q)$ . Therefore, the above program defines a relation  $q$  that is true of  $a$  and  $b$ ; a relation  $p$  that is true of a unary first-order relation  $Q$  only if  $Q$  is true of  $a$ ; and a relation  $id$  that is true of a unary first-order relation  $R$  and a data object  $X$  only if  $R$  is true of  $X$ . The ground instantiation we obtain is the following infinite program:

$$\begin{aligned} & q(a) . \\ & q(b) . \\ & p(q) : \neg q(a) . \\ & id(q)(a) : \neg q(a) . \\ & id(q)(b) : \neg q(b) . \\ & p(id(q)) : \neg id(q)(a) . \\ & id(id(q))(a) : \neg id(q)(a) . \\ & id(id(q))(b) : \neg id(q)(b) . \\ & \dots \end{aligned}$$

One can now treat the new program as an infinite propositional one (i.e., each ground atom can be seen as a propositional variable). This implies that we can use the standard least fixed-point construction of classical logic programming (see for example [18]) in order to compute the set of atoms that should be taken as “true”. In our example, the least fixed-point will contain atoms such as  $q(a)$ ,  $q(b)$ ,  $p(q)$ ,  $id(q)(a)$ ,  $id(q)(b)$ ,  $p(id(q))$ , and so on.

The main contribution of Bezem’s work was that he established that the least fixed-point semantics of the ground instantiation of every positive higher-order logic program of the language considered in [2, 3], is *extensional* in a sense that can be intuitively explained as follows. In the above example  $q$  and  $\text{id}(q)$  can be considered equal since they are both true of exactly the constants  $a$  and  $b$ . Therefore, we would expect that (for example) if  $p(q)$  is true then  $p(\text{id}(q))$  is also true, because  $q$  and  $\text{id}(q)$  should be considered as interchangeable. This property of “interchangeability” is formally defined in [2, 3] and it is demonstrated that it holds in the least fixed-point of the immediate consequence operator of the ground instantiation of every program.

The key idea behind extending Bezem’s semantics in order to apply to higher-order logic programs with negation, is straightforward to state: given such a program, we first take its ground instantiation. The resulting program is a (possibly infinite) propositional program with negation, and therefore we can compute its semantics in any standard way that exists for obtaining the meaning of such programs. For example, one could use the well-founded semantics [14] or the stable model semantics [15], and then proceed to show that the well-founded model (respectively, each stable model) is extensional in the sense of [2, 3] (informally described above). Instead of using the well-founded or the stable model semantics, we have chosen to use a relatively recent proposal for assigning meaning to classical logic programs with negation, namely the infinite-valued semantics [20]. As it has been demonstrated in [20], the infinite-valued semantics is compatible with the well-founded: if we collapse the infinite-valued model to three truth values, we get the well-founded one. There are three main reasons for choosing to proceed with the infinite-valued approach (instead of the well-founded or the stable model approaches):

- Despite the close connections between the well-founded and the infinite-valued approaches, it has recently been demonstrated [22] that the well-founded-based adaptation of the technique of [2, 3] is *not* extensional.<sup>1</sup> Moreover, as we demonstrate in Section 8, a stable-models-based adaptation of Bezem’s technique, does *not* in general lead to extensional models.
- An extension of the infinite-valued approach was used in [6] to give the first extensional semantics for higher-order logic programs with negation. By developing our present approach using the same underlying logic, we facilitate the future comparison between the two approaches.
- As it was recently demonstrated in [13, 5], the infinite-valued approach satisfies all identities of *iteration theories* [4], while the well-founded semantics does not. Since iteration theories (intuitively) provide an abstract framework for the evaluation of the merits of various semantic approaches for languages that involve recursion, the results just mentioned give an extra incentive for the further study and use of the infinite-valued approach.

We demonstrate that the infinite-valued semantics of the ground instantiation of every higher-order logic program with negation, is extensional. In this way we extend the results of [2, 3] which applied only to positive programs. The proof of extensionality is quite intricate and is performed by a tedious induction on the approximations of the minimum infinite-valued model. As an immediate application of our result, we show how one can define the notions of *stratification* and *local stratification* for higher-order logic programs with negation.

---

<sup>1</sup>The result in [22] was obtained while the present paper was under review.

## 3. THE INFINITE-VALUED SEMANTICS

In this section we give an overview of the infinite-valued approach of [20]. As in [20], we consider (possibly countably infinite) propositional programs, consisting of clauses of the form  $\mathbf{p} \leftarrow \mathbf{L}_1, \dots, \mathbf{L}_n$ , where each  $\mathbf{L}_i$  is either a propositional variable or the negation of a propositional variable; the  $\mathbf{L}_i$  will be called *literals*, *negative* if they have negation and *positive* otherwise. For some technical reasons that will be explained just after Definition 4.9, we allow a positive literal  $\mathbf{L}_i$  to also be one of the constants **true** and **false**.

The key idea of the infinite-valued approach is that, in order to give a logical semantics to negation-as-failure and to distinguish it from ordinary negation, one needs to extend the domain of truth values. For example, consider the program:

$$\begin{array}{l} \mathbf{p} \leftarrow \\ \mathbf{r} \leftarrow \sim \mathbf{p} \\ \mathbf{s} \leftarrow \sim \mathbf{q} \\ \mathbf{t} \leftarrow \sim \mathbf{t} \end{array}$$

According to negation-as-failure, both  $\mathbf{p}$  and  $\mathbf{s}$  receive the value *True*. However,  $\mathbf{p}$  seems “truer” than  $\mathbf{s}$  because there is a rule which says so, whereas  $\mathbf{s}$  is true only because we are never obliged to make  $\mathbf{q}$  true. In a sense,  $\mathbf{s}$  is true only by default. For this reason, it was proposed in [20] to introduce a “default” truth value  $T_1$  just below the “real” true  $T_0$ , and (by symmetry) a weaker false value  $F_1$  just above (“not as false as”) the real false  $F_0$ . Then, negation-as-failure is a combination of ordinary negation with a weakening. Thus  $\sim F_0 = T_1$  and  $\sim T_0 = F_1$ . Since negations can be iterated, the new truth domain has a sequence  $\dots, T_3, T_2, T_1$  of weaker and weaker truth values below  $T_0$  but above a neutral value 0; and a mirror image sequence  $F_1, F_2, F_3, \dots$  above  $F_0$  and below 0. Since our propositional programs are possibly countably infinite, we need a  $T_\alpha$  and a  $F_\alpha$  for every countable ordinal  $\alpha$ . The intermediate truth value 0 is needed for certain atoms that have a “pathological” negative dependence on themselves (such as  $\mathbf{t}$  in the above program). In conclusion, our truth domain  $\mathbb{V}$  is shaped as follows:

$$F_0 < F_1 < \dots < F_\omega < \dots < F_\alpha < \dots < 0 < \dots < T_\alpha < \dots < T_\omega < \dots < T_1 < T_0$$

and the notion of “Herbrand interpretation of a program” can be generalized:

**Definition 3.1.** An (infinite-valued) *interpretation*  $I$  of a propositional program  $\mathbf{P}$  is a function from the set of propositional variables that appear in  $\mathbf{P}$  to the set  $\mathbb{V}$  of truth values.

For example, an infinite-valued interpretation for the program in the beginning of this section is  $I = \{(\mathbf{p}, T_3), (\mathbf{q}, F_0), (\mathbf{r}, 0), (\mathbf{s}, F_2), (\mathbf{t}, T_0)\}$ . As we are going to see later in this section, the interpretation that captures the meaning of the above program is  $J = \{(\mathbf{p}, T_0), (\mathbf{q}, F_0), (\mathbf{r}, F_1), (\mathbf{s}, T_1), (\mathbf{t}, 0)\}$ .

We will use  $\emptyset$  to denote the interpretation that assigns the  $F_0$  value to all propositional variables of a program. If  $v \in \mathbb{V}$  is a truth value, we will use  $I \parallel v$  to denote the set of variables which are assigned the value  $v$  by  $I$ . In order to define the notion of “model”, we need the following definitions:

**Definition 3.2.** Let  $I$  be an interpretation of a given propositional program  $\mathbf{P}$ . For every negative literal  $\sim \mathbf{p}$  appearing in  $\mathbf{P}$  we extend  $I$  as follows:

$$I(\sim \mathbf{p}) = \begin{cases} T_{\alpha+1}, & \text{if } I(\mathbf{p}) = F_\alpha \\ F_{\alpha+1}, & \text{if } I(\mathbf{p}) = T_\alpha \\ 0, & \text{if } I(\mathbf{p}) = 0 \end{cases}$$

Moreover,  $I(\text{true}) = T_0$  and  $I(\text{false}) = F_0$ . Finally, for every conjunction of literals  $L_1, \dots, L_n$  appearing as the body of a clause in  $P$ , we extend  $I$  by  $I(L_1, \dots, L_n) = \min\{I(L_1), \dots, I(L_n)\}$ .

**Definition 3.3.** Let  $P$  be a propositional program and  $I$  an interpretation of  $P$ . Then,  $I$  satisfies a clause  $\mathbf{p} \leftarrow L_1, \dots, L_n$  of  $P$  if  $I(\mathbf{p}) \geq I(L_1, \dots, L_n)$ . Moreover,  $I$  is a *model* of  $P$  if  $I$  satisfies all clauses of  $P$ .

As it is demonstrated in [20], every program has a *minimum* infinite-valued model under an ordering relation  $\sqsubseteq$ , which compares interpretations in a stage-by-stage manner. To formally state this result, the following definitions are necessary:

**Definition 3.4.** The *order* of a truth value is defined as follows:  $order(T_\alpha) = \alpha$ ,  $order(F_\alpha) = \alpha$  and  $order(0) = +\infty$ .

**Definition 3.5.** Let  $I$  and  $J$  be interpretations of a given propositional program  $P$  and  $\alpha$  be a countable ordinal. We write  $I =_\alpha J$ , if for all  $\beta \leq \alpha$ ,  $I \parallel T_\beta = J \parallel T_\beta$  and  $I \parallel F_\beta = J \parallel F_\beta$ . We write  $I \sqsubseteq_\alpha J$ , if for all  $\beta < \alpha$ ,  $I =_\beta J$  and, moreover,  $I \parallel T_\alpha \subseteq J \parallel T_\alpha$  and  $I \parallel F_\alpha \supseteq J \parallel F_\alpha$ . We write  $I \sqsubset_\alpha J$ , if  $I \sqsubseteq_\alpha J$  but  $I =_\alpha J$  does not hold.

**Definition 3.6.** Let  $I$  and  $J$  be interpretations of a given propositional program  $P$ . We write  $I \sqsubset J$ , if there exists a countable ordinal  $\alpha$  such that  $I \sqsubset_\alpha J$ . We write  $I \sqsubseteq J$  if either  $I = J$  or  $I \sqsubset J$ .

It is easy to see [20] that  $\sqsubseteq$  is a partial order,  $\sqsubseteq_\alpha$  is a preorder, and  $=_\alpha$  is an equivalence relation. As in the case of positive programs, the minimum model of a program  $P$  coincides with the least fixed-point of an operator  $T_P$ . This operator is defined through the notion of the “least upper bound” of a set of truth values.

**Definition 3.7.** Let  $S$  be a set with a partial order  $\leq$  and let  $A \subseteq S$ . We say that  $u \in S$  is an *upper bound* of  $A$ , if for every  $v \in A$  we have  $v \leq u$ . Moreover,  $u$  is called the *least upper bound* of  $A$ , if for every upper bound  $u'$  of  $A$  we have  $u \leq u'$ .

If the least upper bound of  $A$  exists then it is unique and we denote it by  $\text{lub}(A)$ . In [20] it is shown that every subset of  $\mathbb{V}$  has a least upper bound and the operator  $T_P$  can then be defined as below:

**Definition 3.8.** Let  $P$  be a propositional program and let  $I$  be an interpretation of  $P$ . The *immediate consequence operator*  $T_P$  of  $P$  is defined as follows:

$$T_P(I)(\mathbf{p}) = \text{lub}(\{I(L_1, \dots, L_n) \mid \mathbf{p} \leftarrow L_1, \dots, L_n \in P\})$$

The least fixed-point  $M_P$  of  $T_P$  is constructed as follows. We start with  $\emptyset$ , namely the interpretation that assigns to every propositional variable of  $P$  the value  $F_0$ . We iterate  $T_P$  on  $\emptyset$  until the set of variables having a  $F_0$  value and the set of variables having a  $T_0$  value, stabilize. Then we reset the values of all remaining variables to  $F_1$ . The procedure is repeated until the  $F_1$  and  $T_1$  values stabilize, and we reset the remaining variables to  $F_2$ , and so on. It is shown in [20] that there exists a countable ordinal  $\delta$  for which this process will not produce any new variables having  $F_\delta$  or  $T_\delta$  values. At this point we reset all remaining variables to 0. The following definitions formalize this process.

**Definition 3.9.** Let  $P$  be a propositional program and let  $I$  be an interpretation of  $P$ . For each countable ordinal  $\alpha$ , we define the interpretation  $T_{P,\alpha}^\omega(I)$  as follows:

$$T_{P,\alpha}^\omega(I)(p) = \begin{cases} I(p), & \text{if } \text{order}(I(p)) < \alpha \\ T_\alpha, & \text{if } p \in \bigcup_{n < \omega} (T_P^n(I) \parallel T_\alpha) \\ F_\alpha, & \text{if } p \in \bigcap_{n < \omega} (T_P^n(I) \parallel F_\alpha) \\ F_{\alpha+1}, & \text{otherwise} \end{cases}$$

**Definition 3.10.** Let  $P$  be a propositional program. For each countable ordinal  $\alpha$ , we define  $M_\alpha = T_{P,\alpha}^\omega(I_\alpha)$  where  $I_0 = \emptyset$ ,  $I_\alpha = M_{\alpha-1}$  if  $\alpha$  is a successor ordinal, and

$$I_\alpha(p) = \begin{cases} M_\beta(p), & \text{if } \text{order}(M_\beta(p)) = \beta \text{ for some } \beta < \alpha \\ F_\alpha, & \text{otherwise} \end{cases}$$

if  $\alpha$  is a limit ordinal. The  $M_0, M_1, \dots, M_\alpha, \dots$  are called the *approximations* to the minimum model of  $P$ .

In [20] it is shown that the above sequence of approximations is well-defined. We will make use of the following lemma from [20]:

**Lemma 3.11.** *Let  $P$  be a propositional program and let  $\alpha$  be a countable ordinal. For all  $n < \omega$ ,  $T_P^n(I_\alpha) \sqsubseteq_\alpha M_\alpha$ .*

The following lemma from [20] states that there exists a certain countable ordinal, after which new approximations do not introduce new truth values:

**Lemma 3.12.** *Let  $P$  be a propositional program. Then, there exists a countable ordinal  $\delta$ , called the depth of  $P$ , such that:*

- (1) *for all countable ordinals  $\gamma \geq \delta$ ,  $M_\gamma \parallel T_\gamma = \emptyset$  and  $M_\gamma \parallel F_\gamma = \emptyset$ ;*
- (2) *for all  $\beta < \delta$ ,  $M_\beta \parallel T_\beta \neq \emptyset$  or  $M_\beta \parallel F_\beta \neq \emptyset$ .*

Given a propositional program  $P$  that has depth  $\delta$ , we define the following interpretation  $M_P$ :

$$M_P(p) = \begin{cases} M_\delta(p), & \text{if } \text{order}(M_\delta(p)) < \delta \\ 0, & \text{otherwise} \end{cases}$$

The following two theorems from [20], establish interesting properties of  $M_P$ :

**Theorem 3.13.** *The infinite-valued interpretation  $M_P$  is a model of  $P$ . Moreover, it is the least (with respect to  $\sqsubseteq$ ) among all the infinite-valued models of  $P$ .*

**Theorem 3.14.** *The interpretation  $N_P$  obtained by collapsing all true values of  $M_P$  to True and all false values to False, coincides with the well-founded model of  $P$ .*

The next lemma states a fact already implied earlier, namely that new approximations do not affect the sets of variables stabilized by the preceding ones.

**Lemma 3.15.** *Let  $P$  be a propositional program and let  $\alpha$  be a countable ordinal. For all countable ordinals  $\beta > \alpha$ ,  $M_\alpha =_\alpha M_\beta$ . Moreover,  $M_\alpha =_\alpha M_P$ .*

4. THE SYNTAX OF  $\mathcal{H}$ 

In this section we define the syntax of the language  $\mathcal{H}$  that we use throughout the paper.  $\mathcal{H}$  is based on a simple type system with two base types:  $o$ , the boolean domain, and  $\iota$ , the domain of data objects. The composite types are partitioned into three classes: functional (assigned to function symbols), predicate (assigned to predicate symbols) and argument (assigned to parameters of predicates).

**Definition 4.1.** A type can either be *functional*, *predicate*, or *argument*, denoted by  $\sigma$ ,  $\pi$  and  $\rho$  respectively and defined as:

$$\begin{aligned}\sigma &:= \iota \mid (\iota \rightarrow \sigma) \\ \pi &:= o \mid (\rho \rightarrow \pi) \\ \rho &:= \iota \mid \pi\end{aligned}$$

We will use  $\tau$  to denote an arbitrary type (either functional, predicate, or argument). As usual, the binary operator  $\rightarrow$  is right-associative. A functional type that is different than  $\iota$  will often be written in the form  $\iota^n \rightarrow \iota$ ,  $n \geq 1$ . Moreover, it can be easily seen that every predicate type  $\pi$  can be written in the form  $\rho_1 \rightarrow \dots \rightarrow \rho_n \rightarrow o$ ,  $n \geq 0$  (for  $n = 0$  we assume that  $\pi = o$ ). We proceed by defining the syntax of  $\mathcal{H}$ :

**Definition 4.2.** The *alphabet* of  $\mathcal{H}$  consists of the following:

- (1) Predicate variables of every predicate type  $\pi$  (denoted by capital letters such as  $Q, R, S, \dots$ ).
- (2) Individual variables of type  $\iota$  (denoted by capital letters such as  $X, Y, Z, \dots$ ).
- (3) Predicate constants of every predicate type  $\pi$  (denoted by lowercase letters such as  $p, q, r, \dots$ ).
- (4) Individual constants of type  $\iota$  (denoted by lowercase letters such as  $a, b, c, \dots$ ).
- (5) Function symbols of every functional type  $\sigma \neq \iota$  (denoted by lowercase letters such as  $f, g, h, \dots$ ).
- (6) The inverse implication constant  $\leftarrow$ , the negation constant  $\sim$ , the comma, the left and right parentheses, and the equality constant  $\approx$  for comparing terms of type  $\iota$ .

Arbitrary variables will be usually denoted by  $V$  and its subscripted versions.

**Definition 4.3.** The set of *terms* of  $\mathcal{H}$  is defined as follows:

- Every predicate variable (respectively, predicate constant) of type  $\pi$  is a term of type  $\pi$ ; every individual variable (respectively, individual constant) of type  $\iota$  is a term of type  $\iota$ ;
- if  $f$  is an  $n$ -ary function symbol and  $E_1, \dots, E_n$  are terms of type  $\iota$  then  $(f E_1 \dots E_n)$  is a term of type  $\iota$ ;
- if  $E_1$  is a term of type  $\rho \rightarrow \pi$  and  $E_2$  a term of type  $\rho$  then  $(E_1 E_2)$  is a term of type  $\pi$ .

**Definition 4.4.** The set of *expressions* of  $\mathcal{H}$  is defined as follows:

- A term of type  $\rho$  is an expression of type  $\rho$ ;
- if  $E$  is a term of type  $o$  then  $(\sim E)$  is an expression of type  $o$ ;
- if  $E_1$  and  $E_2$  are terms of type  $\iota$ , then  $(E_1 \approx E_2)$  is an expression of type  $o$ .

We write  $vars(E)$  to denote the set of all the variables in  $E$ . Expressions (respectively, terms) that have no variables will often be referred to as *ground expressions* (respectively, *ground terms*). We will omit parentheses when no confusion arises. To denote that an expression  $E$  has type  $\rho$  we will often write  $E : \rho$ . Terms of type  $o$  will often be referred to



as *atoms*. Expressions of type  $o$  that do not contain negation, i.e. atoms and expressions of the form  $(E_1 \approx E_2)$ , will be called *positive literals*, while expressions of the form  $(\sim E)$  will be called *negative literals*. A *literal* is either a positive literal or a negative literal.

**Definition 4.5.** A *clause* of  $\mathcal{H}$  is a formula  $p V_1 \cdots V_n \leftarrow L_1, \dots, L_m$ , where  $p$  is a predicate constant of type  $\rho_1 \rightarrow \cdots \rightarrow \rho_n \rightarrow o$ ,  $V_1, \dots, V_n$  are distinct variables of types  $\rho_1, \dots, \rho_n$  respectively and  $L_1, \dots, L_m$  are literals. The term  $p V_1 \cdots V_n$  is called the *head* of the clause, the variables  $V_1, \dots, V_n$  are the *formal parameters* of the clause and the conjunction  $L_1, \dots, L_m$  is its *body*. A *program*  $P$  of  $\mathcal{H}$  is a finite set of clauses.

**Example 4.6.** The program below defines the **subset** relation over unary predicates:

$$\begin{aligned} \text{subset } S1 \ S2 & :- \sim(\text{nsubset } S1 \ S2). \\ \text{nsubset } S1 \ S2 & :- S1 \ X, \sim(S2 \ X). \end{aligned}$$

The variables **S1** and **S2** are both predicate variables of type  $\iota \rightarrow o$ , while **X** is an individual variable (i.e., it is of type  $\iota$ ). Given predicates  $p$  and  $q$  of type  $\iota \rightarrow o$ , **subset**  $p \ q$  is true if  $p$  is a subset of  $q$ .  $\square$

In the following, we will often talk about the “ground instantiation of a program”. This notion is formally defined below.

**Definition 4.7.** A *substitution*  $\theta$  is a finite set of the form  $\{V_1/E_1, \dots, V_n/E_n\}$  where the  $V_i$ 's are different variables and each  $E_i$  is a term having the same type as  $V_i$ . We write  $dom(\theta)$  to denote the domain  $\{V_1, \dots, V_n\}$  of  $\theta$ . If all the expressions  $E_1, \dots, E_n$  are ground terms,  $\theta$  is called a *ground substitution*.

We can now define the application of a substitution to an expression.

**Definition 4.8.** Let  $\theta$  be a substitution and  $E$  be an expression. Then,  $E\theta$  is an expression obtained from  $E$  as follows:

- $E\theta = E$  if  $E$  is a predicate constant or individual constant;
- $V\theta = \theta(V)$  if  $V \in dom(\theta)$ ; otherwise,  $V\theta = V$ ;
- $(f E_1 \cdots E_n)\theta = (f E_1\theta \cdots E_n\theta)$ ;
- $(E_1 E_2)\theta = (E_1\theta E_2\theta)$ ;
- $(\sim E)\theta = (\sim E\theta)$ ;
- $(E_1 \approx E_2)\theta = (E_1\theta \approx E_2\theta)$ .

If  $\theta$  is a ground substitution such that  $vars(E) \subseteq dom(\theta)$ , then the ground expression  $E\theta$  is called a *ground instance* of  $E$ .

**Definition 4.9.** Let  $P$  be a program. A *ground instance of a clause*  $p V_1 \cdots V_n \leftarrow L_1, \dots, L_m$  of  $P$  is a formula  $(p V_1 \cdots V_n)\theta \leftarrow L_1\theta, \dots, L_m\theta$ , where  $\theta$  is a ground substitution whose domain is the set of all variables that appear in the clause, such that for every  $V \in dom(\theta)$  with  $V : \rho$ ,  $\theta(V)$  is a ground expression of type  $\rho$  that has been formed with predicate constants, function symbols, and individual constants that appear in  $P$ . The *ground instantiation of a program*  $P$ , denoted by  $Gr(P)$ , is the (possibly infinite) set that contains all the ground instances of the clauses of  $P$ .

In the rest of the paper, we will be extensively using the notion of ground instantiation. Notice that in the body of a clause of the ground instantiation, there may exist ground expressions of the form  $(E_1 \approx E_2)$ . In the case where the two expressions  $E_1$  and  $E_2$  are syntactically identical, the expression  $(E_1 \approx E_2)$  will be treated as the constant **true**, and otherwise as the constant **false**.

## 5. THE SEMANTICS OF $\mathcal{H}$

In this section we develop the semantics of  $\mathcal{H}$ . Our developments generalize the semantics of [2, 3] for positive higher-order logic programs, to programs with negation. Notice that the semantics of [2, 3] is based on classical two-valued logic, while ours on the infinite-valued logic of Section 3.

In order to interpret the programs of  $\mathcal{H}$ , we need to specify the semantic domains in which the expressions of each type  $\tau$  are assigned their meanings. We adopt the approach of [2, 3]. More specifically, the following definition implies that the expressions of predicate types should be understood as representing functions. We use  $[S_1 \rightarrow S_2]$  to denote the set of (possibly partial) functions from a set  $S_1$  to a set  $S_2$ . The possibility to have a partial function arises due to a technicality which is explained in the remark just above Definition 5.3.

**Definition 5.1.** A *functional type structure*  $\mathcal{S}$  for  $\mathcal{H}$  consists of two non-empty sets  $D$  and  $A$  together with an assignment  $\llbracket \tau \rrbracket$  to each type  $\tau$  of  $\mathcal{H}$ , so that the following are satisfied:

- $\llbracket \iota \rrbracket = D$ ;
- $\llbracket \iota^n \rightarrow \iota \rrbracket = D^n \rightarrow D$ ;
- $\llbracket o \rrbracket = A$ ;
- $\llbracket \rho \rightarrow \pi \rrbracket \subseteq \llbracket [\rho] \rrbracket \rightarrow \llbracket [\pi] \rrbracket$ .

Given a functional type structure  $\mathcal{S}$ , any function  $val : \llbracket o \rrbracket \rightarrow \mathbb{V}$  will be called an *infinite-valued valuation function* (or simply *valuation function*) for  $\mathcal{S}$ .

It is customary in the study of the semantics of logic programming languages to restrict attention to *Herbrand interpretations*. Given a program  $P$ , a Herbrand interpretation is one that has as its underlying universe the so-called *Herbrand universe* of  $P$ :

**Definition 5.2.** For a program  $P$ , we define the *Herbrand universe* for every argument type  $\rho$ , denoted by  $U_{P,\rho}$ , to be the set of all ground terms of type  $\rho$  that can be formed out of the individual constants, function symbols, and predicate constants in the program. Moreover, we define  $U_{P,o}^+$  to be the set of all ground expressions of type  $o$ , that can be formed out of the above symbols, i.e. the set  $U_{P,o}^+ = U_{P,o} \cup \{(E_1 \approx E_2) \mid E_1, E_2 \in U_{P,\iota}\} \cup \{(\sim E) \mid E \in U_{P,o}\}$ .

Following [2, 3], we take  $D$  and  $A$  in Definition 5.1 to be equal to  $U_{P,\iota}$  and  $U_{P,o}^+$  respectively. Then, each element of  $U_{P,\rho \rightarrow \pi}$  can itself be perceived as a function mapping elements of  $\llbracket [\rho] \rrbracket$  to elements of  $\llbracket [\pi] \rrbracket$ , through syntactic application mapping. That is,  $E \in U_{P,\rho \rightarrow \pi}$  can be viewed as the function mapping each term  $E' \in U_{P,\rho}$  to the term  $EE' \in U_{P,\pi}$ . Similarly, every  $n$ -ary function symbol  $f$  appearing in  $P$  can be viewed as the function mapping each element  $(E_1, \dots, E_n) \in U_{P,\iota}^n$  to the term  $(f E_1 \dots E_n) \in U_{P,\iota}$ .

**Remark:** There is a small technicality here which we need to clarify. In the case where  $\rho = o$ ,  $E \in U_{P,o \rightarrow \pi}$  is a partial function because it maps elements of  $U_{P,o}$  (and not of  $U_{P,o}^+$ ) to elements of  $U_{P,\pi}$ ; this is due to the fact that our syntax does not allow an expression of type  $o \rightarrow \pi$  to take as argument an expression of the form  $(E_1 \approx E_2)$  nor of the form  $(\sim E)$ . In all other cases (i.e., when  $\rho \neq o$ ),  $E$  represents a total function.

**Definition 5.3.** A *Herbrand interpretation*  $I$  of a program  $P$  consists of

- (1) a functional type structure  $\mathcal{S}_I$ , such that  $D = U_{P,\iota}$ ,  $A = U_{P,o}^+$  and  $\llbracket [\rho \rightarrow \pi] \rrbracket = U_{P,\rho \rightarrow \pi}$  for every predicate type  $\rho \rightarrow \pi$ ;

- (2) an assignment to each individual constant  $c$  in  $P$ , of the element  $I(c) = c$ ; to each predicate constant  $p$  in  $P$ , of the element  $I(p) = p$ ; to each function symbol  $f$  in  $P$ , of the element  $I(f) = f$ ;
- (3) a valuation function  $val_I(\cdot)$  for  $\mathcal{S}_I$ , assigning to each element of  $U_{P,o}^+$  an element in  $\mathbb{V}$ , while satisfying the following:
- for all  $E_1, E_2 \in U_{P,\iota}$ ,  $val_I((E_1 \approx E_2)) = \begin{cases} F_0, & \text{if } E_1 \neq E_2; \\ T_0, & \text{if } E_1 = E_2; \end{cases}$
  - for all  $E \in U_{P,o}$ ,  $val_I((\sim E)) = \begin{cases} T_{\alpha+1}, & \text{if } val_I(E) = F_\alpha \\ F_{\alpha+1}, & \text{if } val_I(E) = T_\alpha. \\ 0, & \text{if } val_I(E) = 0 \end{cases}$

We call  $val_I(\cdot)$  the *valuation function of  $I$*  and omit the reference to  $\mathcal{S}_I$ , since the latter is common to all Herbrand interpretations of a program. In fact, individual Herbrand interpretations are only set apart by their valuation functions.

**Definition 5.4.** A *Herbrand state* (or simply *state*)  $s$  of a program  $P$  is a function that assigns to each variable  $V$  of type  $\rho$  an element of  $U_{P,\rho}$ .

Given a Herbrand interpretation  $I$  and state  $s$ , we can define the semantics of expressions with respect to  $I$  and  $s$ .

**Definition 5.5.** Let  $P$  be a program. Also, let  $I$  be a Herbrand interpretation and  $s$  a Herbrand state of  $P$ . Then the semantics of expressions with respect to  $I$  and  $s$  is defined as follows:

- $\llbracket c \rrbracket_s(I) = I(c) = c$ , for every individual constant  $c$ ;
- $\llbracket p \rrbracket_s(I) = I(p) = p$ , for every predicate constant  $p$ ;
- $\llbracket V \rrbracket_s(I) = s(V)$ , for every variable  $V$ ;
- $\llbracket (f E_1 \cdots E_n) \rrbracket_s(I) = (I(f) \llbracket E_1 \rrbracket_s(I) \cdots \llbracket E_n \rrbracket_s(I)) = (f \llbracket E_1 \rrbracket_s(I) \cdots \llbracket E_n \rrbracket_s(I))$ , for every  $n$ -ary function symbol  $f$ ;
- $\llbracket (E_1 E_2) \rrbracket_s(I) = (\llbracket E_1 \rrbracket_s(I) \llbracket E_2 \rrbracket_s(I))$ ;
- $\llbracket (E_1 \approx E_2) \rrbracket_s(I) = (\llbracket E_1 \rrbracket_s(I) \approx \llbracket E_2 \rrbracket_s(I))$ ;
- $\llbracket (\sim E) \rrbracket_s(I) = (\sim \llbracket E \rrbracket_s(I))$ .

Since we are dealing with Herbrand interpretations, it is easy to see that for every Herbrand state  $s$  and ground expression  $E$ , we have  $\llbracket E \rrbracket_s(I) = E$ . Therefore, if  $E$  is a ground literal, we can write  $val_I(E)$  instead of  $val_I(\llbracket E \rrbracket_s(I))$ . Stretching this abuse of notation a little further, we can extend a valuation function to assign truth values to ground conjunctions of literals:

**Definition 5.6.** Let  $P$  be a program and  $I$  be a Herbrand interpretation of  $P$ . We define  $val_I(L_1, \dots, L_n) = \min\{val_I(L_1), \dots, val_I(L_n)\}$  for all  $L_1, \dots, L_n \in U_{P,o}^+$ .

Based on the above definition, we can define the concept of Herbrand models for our higher-order programs in the same way as in classical logic programming.

**Definition 5.7.** Let  $P$  be a program and  $I$  be a Herbrand interpretation of  $P$ . We say  $I$  is a *model* of  $P$  if  $val_I(\llbracket A \rrbracket_s(I)) \geq val_I(\llbracket L_1 \rrbracket_s(I), \dots, \llbracket L_m \rrbracket_s(I))$  holds for every clause  $A \leftarrow L_1, \dots, L_m$  and every Herbrand state  $s$  of  $P$ .

Bezem's semantics is based on the observation that, given a positive higher-order program, we can use the minimum model semantics of its ground instantiation as a (two-valued)

valuation function defining a Herbrand interpretation for the initial program itself. We use the same idea for  $\mathcal{H}$  programs; the only difference is that we employ the infinite-valued model of the ground instantiation of the program as the valuation function.

**Definition 5.8.** Let  $P$  be a program. Also, let  $\text{Gr}(P)$  be the ground instantiation of  $P$  and let  $M_{\text{Gr}(P)}$  be the infinite-valued model of  $\text{Gr}(P)$ . We define  $\mathcal{M}_P$  to be the Herbrand interpretation of  $P$  such that  $\text{val}_{\mathcal{M}_P}(A) = M_{\text{Gr}(P)}(A)$  for every  $A \in U_{P,o}$ .

We adopt the notation  $I \parallel v$  from Section 3, to signify the set of atoms which are assigned a certain truth value  $v \in \mathbb{V}$  by a Herbrand interpretation  $I$ ; that is,  $I \parallel v = \{A \mid A \in U_{P,o} \text{ and } \text{val}_I(A) = v\}$ . Then the relations  $\sqsubseteq_\alpha$ ,  $\sqsubset_\alpha$ ,  $=_\alpha$ ,  $\sqsubseteq$  and  $\sqsubset$  on Herbrand interpretations of a higher-order program can be defined in exactly the same manner as in Section 3.

The next theorem verifies that our semantics is well-defined, in the sense that the interpretation  $\mathcal{M}_P$ , which we chose as the meaning of a program  $P$ , is indeed a model of  $P$ . In fact it is the minimum, with respect to  $\sqsubseteq$ , model of  $P$ .

**Theorem 5.9.**  $\mathcal{M}_P$  is the minimum (with respect to  $\sqsubseteq$ ) Herbrand model of  $P$ .

*Proof.* Let  $\text{Gr}(P)$  be the ground instantiation of  $P$  and  $M_{\text{Gr}(P)}$  be the infinite-valued model of  $\text{Gr}(P)$ . Recall (Definition 5.4) that a Herbrand state  $s$  of  $P$  assigns to each variable  $V$  of type  $\rho$  an element of  $U_{P,\rho}$  and that this (semantic) element is in fact a ground expression of the same type. Moreover, recall (Definition 4.7) that a ground substitution, which includes  $V$  in its domain, also assigns a ground expression of type  $\rho$  to  $V$ ; and that such an expression coincides with its own meaning, under any Herbrand interpretation and any state. Therefore, for every Herbrand state  $s$  of  $P$  there exists a ground substitution  $\theta$  such that  $s(V) = \llbracket \theta(V) \rrbracket_{s'}(\mathcal{M}_P)$  for all states  $s'$  and variables  $V$  in  $P$ . Also, for every clause  $A \leftarrow L_1, \dots, L_m$  in  $P$  there exists a respective ground instance  $A\theta \leftarrow L_1\theta, \dots, L_m\theta$  in  $\text{Gr}(P)$ . As  $M_{\text{Gr}(P)}$  is a model of  $\text{Gr}(P)$ ,  $M_{\text{Gr}(P)}(A\theta) \geq \min\{M_{\text{Gr}(P)}(L_1\theta), \dots, M_{\text{Gr}(P)}(L_m\theta)\}$ . By definition,  $\text{val}_{\mathcal{M}_P}(A\theta) = M_{\text{Gr}(P)}(A\theta)$  and  $\text{val}_{\mathcal{M}_P}(L_i\theta) = M_{\text{Gr}(P)}(L_i\theta)$  for all  $i \leq m$ . Moreover, it is easy to see (by a trivial induction on the structure of the expression) that  $A\theta = \llbracket A \rrbracket_s(\mathcal{M}_P)$ , which implies that  $\text{val}_{\mathcal{M}_P}(A\theta) = \text{val}_{\mathcal{M}_P}(\llbracket A \rrbracket_s(\mathcal{M}_P))$ . Similarly,  $\text{val}_{\mathcal{M}_P}(L_i\theta) = \text{val}_{\mathcal{M}_P}(\llbracket L_i \rrbracket_s(\mathcal{M}_P))$ , for all  $i \leq m$ . Then  $\text{val}_{\mathcal{M}_P}(\llbracket A \rrbracket_s(\mathcal{M}_P)) \geq \min\{\text{val}_{\mathcal{M}_P}(\llbracket L_1 \rrbracket_s(\mathcal{M}_P)), \dots, \text{val}_{\mathcal{M}_P}(\llbracket L_m \rrbracket_s(\mathcal{M}_P))\}$  follows immediately and implies that  $\mathcal{M}_P$  is a model of  $P$ . To see that it is minimum, assume there exists a model  $\mathcal{M}$  of the higher-order program  $P$ , distinct from  $\mathcal{M}_P$ , which does not satisfy  $\mathcal{M}_P \sqsubset \mathcal{M}$ . Then the valuation function  $\text{val}_{\mathcal{M}}(\cdot)$  of  $\mathcal{M}$  defines a (first-order) interpretation  $M$  for the ground instantiation  $\text{Gr}(P)$  of  $P$ , if, for every ground atom  $A$ , we take  $M(A) = \text{val}_{\mathcal{M}}(A)$ . It is obvious that  $M_{\text{Gr}(P)} \not\sqsubseteq M$ , since  $M(A) = \text{val}_{\mathcal{M}}(A)$  and  $\text{val}_{\mathcal{M}_P}(A) = M_{\text{Gr}(P)}(A)$  for every ground atom  $A$  imply that  $M \parallel v = \mathcal{M} \parallel v$  and  $M_{\text{Gr}(P)} \parallel v = \mathcal{M}_P \parallel v$  for every truth value  $v \in \mathbb{V}$ . Also,  $M$  is distinct from  $M_{\text{Gr}(P)}$ , since  $M(B) = \text{val}_{\mathcal{M}}(B) \neq \text{val}_{\mathcal{M}_P}(B) = M_{\text{Gr}(P)}(B)$  for at least one ground atom  $B$ . Next we demonstrate that  $M$  is a model of the ground instantiation  $\text{Gr}(P)$  of  $P$ , which is of course a contradiction, since  $M_{\text{Gr}(P)}$  is the minimum model of  $\text{Gr}(P)$  and  $M_{\text{Gr}(P)} \sqsubseteq M$  should hold. Indeed, every clause in  $\text{Gr}(P)$  is a ground instance of a clause  $A \leftarrow L_1, \dots, L_m$  in  $P$  and is therefore of the form  $A\theta \leftarrow L_1\theta, \dots, L_m\theta$  for some ground substitution  $\theta$ . Consider a Herbrand state  $s$ , such that  $s(V) = \theta(V)$  for every variable  $V$  in  $P$ . Because we assumed the higher-order interpretation  $\mathcal{M}$  to be a model of  $P$ , we have that  $\text{val}_{\mathcal{M}}(\llbracket A \rrbracket_s(\mathcal{M})) \geq \min\{\text{val}_{\mathcal{M}}(\llbracket L_1 \rrbracket_s(\mathcal{M})), \dots, \text{val}_{\mathcal{M}}(\llbracket L_m \rrbracket_s(\mathcal{M}))\}$ . Again, it is easy to see that  $A\theta = \llbracket A \rrbracket_s(\mathcal{M})$  and therefore  $\text{val}_{\mathcal{M}}(A\theta) = \text{val}_{\mathcal{M}}(\llbracket A \rrbracket_s(\mathcal{M}))$ . Similarly,  $\text{val}_{\mathcal{M}}(L_i\theta) = \text{val}_{\mathcal{M}}(\llbracket L_i \rrbracket_s(\mathcal{M}))$  for all

$i \leq m$ . Additionally, by the construction of  $M$ ,  $val_{\mathcal{M}}(A\theta) = M(A\theta)$  and  $val_{\mathcal{M}}(L_i\theta) = M(L_i\theta)$  for all  $i \leq m$ , so  $M(A\theta) \geq \min\{M(L_1\theta), \dots, M(L_m\theta)\}$ , which implies that  $M$  is a model of  $\text{Gr}(\text{P})$ . Because we reached a contradiction,  $\mathcal{M}_{\text{P}}$  must be the minimum model of  $\text{P}$ .  $\square$

## 6. EXTENSIONALITY OF THE PROPOSED SEMANTICS

In this section we show that the infinite-valued model we defined in the previous section enjoys the extensionality property, as this was defined in [2]. For an intuitive explanation of this property, the reader should consult again the example of Section 2. In order to formally define this notion, Bezem introduced [2, 3] relations  $\cong_{val, \tau}$  over the set of expressions of a given type  $\tau$  and under a given valuation function  $val$ . These relations intuitively express extensional equality of type  $\tau$ . For the purposes of this paper only extensional equality of argument types will be needed, for which the formal definition is as follows:

**Definition 6.1.** Let  $\mathcal{S}$  be a functional type structure and  $val$  be a valuation function for  $\mathcal{S}$ . For every argument type  $\rho$  we define the relations  $\cong_{val, \rho}$  on  $\llbracket \rho \rrbracket$  as follows: Let  $d, d' \in \llbracket \rho \rrbracket$ ; then  $d \cong_{val, \rho} d'$  if and only if

- (1)  $\rho = \iota$  and  $d = d'$ , or
- (2)  $\rho = o$  and  $val(d) = val(d')$ , or
- (3)  $\rho = \rho' \rightarrow \pi$  and  $d e \cong_{val, \pi} d' e'$  for all  $e, e' \in \llbracket \rho' \rrbracket$ , such that  $e \cong_{val, \rho'} e'$  and  $d e, d' e'$  are both defined.

One can easily verify that, for all  $d, d' \in \llbracket \rho_1 \rightarrow \dots \rightarrow \rho_n \rightarrow o \rrbracket$ ,  $e_1, e'_1 \in \llbracket \rho_1 \rrbracket, \dots, e_n, e'_n \in \llbracket \rho_n \rrbracket$ , if  $d \cong_{val, \rho_1 \rightarrow \dots \rightarrow \rho_n \rightarrow o} d'$ ,  $e_1 \cong_{val, \rho_1} e'_1, \dots, e_n \cong_{val, \rho_n} e'_n$  and  $d e_1 \dots e_n, d' e'_1 \dots e'_n$  are both defined, then  $val(d e_1 \dots e_n) = val(d' e'_1 \dots e'_n)$ .

Generally, it is not guaranteed that such relations will be equivalence relations; rather they are partial equivalences (they are shown in [2] to be symmetric and transitive). However, we are going to see that the minimum model of a program defines true equivalence relations for all types  $\tau$ .

**Definition 6.2.** Let  $\text{P}$  be a program and let  $I$  be a Herbrand interpretation of  $\text{P}$ . We say  $I$  is *extensional* if for all argument types  $\rho$  the relations  $\cong_{val_I, \rho}$  are reflexive, i.e. for all  $E \in \llbracket \rho \rrbracket$ , it holds that  $E \cong_{val_I, \rho} E$ .

**Theorem 6.3** (Extensionality).  $\mathcal{M}_{\text{P}}$  is extensional.

*Proof.* Since the valuation function of  $\mathcal{M}_{\text{P}}$  is  $M_{\text{Gr}(\text{P})}$ , essentially we need to show that  $E \cong_{M_{\text{Gr}(\text{P})}, \rho} E$ , for every ground expression  $E$  of every argument type  $\rho$ . We perform an induction on the structure of  $\rho$ . For the base types  $\iota$  and  $o$  the statement holds by definition. For the induction step, we prove the statement for a predicate type  $\pi = \rho_1 \rightarrow \dots \rightarrow \rho_m \rightarrow o$ , assuming that it holds for all types simpler than  $\pi$  (i.e., for the types  $\rho_1, \dots, \rho_m, o$  and, recursively, the types that are simpler than  $\rho_1, \dots, \rho_m$ ). Let  $A$  be any atom of the following form:  $A$  is headed by a predicate constant and all variables in  $vars(A)$  are of types simpler than  $\pi$ . Let  $\theta, \theta'$  be ground substitutions, such that  $vars(A) \subseteq dom(\theta), dom(\theta')$  and  $\theta(V) \cong_{M_{\text{Gr}(\text{P})}, \rho} \theta'(V)$  for any  $V : \rho$  in  $vars(A)$ . We claim it suffices to show the following two properties  $P_1(\alpha)$  and  $P_2(\alpha)$ , for all ordinals  $\alpha$ :

- $P_1(\alpha)$ : if  $M_\alpha(A\theta) = T_\alpha$  then  $M_{\text{Gr}(\text{P})}(A\theta') = T_\alpha$ ;
- $P_2(\alpha)$ : if  $M_\alpha(A\theta) = F_\alpha$  then  $M_{\text{Gr}(\text{P})}(A\theta') = F_\alpha$ .

To see why proving the above properties is enough to establish that  $\mathbf{E} \cong_{M_{\text{Gr}(\mathbf{P}),\pi}} \mathbf{E}$ , observe the following: first of all, we assumed that  $\pi$  is of the form  $\rho_1 \rightarrow \dots \rightarrow \rho_m \rightarrow o$ , so if  $\mathbf{V}_1 : \rho_1, \dots, \mathbf{V}_m : \rho_m$  are variables, then  $\mathbf{E} \mathbf{V}_1 \dots \mathbf{V}_m$  is an atom of the form described above. Also, by Lemma 3.15 we have that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{E} \theta(\mathbf{V}_1) \dots \theta(\mathbf{V}_m)) = T_\alpha$  if and only if  $M_\alpha(\mathbf{E} \theta(\mathbf{V}_1) \dots \theta(\mathbf{V}_m)) = T_\alpha$ . If  $P_1(\alpha)$  holds, the latter implies that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{E} \theta'(\mathbf{V}_1) \dots \theta'(\mathbf{V}_m)) = T_\alpha$ . Because the relations  $\cong_{M_{\text{Gr}(\mathbf{P}),\rho_i}}$  are symmetric,  $\theta$  and  $\theta'$  are interchangeable. Therefore the same argument can be used to infer the reverse implication, i.e.  $M_{\text{Gr}(\mathbf{P})}(\mathbf{E} \theta'(\mathbf{V}_1) \dots \theta'(\mathbf{V}_m)) = T_\alpha \Rightarrow M_{\text{Gr}(\mathbf{P})}(\mathbf{E} \theta(\mathbf{V}_1) \dots \theta(\mathbf{V}_m)) = T_\alpha$ , thus resulting to an equivalence. If  $P_2(\alpha)$  holds, the analogous equivalence can be shown for the value  $F_\alpha$ , in the same way. Finally, the equivalence for the 0 value follows by a simple elimination argument: if for example  $M_{\text{Gr}(\mathbf{P})}(\mathbf{E} \theta(\mathbf{V}_1) \dots \theta(\mathbf{V}_m)) = 0$ , we make the assumption that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{E} \theta'(\mathbf{V}_1) \dots \theta'(\mathbf{V}_m)) = T_\alpha$  (respectively,  $F_\alpha$ ) for some ordinal  $\alpha$ . Then, by Lemma 3.15,  $M_\alpha(\mathbf{E} \theta'(\mathbf{V}_1) \dots \theta'(\mathbf{V}_m)) = T_\alpha$  (respectively,  $F_\alpha$ ), so if property  $P_1(\alpha)$  (respectively,  $P_2(\alpha)$ ) holds, it gives us that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{E} \theta(\mathbf{V}_1) \dots \theta(\mathbf{V}_m)) = T_\alpha$  (respectively,  $F_\alpha$ ), which is a contradiction. It follows that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{E} \theta'(\mathbf{V}_1) \dots \theta'(\mathbf{V}_m)) = 0$ . Again, we can show the reverse implication by the same argument.

We will proceed by a second induction on  $\alpha$ .

**Second Induction Basis ( $\alpha = 0$ ):** We have  $M_0 = T_{\mathbf{P},0}^\omega(\emptyset)$ . Observe that  $T_{\mathbf{P},0}^\omega(\emptyset)(\mathbf{A}\theta)$  will evaluate to  $T_0$  if and only if there exists some  $n < \omega$  for which  $T_{\mathbf{P}}^n(\emptyset)(\mathbf{A}\theta) = T_0$ . On the other hand, it will evaluate to  $F_0$  if and only if there does not exist a  $n < \omega$  for which  $T_{\mathbf{P}}^n(\emptyset)(\mathbf{A}\theta) \neq F_0$ . Therefore, in order to prove  $P_1(0)$  and  $P_2(0)$ , we first need to perform a third induction on  $n$  and prove the following two properties:

- $P'_1(0, n)$ : if  $T_{\mathbf{P}}^n(\emptyset)(\mathbf{A}\theta) = T_0$  then  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta') = T_0$ ;
- $P'_2(0, n)$ : if  $T_{\mathbf{P}}^n(\emptyset)(\mathbf{A}\theta) > F_0$  then  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta') > F_0$ .

**Third Induction Basis ( $n = 0$ ):** Both  $P'_1(0, 0)$  and  $P'_2(0, 0)$  hold vacuously, since  $T_{\mathbf{P}}^0(\emptyset) = \emptyset$ , i.e. the interpretation that assigns  $F_0$  to every atom.

**Third Induction Step ( $n + 1$ ):** First we show  $P'_1(0, n + 1)$ , assuming that  $P'_1(0, n)$  holds.

If  $T_{\mathbf{P}}^{n+1}(\emptyset)(\mathbf{A}\theta) = T_0$ , then there exists a clause  $\mathbf{A}\theta \leftarrow \mathbf{L}_1, \dots, \mathbf{L}_k$  in  $\text{Gr}(\mathbf{P})$  such that for each  $i \leq k$ ,  $T_{\mathbf{P}}^n(\emptyset)(\mathbf{L}_i) = T_0$ . This implies that each  $\mathbf{L}_i$  is a positive literal, since a negative one cannot be assigned the value  $T_0$  in any interpretation. This clause is a ground instance of a clause  $\mathbf{p}\mathbf{V}_1 \dots \mathbf{V}_r \leftarrow \mathbf{K}_1, \dots, \mathbf{K}_k$  in the higher-order program and there exists a substitution  $\theta''$ , such that  $(\mathbf{p}\mathbf{V}_1 \dots \mathbf{V}_r)\theta'' = \mathbf{A}$  and, for any variable  $\mathbf{V} \notin \{\mathbf{V}_1, \dots, \mathbf{V}_r\}$  appearing in the body of the clause,  $\theta''(\mathbf{V})$  is an appropriate ground term, so that  $\mathbf{L}_i = \mathbf{K}_i\theta''\theta$  for all  $i \leq k$ . Observe that the variables appearing in the clause  $(\mathbf{p}\mathbf{V}_1 \dots \mathbf{V}_r)\theta'' \leftarrow \mathbf{K}_1\theta'', \dots, \mathbf{K}_k\theta''$  are exactly the variables appearing in  $\mathbf{A}$  and they are all of types simpler than  $\pi$ . We distinguish the following cases for each  $\mathbf{K}_i\theta''$ ,  $i \leq k$ :

- (1)  $\mathbf{K}_i\theta''$  is of the form  $(\mathbf{E}_1 \approx \mathbf{E}_2)$ : By definition,  $T_{\mathbf{P}}^n(\emptyset)(\mathbf{L}_i) = T_{\mathbf{P}}^n(\emptyset)(\mathbf{K}_i\theta''\theta) = T_0$  implies that  $\mathbf{E}_1\theta = \mathbf{E}_2\theta$ . Since  $\mathbf{E}_1$  and  $\mathbf{E}_2$  are expressions of type  $\iota$ , all variables in  $\mathbf{E}_1$  and  $\mathbf{E}_2$  are also of type  $\iota$  and, because  $\cong_{M_{\text{Gr}(\mathbf{P}),\iota}}$  is defined as equality, we will have  $\mathbf{E}_1\theta = \mathbf{E}_1\theta'$  and  $\mathbf{E}_2\theta = \mathbf{E}_2\theta'$ . Therefore  $\mathbf{E}_1\theta' = \mathbf{E}_2\theta'$  and  $M_{\text{Gr}(\mathbf{P})}(\mathbf{K}_i\theta''\theta') = T_0$  will also hold.
- (2)  $\mathbf{K}_i\theta''$  is an atom and starts with a predicate constant: As we observed, the variables appearing in  $\mathbf{K}_i\theta''$  are of types simpler than  $\pi$ . By the third induction hypothesis,  $\mathbf{K}_i\theta''$  satisfies property  $P'_1(0, n)$  and therefore  $T_{\mathbf{P}}^n(\emptyset)(\mathbf{L}_i) = T_{\mathbf{P}}^n(\emptyset)(\mathbf{K}_i\theta''\theta) = T_0$  implies that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{K}_i\theta''\theta') = T_0$ .
- (3)  $\mathbf{K}_i\theta''$  is an atom and starts with a predicate variable: Let  $\mathbf{K}_i\theta'' = \mathbf{V}\mathbf{E}_1 \dots \mathbf{E}_{m'}$  for some  $\mathbf{V} \in \text{vars}(\mathbf{A})$ . Then  $\mathbf{B} = \theta(\mathbf{V})\mathbf{E}_1 \dots \mathbf{E}_{m'}$  is an atom that begins with a predicate

constant and, by  $\text{vars}(K_i\theta'') \subseteq \text{vars}(A)$ , all of the variables of  $B$  are of types simpler than  $\pi$ . Also,  $T_{\mathbf{P}}^n(\emptyset)(B\theta) = T_{\mathbf{P}}^n(\emptyset)(K_i\theta''\theta) = T_0$ , which, by property  $P'_1(0, n)$  yields that  $M_{\text{Gr}(\mathbf{P})}(B\theta') = M_{\text{Gr}(\mathbf{P})}(\theta(\mathbf{V}) E_1\theta' \cdots E_{m'}\theta') = T_0$  (1). Observe that the types of all arguments of  $\theta(\mathbf{V})$ , i.e. the types of  $E_j\theta'$  for all  $j \leq m'$ , are simpler than the type of  $\mathbf{V}$  and consequently, since  $\mathbf{V} \in \text{vars}(A)$ , simpler than  $\pi$ . For each  $j \leq m'$ , let  $\rho_j$  be the type of  $E_j$  and let  $\rho$  be the type of  $\mathbf{V}$ ; by the first induction hypothesis,  $E_j\theta' \cong_{M_{\text{Gr}(\mathbf{P}), \rho_j}} E_j\theta''$ . Moreover, by assumption we have that  $\theta(\mathbf{V}) \cong_{M_{\text{Gr}(\mathbf{P}), \rho}} \theta'(\mathbf{V})$ . Then, by definition  $M_{\text{Gr}(\mathbf{P})}(\theta(\mathbf{V}) E_1\theta' \cdots E_{m'}\theta') = M_{\text{Gr}(\mathbf{P})}(\theta'(\mathbf{V}) E_1\theta' \cdots E_{m'}\theta')$  and, by (1),  $M_{\text{Gr}(\mathbf{P})}(\theta'(\mathbf{V}) E_1\theta' \cdots E_{m'}\theta') = T_0$ .

In conclusion, the clause  $A\theta' \leftarrow K_1\theta''\theta', \dots, K_k\theta''\theta'$  is in  $\text{Gr}(\mathbf{P})$  and for each  $i \leq k$ , we have  $M_{\text{Gr}(\mathbf{P})}(K_i\theta''\theta') = T_0$ , therefore  $M_{\text{Gr}(\mathbf{P})}(A\theta') = T_0$  must also hold.

This concludes the proof for  $P'_1(0, n)$ . Next we prove  $P'_2(0, n+1)$ , assuming  $P'_2(0, n)$  holds. If  $T_{\mathbf{P}}^{n+1}(\emptyset)(A\theta) > F_0$ , then there exists a clause  $A\theta \leftarrow L_1, \dots, L_k$  in  $\text{Gr}(\mathbf{P})$  such that for each  $i \leq k$ ,  $T_{\mathbf{P}}^n(\emptyset)(L_i) > F_0$ . This clause is a ground instance of a clause  $\mathbf{pV}_1 \cdots \mathbf{V}_r \leftarrow K_1, \dots, K_k$  in the higher-order program and there exists a substitution  $\theta''$ , such that  $(\mathbf{pV}_1 \cdots \mathbf{V}_r)\theta'' = A$  and, for any variable  $\mathbf{V} \notin \{\mathbf{V}_1, \dots, \mathbf{V}_r\}$  appearing in the body of the clause,  $\theta''(\mathbf{V})$  is an appropriate ground term, so that  $L_i = K_i\theta''\theta$  for all  $i \leq k$ . Observe that the variables appearing in the clause  $(\mathbf{pV}_1 \cdots \mathbf{V}_r)\theta'' \leftarrow K_1\theta'', \dots, K_k\theta''$  are exactly the variables appearing in  $A$  and they are all of types simpler than  $\pi$ . We can distinguish the following cases for each  $K_i\theta''$ ,  $i \leq k$ :

- (1)  $K_i\theta''$  is a positive literal: A positive literal may take one of the three forms that we examined in our proof for property  $P'_1(0, n+1)$ . We can show that  $M_{\text{Gr}(\mathbf{P})}(K_i\theta''\theta') > F_0$  by the same arguments we used in each case.
- (2)  $K_i\theta''$  is a negative literal: A negative literal cannot be assigned the value  $F_0$  in any interpretation. Therefore  $M_{\text{Gr}(\mathbf{P})}(K_i\theta''\theta') > F_0$  holds by definition.

In conclusion, the clause  $A\theta' \leftarrow K_1\theta''\theta', \dots, K_k\theta''\theta'$  is in  $\text{Gr}(\mathbf{P})$  and for each  $i \leq k$ , we have  $M_{\text{Gr}(\mathbf{P})}(K_i\theta''\theta') > F_0$ , therefore  $M_{\text{Gr}(\mathbf{P})}(A\theta') > F_0$  must also hold.

This concludes the proof for  $P'_2(0, n)$ . We will now use properties  $P'_1(0, n)$  and  $P'_2(0, n)$  in order to show  $P_1(0)$  and  $P_2(0)$ . By definition, if  $M_0(A\theta) = T_{\mathbf{P}, 0}^\omega(\emptyset)(A\theta) = T_0$ , then there exists some  $n < \omega$  such that  $T_{\mathbf{P}}^n(\emptyset)(A\theta) = T_0$ . Applying  $P'_1(0, n)$  to  $A\theta$  we immediately conclude that  $M_{\text{Gr}(\mathbf{P})}(A\theta') = T_0$ , which establishes property  $P_1(0)$ . Now let  $M_0(A\theta) = F_0$  and assume  $M_{\text{Gr}(\mathbf{P})}(A\theta') \neq F_0$ . By Lemma 3.15, the latter can only hold if  $M_0(A\theta') = T_{\mathbf{P}, 0}^\omega(\emptyset)(A\theta') \neq F_0$  and this, in turn, means that there exists at least one  $n < \omega$  such that  $T_{\mathbf{P}}^n(\emptyset)(A\theta') > F_0$ . Then, reversing the roles of  $\theta$  and  $\theta'$ , we can apply property  $P'_2(0, n)$  to  $A\theta'$  and conclude that  $M_{\text{Gr}(\mathbf{P})}(A\theta) > F_0$ , which, again by Lemma 3.15, contradicts  $M_0(A\theta) = F_0$ . Therefore it must be  $M_{\text{Gr}(\mathbf{P})}(A\theta') = F_0$ .

**Second Induction Step:** Now we prove properties  $P_1(\alpha)$  and  $P_2(\alpha)$  for an arbitrary countable ordinal  $\alpha$ , assuming that  $P_1(\beta)$  and  $P_2(\beta)$  hold for all  $\beta < \alpha$ .

We have  $M_\alpha = T_{\mathbf{P}, \alpha}^\omega(I_\alpha)$ . Again, we first perform a third induction on  $n$  and prove two auxiliary properties, as follows:

$P'_1(\alpha, n)$ : if  $T_{\mathbf{P}}^n(I_\alpha)(A\theta) \geq T_\alpha$  then  $M_{\text{Gr}(\mathbf{P})}(A\theta') \geq T_\alpha$ ;

$P'_2(\alpha, n)$ : if  $T_{\mathbf{P}}^n(I_\alpha)(A\theta) > F_\alpha$  then  $M_{\text{Gr}(\mathbf{P})}(A\theta') > F_\alpha$ .

**Third Induction Basis ( $n = 0$ ):** We have  $T_{\mathbf{P}}^0(I_\alpha) = I_\alpha$ . Observe that, whether  $\alpha$  is a successor or a limit ordinal,  $I_\alpha$  does not assign to any atom the value  $T_\alpha$  or any value that is greater than  $F_\alpha$  and smaller than  $T_\alpha$ . So, whether we assume  $T_{\mathbf{P}}^0(I_\alpha)(A\theta) \geq T_\alpha$  or  $T_{\mathbf{P}}^0(I_\alpha)(A\theta) > F_\alpha$ , it must be  $T_{\mathbf{P}}^0(I_\alpha)(A\theta) = T_\beta$  for some ordinal  $\beta < \alpha$ . By Lemma

3.11,  $M_\alpha(A\theta) = T_\beta$  and so, by Lemma 3.15,  $M_\beta(A\theta) = T_\beta$ . Then, by the second induction hypothesis, property  $P_1(\beta)$  holds and yields  $M_{\text{Gr}(\mathcal{P})}(A\theta') = T_\beta$ , which implies both  $M_{\text{Gr}(\mathcal{P})}(A\theta') \geq T_\alpha$  and  $M_{\text{Gr}(\mathcal{P})}(A\theta') > F_\alpha$ . Therefore property  $P'_1(\alpha, 0)$  and property  $P'_2(\alpha, 0)$  also hold.

**Third Induction Step** ( $n + 1$ ): First we show  $P'_1(\alpha, n + 1)$ , assuming that  $P'_1(\alpha, n)$  holds.

If  $T_{\mathcal{P}}^{n+1}(I_\alpha)(A\theta) \geq T_\alpha$ , then there exists a clause  $A\theta \leftarrow L_1, \dots, L_k$  in  $\text{Gr}(\mathcal{P})$  such that for each  $i \leq k$ ,  $T_{\mathcal{P}}^n(I_\alpha)(L_i) \geq T_\alpha$ . This clause is a ground instance of a clause  $\mathfrak{p}V_1 \cdots V_r \leftarrow K_1, \dots, K_k$  in the higher-order program and there exists a substitution  $\theta''$ , such that  $(\mathfrak{p}V_1 \cdots V_r)\theta'' = A$  and, for any variable  $V \notin \{V_1, \dots, V_r\}$  appearing in the body of the clause,  $\theta''(V)$  is an appropriate ground term, so that  $L_i = K_i\theta''\theta$  for all  $i \leq k$ . As we observed earlier, the variables appearing in the clause  $(\mathfrak{p}V_1 \cdots V_r)\theta'' \leftarrow K_1\theta'', \dots, K_k\theta''$  are exactly the variables appearing in  $A$  and they are all of types simpler than  $\pi$ . We can distinguish the following cases for each  $K_i\theta''$ :

- (1)  $K_i\theta''$  is a positive literal: A positive literal may take one of the three forms that we examined in our proof for property  $P'_1(0, n + 1)$ . We can show that  $M_{\text{Gr}(\mathcal{P})}(K_i\theta''\theta') \geq T_\alpha$  by the same arguments we used in each case.
- (2)  $K_i\theta''$  is a negative literal and its atom starts with a predicate constant: Let  $K_i\theta''$  be of the form  $\sim B$ , where  $B$  is an atom that starts with a predicate constant. It is  $T_{\mathcal{P}}^n(I_\alpha)(\sim B\theta) = T_{\mathcal{P}}^n(I_\alpha)(K_i\theta''\theta) = T_{\mathcal{P}}^n(I_\alpha)(L_i) \geq T_\alpha$ . Then  $T_{\mathcal{P}}^n(I_\alpha)(B\theta) < F_\alpha$ , i.e.  $T_{\mathcal{P}}^n(I_\alpha)(B\theta) = F_\beta$  for some ordinal  $\beta < \alpha$ . By Lemma 3.11,  $M_\alpha(B\theta) = F_\beta$  and thus, by Lemma 3.15,  $M_\beta(B\theta) = F_\beta$ . By  $\text{vars}(K_i\theta'') \subseteq \text{vars}(A)$ , all the variables of  $B$  are of types simpler than  $\pi$ , so we can apply the second induction hypothesis, in particular property  $P_2(\beta)$ , to  $B\theta$  and conclude that  $M_{\text{Gr}(\mathcal{P})}(B\theta') = F_\beta$ . Then  $M_{\text{Gr}(\mathcal{P})}(K_i\theta''\theta') = M_{\text{Gr}(\mathcal{P})}(\sim B\theta') = T_{\beta+1} \geq T_\alpha$ .
- (3)  $K_i\theta''$  is a negative literal and its atom starts with a predicate variable: Let  $K_i\theta'' = \sim(V E_1 \cdots E_{m'})$  for some  $V \in \text{vars}(A)$ . Then  $B = \theta(V) E_1 \cdots E_{m'}$  is an atom that begins with a predicate constant and, by  $\text{vars}(K_i\theta'') \subseteq \text{vars}(A)$ , all the variables of  $B$  are of types simpler than  $\pi$ . Also,  $T_{\mathcal{P}}^n(I_\alpha)(\sim B\theta) = T_{\mathcal{P}}^n(I_\alpha)(K_i\theta''\theta) = T_{\mathcal{P}}^n(I_\alpha)(L_i) \geq T_\alpha$ . Then  $T_{\mathcal{P}}^n(I_\alpha)(B\theta) < F_\alpha$ , i.e.  $T_{\mathcal{P}}^n(I_\alpha)(B\theta) = F_\beta$  for some ordinal  $\beta < \alpha$ . By Lemma 3.11,  $M_\alpha(B\theta) = F_\beta$  and thus, by Lemma 3.15,  $M_\beta(B\theta) = F_\beta$ . By the second induction hypothesis, property  $P_2(\beta)$  gives us that  $M_{\text{Gr}(\mathcal{P})}(B\theta') = M_{\text{Gr}(\mathcal{P})}(\theta(V) E_1\theta' \cdots E_{m'}\theta') = F_\beta$  (1). The types of all arguments of  $\theta(V)$ , i.e. the types of  $E_j\theta'$  for all  $j \leq m'$ , are simpler than the type of  $V$  and consequently, since  $V \in \text{vars}(A)$ , simpler than  $\pi$ . For each  $j \leq m'$ , let  $\rho_j$  be the type of  $E_j$  and let  $\rho$  be the type of  $V$ ; by the first induction hypothesis,  $E_j\theta' \cong_{M_{\text{Gr}(\mathcal{P}), \rho_j}} E_j\theta'$ . Moreover, by assumption we have that  $\theta(V) \cong_{M_{\text{Gr}(\mathcal{P}), \rho}} \theta'(V)$ . Then, by definition  $M_{\text{Gr}(\mathcal{P})}(\theta(V) E_1\theta' \cdots E_{m'}\theta') = M_{\text{Gr}(\mathcal{P})}(\theta'(V) E_1\theta' \cdots E_{m'}\theta')$  and, by (1),  $M_{\text{Gr}(\mathcal{P})}(\theta'(V) E_1\theta' \cdots E_{m'}\theta') = F_\beta$ . Therefore, it follows that  $M_{\text{Gr}(\mathcal{P})}(K_i\theta''\theta') = M_{\text{Gr}(\mathcal{P})}(\sim(\theta'(V) E_1\theta' \cdots E_{m'}\theta')) = T_{\beta+1} \geq T_\alpha$ .

We can conclude that  $M_{\text{Gr}(\mathcal{P})}(A\theta') \geq T_\alpha$ , as the clause  $A\theta' \leftarrow K_1\theta''\theta', \dots, K_k\theta''\theta'$  is in  $\text{Gr}(\mathcal{P})$  and we have shown that, for each  $i \leq k$ ,  $M_{\text{Gr}(\mathcal{P})}(K_i\theta''\theta') \geq T_\alpha$ .

This concludes the proof of  $P'_1(\alpha, n)$ . Next we prove  $P'_2(\alpha, n + 1)$ , assuming  $P'_2(\alpha, n)$  holds. If  $T_{\mathcal{P}}^{n+1}(I_\alpha)(A\theta) > F_\alpha$ , then there exists a clause  $A\theta \leftarrow L_1, \dots, L_k$  in  $\text{Gr}(\mathcal{P})$  such that for each  $i \leq k$ ,  $T_{\mathcal{P}}^n(I_\alpha)(L_i) > F_\alpha$ . This clause is a ground instance of a clause  $\mathfrak{p}V_1 \cdots V_r \leftarrow K_1, \dots, K_k$  in the higher-order program and there exists a substitution  $\theta''$  such that  $(\mathfrak{p}V_1 \cdots V_r)\theta'' = A$  and, for any variable  $V \notin \{V_1, \dots, V_r\}$  appearing in the body of the clause,  $\theta''(V)$  is an appropriate ground term, so that  $L_i = K_i\theta''\theta$  for all



$i \leq k$ . Again we observe that the variables appearing in the clause  $(\mathbf{p} \vee_1 \cdots \vee_r) \theta'' \leftarrow \mathbf{K}_1 \theta'', \dots, \mathbf{K}_k \theta''$  are exactly the variables appearing in  $\mathbf{A}$ , which are all of types simpler than  $\pi$ , and distinguish the following cases for each  $\mathbf{K}_i \theta''$ :

- (1)  $\mathbf{K}_i \theta''$  is a positive literal: A positive literal may take one of the three forms that we examined in our proof for property  $P'_1(0, n+1)$ . We can show that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{K}_i \theta'' \theta') > F_\alpha$  by the same arguments we used in each case.
- (2)  $\mathbf{K}_i \theta''$  is a negative literal and its atom starts with a predicate constant: Let  $\mathbf{K}_i \theta''$  be of the form  $\sim \mathbf{B}$ , where  $\mathbf{B}$  is an atom that starts with a predicate constant and, by  $\text{vars}(\mathbf{K}_i \theta'') \subseteq \text{vars}(\mathbf{A})$ , all the variables of  $\mathbf{B}$  are of types simpler than  $\pi$ . It is  $T_{\mathbf{P}}^n(I_\alpha)(\sim \mathbf{B} \theta) = T_{\mathbf{P}}^n(I_\alpha)(\mathbf{K}_i \theta'' \theta) = T_{\mathbf{P}}^n(I_\alpha)(L_i) > F_\alpha$  and we claim that  $M_{\text{Gr}(\mathbf{P})}(\sim \mathbf{B} \theta') > F_\alpha$ . For the sake of contradiction, assume that  $M_{\text{Gr}(\mathbf{P})}(\sim \mathbf{B} \theta') \leq F_\alpha$ . Then  $M_{\text{Gr}(\mathbf{P})}(\mathbf{B} \theta') > T_\alpha$ , i.e.  $M_{\text{Gr}(\mathbf{P})}(\mathbf{B} \theta') = T_\beta$  for some ordinal  $\beta < \alpha$ . Therefore, Lemma 3.15 implies that  $M_\beta(\mathbf{B} \theta') = T_\beta$ . This means that, if we reverse the roles of  $\theta$  and  $\theta'$ , we can apply the second induction hypothesis to  $\mathbf{B} \theta'$  and use property  $P_1(\beta)$  to conclude that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{B} \theta) = T_\beta$ . By Lemma 3.15, this implies that  $M_\alpha(\mathbf{B} \theta) = T_\beta$  and, by Lemma 3.11,  $T_{\mathbf{P}}^n(I_\alpha)(\mathbf{B} \theta) = T_\beta$ . Then  $T_{\mathbf{P}}^n(I_\alpha)(\sim \mathbf{B} \theta) = F_{\beta+1} \leq F_\alpha$ , which is obviously a contradiction. So it must be  $M_{\text{Gr}(\mathbf{P})}(\mathbf{K}_i \theta'' \theta') = M_{\text{Gr}(\mathbf{P})}(\sim \mathbf{B} \theta') > F_\alpha$ .
- (3)  $\mathbf{K}_i \theta''$  is a negative literal and its atom starts with a predicate variable: Let  $\mathbf{K}_i \theta'' = \sim(\mathbf{V} \mathbf{E}_1 \cdots \mathbf{E}_{m'})$  for some  $\mathbf{V} \in \text{vars}(\mathbf{A})$ . Then  $\mathbf{B} = \theta(\mathbf{V}) \mathbf{E}_1 \cdots \mathbf{E}_{m'}$  is an atom that begins with a predicate constant and, by  $\text{vars}(\mathbf{K}_i \theta'') \subseteq \text{vars}(\mathbf{A})$ , all the variables of  $\mathbf{B}$  are of types simpler than  $\pi$ . Also,  $T_{\mathbf{P}}^n(I_\alpha)(\sim \mathbf{B} \theta) = T_{\mathbf{P}}^n(I_\alpha)(\mathbf{K}_i \theta'' \theta) = T_{\mathbf{P}}^n(I_\alpha)(L_i) > F_\alpha$ . We claim that  $M_{\text{Gr}(\mathbf{P})}(\sim \mathbf{B} \theta') > F_\alpha$ . Again, assume that this is not so, i.e.  $M_{\text{Gr}(\mathbf{P})}(\sim \mathbf{B} \theta') \leq F_\alpha$ ; then  $M_{\text{Gr}(\mathbf{P})}(\mathbf{B} \theta') = T_\beta > T_\alpha$  for some ordinal  $\beta < \alpha$ . By Lemma 3.15,  $M_\beta(\mathbf{B} \theta') = T_\beta$  and the second induction hypothesis applies. So if we reverse the roles of  $\theta$  and  $\theta'$  property  $P_1(\beta)$  gives us that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{B} \theta) = T_\beta$ . By Lemma 3.15,  $M_\alpha(\mathbf{B} \theta) = T_\beta$  and thus, by Lemma 3.11,  $T_{\mathbf{P}}^n(I_\alpha)(\mathbf{B} \theta) = T_\beta$ . This is a contradiction, as it implies that  $T_{\mathbf{P}}^n(I_\alpha)(\sim \mathbf{B} \theta) = F_{\beta+1} \leq F_\alpha$ . Therefore it must hold that  $M_{\text{Gr}(\mathbf{P})}(\sim \mathbf{B} \theta') = M_{\text{Gr}(\mathbf{P})}(\sim(\theta(\mathbf{V}) \mathbf{E}_1 \theta' \cdots \mathbf{E}_{m'} \theta')) > F_\alpha$  (1). The types of all arguments of  $\theta(\mathbf{V})$ , i.e. the types of  $\mathbf{E}_j \theta'$  for all  $j \leq m'$ , are simpler than the type of  $\mathbf{V}$  and therefore simpler than  $\pi$ . For each  $j \leq m'$ , let  $\rho_j$  be the type of  $\mathbf{E}_j$  and let  $\rho$  be the type of  $\mathbf{V}$ ; by the first induction hypothesis,  $\mathbf{E}_j \theta' \cong_{M_{\text{Gr}(\mathbf{P}), \rho_j}} \mathbf{E}_j \theta'$  and by assumption  $\theta(\mathbf{V}) \cong_{M_{\text{Gr}(\mathbf{P}), \rho}} \theta'(\mathbf{V})$ . Then, by definition  $M_{\text{Gr}(\mathbf{P})}(\theta(\mathbf{V}) \mathbf{E}_1 \theta' \cdots \mathbf{E}_{m'} \theta') = M_{\text{Gr}(\mathbf{P})}(\theta'(\mathbf{V}) \mathbf{E}_1 \theta' \cdots \mathbf{E}_{m'} \theta')$  and, consequently,  $M_{\text{Gr}(\mathbf{P})}(\sim(\theta(\mathbf{V}) \mathbf{E}_1 \theta' \cdots \mathbf{E}_{m'} \theta')) = M_{\text{Gr}(\mathbf{P})}(\sim(\theta'(\mathbf{V}) \mathbf{E}_1 \theta' \cdots \mathbf{E}_{m'} \theta'))$ . Therefore by (1),  $M_{\text{Gr}(\mathbf{P})}(\sim(\theta'(\mathbf{V}) \mathbf{E}_1 \theta' \cdots \mathbf{E}_{m'} \theta')) > F_\alpha$  and so we can conclude that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{K}_i \theta'' \theta') = M_{\text{Gr}(\mathbf{P})}(\sim(\theta'(\mathbf{V}) \mathbf{E}_1 \theta' \cdots \mathbf{E}_{m'} \theta')) > F_\alpha$ .

Observe that the clause  $\mathbf{A} \theta' \leftarrow \mathbf{K}_1 \theta'' \theta', \dots, \mathbf{K}_k \theta'' \theta'$  is in  $\text{Gr}(\mathbf{P})$  and we have shown that, for each  $i \leq k$ ,  $M_{\text{Gr}(\mathbf{P})}(\mathbf{K}_i \theta'' \theta') > F_\alpha$ . So  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A} \theta') > F_\alpha$  must also hold.

This concludes the proof for  $P'_2(\alpha, n)$ . We will now use properties  $P'_1(\alpha, n)$  and  $P'_2(\alpha, n)$  in order to show  $P_1(\alpha)$  and  $P_2(\alpha)$ . By definition, if  $M_\alpha(\mathbf{A} \theta) = T_{\mathbf{P}, \alpha}^\omega(I_\alpha)(\mathbf{A} \theta) = T_\alpha$ , then there exists some  $n < \omega$  such that  $T_{\mathbf{P}}^n(I_\alpha)(\mathbf{A} \theta) = T_\alpha$ . As we have shown above, property  $P'_1(\alpha, n)$  yields  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A} \theta') \geq T_\alpha$ . However, if it was  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A} \theta') = T_\beta > T_\alpha$  for some ordinal  $\beta < \alpha$ , then, by Lemma 3.15 we would also have  $M_\beta(\mathbf{A} \theta') = T_\beta$ . By the second induction hypothesis we would be able to apply property  $P_1(\beta)$  to  $\mathbf{A} \theta'$  and infer that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A} \theta) = T_\beta$ , which, again by Lemma 3.15, contradicts  $M_\alpha(\mathbf{A} \theta) = T_\alpha$ . So  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A} \theta')$  can only be equal to  $T_\alpha$  and property  $P_1(\alpha)$  holds. Now let  $M_\alpha(\mathbf{A} \theta) = F_\alpha$

and assume  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta') \neq F_\alpha$ , i.e. either  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta') < F_\alpha$  or  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta') > F_\alpha$ . In the first case,  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta') = F_\beta$  and, by Lemma 3.15,  $M_\beta(\mathbf{A}\theta') = F_\beta$  for some  $\beta < \alpha$ . By the second induction hypothesis, property  $P_2(\beta)$  gives us that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta) = F_\beta < F_\alpha$ . In the second case, Lemma 3.15 implies that  $M_\alpha(\mathbf{A}\theta') = T_{\mathbf{P},\alpha}^\omega(I_\alpha)(\mathbf{A}\theta') > F_\alpha$  and this, in turn, means that there exists at least one  $n < \omega$  such that  $T_{\mathbf{P}}^n(I_\alpha)(\mathbf{A}\theta') > F_\alpha$ . Then, reversing the roles of  $\theta$  and  $\theta'$ , we can apply property  $P_2'(\alpha, n)$  to  $\mathbf{A}\theta'$  and conclude that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta) > F_\alpha$ . In both cases, our conclusion constitutes a contradiction, because, by Lemma 3.15,  $M_\alpha(\mathbf{A}\theta) = F_\alpha$  implies that  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta) = F_\alpha$ . Therefore it must also be  $M_{\text{Gr}(\mathbf{P})}(\mathbf{A}\theta') = F_\alpha$  and this proves property  $P_2(\alpha)$ .  $\square$

## 7. STRATIFIED AND LOCALLY STRATIFIED PROGRAMS

In this section we define the notions of *stratified* and *locally stratified* programs and argue that atoms of such programs never obtain the truth value 0 under the proposed semantics. The notion of local stratification is a straightforward generalization of the corresponding notion for classical (first-order) logic programs. However, the notion of stratification is a genuine extension of the corresponding notion for first-order programs.

**Definition 7.1.** A program  $\mathbf{P}$  is called *locally stratified* if and only if it is possible to decompose the Herbrand base  $U_{\mathbf{P},o}$  of  $\mathbf{P}$  into disjoint sets (called *strata*)  $S_1, S_2, \dots, S_\alpha, \dots, \alpha < \gamma$ , where  $\gamma$  is a countable ordinal, such that for every clause  $\mathbf{H} \leftarrow \mathbf{A}_1, \dots, \mathbf{A}_m, \sim \mathbf{B}_1, \dots, \sim \mathbf{B}_n$  in  $\text{Gr}(\mathbf{P})$ , we have that for every  $i \leq m$ ,  $\text{stratum}(\mathbf{A}_i) \leq \text{stratum}(\mathbf{H})$  and for every  $i \leq n$ ,  $\text{stratum}(\mathbf{B}_i) < \text{stratum}(\mathbf{H})$ , where *stratum* is a function such that  $\text{stratum}(\mathbf{C}) = \beta$ , if the atom  $\mathbf{C} \in U_{\mathbf{P},o}$  belongs to  $S_\beta$ , and  $\text{stratum}(\mathbf{C}) = 0$ , if  $\mathbf{C} \notin U_{\mathbf{P},o}$  and is of the form  $(\mathbf{E}_1 \approx \mathbf{E}_2)$ .

All atoms in the minimum Herbrand model of a locally stratified program have non-zero values:

**Lemma 7.2.** *Let  $\mathbf{P}$  be a locally stratified logic program. Then, for every atom  $\mathbf{A} \in U_{\mathbf{P},o}$  it holds  $\mathcal{M}_{\mathbf{P}}(\mathbf{A}) \neq 0$ .*

*Proof.* Theorem 3.14 implies that the infinite-valued model  $M_{\text{Gr}(\mathbf{P})}$  of the ground instantiation of  $\mathbf{P}$  assigns the truth value 0 to an atom iff this atom is assigned the truth value 0 by the well-founded model of  $\text{Gr}(\mathbf{P})$ . Notice now that, by Definition 7.1, if  $\mathbf{P}$  is a locally stratified higher-order program, then  $\text{Gr}(\mathbf{P})$  is in turn a locally stratified propositional program (having exactly the same local stratification as  $\mathbf{P}$ ). Recall that the well-founded model of a locally stratified propositional program does not assign the truth value 0 to any atom [14], so neither does  $M_{\text{Gr}(\mathbf{P})}$  or, consequently,  $\mathcal{M}_{\mathbf{P}}$ .  $\square$

Since Definition 7.1 generalizes the corresponding one for classical logic programs, the undecidability result [10] for detecting whether a given program is locally stratified, extends directly to the higher-order case.

**Lemma 7.3.** *The problem of determining whether a given logic program  $\mathbf{P}$  is locally stratified, is undecidable.*

However, there exists a notion of stratification for higher-order logic programs that is decidable and has as a special case the stratification for classical logic programs [1]. In the following definition, a predicate type  $\pi$  is understood to be *greater than* a second predicate type  $\pi'$ , if  $\pi$  is of the form  $\rho_1 \rightarrow \dots \rightarrow \rho_n \rightarrow \pi'$ , where  $n \geq 1$ .

**Definition 7.4.** A program  $P$  is called *stratified* if and only if it is possible to decompose the set of all predicate constants that appear in  $P$  into a finite number  $r$  of disjoint sets (called *strata*)  $S_1, S_2, \dots, S_r$ , such that for every clause  $H \leftarrow A_1, \dots, A_m, \sim B_1, \dots, \sim B_n$  in  $P$ , where the predicate constant of  $H$  is  $p$ , we have:

- (1) for every  $i \leq m$ , if  $A_i$  is a term that starts with a predicate constant  $q$ , then  $\text{stratum}(q) \leq \text{stratum}(p)$ ;
- (2) for every  $i \leq m$ , if  $A_i$  is a term that starts with a predicate variable  $Q$ , then for all predicate constants  $q$  that appear in  $P$  such that the type of  $q$  is greater than or equal to the type of  $Q$ , it holds  $\text{stratum}(q) \leq \text{stratum}(p)$ ;
- (3) for every  $i \leq n$ , if  $B_i$  starts with a predicate constant  $q$ , then  $\text{stratum}(q) < \text{stratum}(p)$ ;
- (4) for every  $i \leq n$ , if  $B_i$  starts with a predicate variable  $Q$ , then for all predicate constants  $q$  that appear in  $P$  such that the type of  $q$  is greater than or equal to the type of  $Q$ , it holds  $\text{stratum}(q) < \text{stratum}(p)$ ;

where for every predicate constant  $r$ ,  $\text{stratum}(r) = i$  if the predicate symbol  $r$  belongs to  $S_i$ .

**Example 7.5.** In the following program:

$$\begin{aligned} p \ Q: & \sim(Q \ a) . \\ q \ X: & \sim(X \approx a) . \end{aligned}$$

the variable  $Q$  is of type  $\iota \rightarrow o$  and  $X$  is of type  $\iota$ . The only predicate constants that appear in the program are  $p$ , which is of type  $(\iota \rightarrow o) \rightarrow o$ , and  $q$ , which is of type  $\iota \rightarrow o$ . Note that the type of  $p$  is neither equal nor greater than the type of  $Q$ , while the type of  $q$  is the same as that of  $Q$ . It is straightforward to see that the program is stratified, if we choose  $S_1 = \{q\}$  and  $S_2 = \{p\}$ . Indeed, for the first clause, we have  $\text{stratum}(p) > \text{stratum}(q)$  and in the second clause there are no predicate constants or predicate variables appearing in its body. However, if  $Q$  and  $X$  are as above and, moreover,  $Y$  is of type  $\iota$ , it can easily be checked that the program:

$$\begin{aligned} p \ Q: & \sim(Q \ a) . \\ q \ X \ Y: & \sim(X \approx a), (Y \approx a), p \ (q \ a) . \end{aligned}$$

is not stratified nor locally stratified, because if the term  $q \ a$  is substituted for  $Q$  we get a circularity through negation. Notice that the type of  $q$  is  $\iota \rightarrow \iota \rightarrow o$  and it is greater than the type of  $Q$  which is  $\iota \rightarrow o$ .  $\square$

Since the set of predicate constants that appear in a program  $P$  is finite, and since the number of predicate constants of the program that have a greater or equal type than the type of a given predicate variable is also finite, it follows that checking whether a given program is stratified, is decidable. Moreover, we have the following theorem:

**Theorem 7.6.** *If  $P$  is stratified then it is locally stratified.*

*Proof.* Consider a decomposition  $S_1, \dots, S_r$  of the set of predicate constants of  $P$  such that the requirements of Definition 7.4 are satisfied. This defines a decomposition  $S'_1, \dots, S'_r$  of the Herbrand base of  $P$ , as follows:

$$S'_i = \{A \in U_{P,o} \mid \text{the leftmost predicate constant of } A \text{ belongs to } S_i\}$$

We show that  $S'_1, \dots, S'_r$  corresponds to a local stratification of  $U_{P,o}$ . Consider a clause in  $P$  of the form  $H' \leftarrow A'_1, \dots, A'_m, \sim B'_1, \dots, \sim B'_n$  and let  $H \leftarrow A_1, \dots, A_m, \sim B_1, \dots, \sim B_n$  be one of its ground instances. Let  $p$  be the predicate constant of  $H$  (and  $H'$ ). Consider any  $A'_i$ . If  $A'_i$  starts with a predicate constant, say  $q_i$ , by Definition 7.4, it is  $\text{stratum}(p) \geq \text{stratum}(q_i)$ .

By the definition of the local stratification decomposition we gave above, it is  $\text{stratum}(\mathbf{H}) = \text{stratum}(\mathbf{p})$  and  $\text{stratum}(\mathbf{A}_i) = \text{stratum}(\mathbf{q}_i)$ , and therefore  $\text{stratum}(\mathbf{H}) \geq \text{stratum}(\mathbf{A}_i)$ . If  $\mathbf{A}'_i$  starts with a predicate variable, say  $\mathbf{Q}$ , then  $\mathbf{Q}$  has been substituted in  $\mathbf{A}'_i$  with a term starting with a predicate constant, say  $\mathbf{q}_i$ , that has a type greater than or equal to that of  $\mathbf{Q}$ . By Definition 7.4, it is  $\text{stratum}(\mathbf{p}) \geq \text{stratum}(\mathbf{q}_i)$  and by the definition of the local stratification decomposition we gave above, it is  $\text{stratum}(\mathbf{H}) = \text{stratum}(\mathbf{p})$  and  $\text{stratum}(\mathbf{A}_i) = \text{stratum}(\mathbf{q}_i)$ , and therefore  $\text{stratum}(\mathbf{H}) \geq \text{stratum}(\mathbf{A}_i)$ . The justification for the case of negative literals, is similar and omitted.  $\square$

## 8. NON-EXTENSIONALITY OF THE STABLE MODELS

It is natural to wonder whether the more traditional approaches to the semantics of negation in logic programming, also lead to extensional semantics when applied to higher-order logic programs with negation under the framework of [2, 3]. The two most widely known such approaches are the *stable model* semantics [15] and the *well-founded* [14] one. It was recently demonstrated [22] that the well-founded approach does not in general lead to an extensional well-founded model when applied to higher-order programs. In this section we demonstrate that the stable model semantics also fails to give extensional models in the general case. We believe that these two negative results reveal the importance of the use of the infinite-valued approach for obtaining extensionality.

The stable model semantics, in its original form introduced in [15], is applied on the ground instantiation of a given first-order logic program with negation, which is a (possibly infinite) propositional program. In this respect, it is not hard to adapt it to apply to the framework of [2, 3], which is based on the ground instantiation of a given higher-order program (which is again a possibly infinite propositional program). For reasons of completeness, we include the definition of stable models [15] (see also [17]).

**Definition 8.1.** Let  $\mathbf{P}$  be a propositional program and let  $I$  be a set of propositional variables. The *reduct*  $\mathbf{P}_I$  of  $\mathbf{P}$  with respect to  $I$ , is the set of rules without negation that can be obtained from  $\mathbf{P}$  by first dropping every rule of  $\mathbf{P}$  that contains a negative literal  $\sim \mathbf{p}$  in its body such that  $\mathbf{p} \in I$ , and then dropping all negative literals from the bodies of all the remaining rules. The set  $I$  is called a *stable model* of  $\mathbf{P}$  if  $I$  coincides with the least model of  $\mathbf{P}_I$ .

To demonstrate the non-extensionality of the stable models approach in the case of higher-order programs, it suffices to find a program that produces non-extensional stable models. The following very simple example does exactly this.

**Example 8.2.** Consider the higher-order program:

$$\begin{aligned} r(Q) &: \sim s(Q). \\ s(Q) &: \sim r(Q). \\ q(a) &. \\ p(a) &. \end{aligned}$$

where the predicate variable  $Q$  of the first and second clause is of type  $\iota \rightarrow o$ . We at first take the ground instantiation of the above program:

$$\begin{aligned} r(p) &: \sim s(p). \\ r(q) &: \sim s(q). \\ s(p) &: \sim r(p). \\ s(q) &: \sim r(q). \\ q(a) &. \\ p(a) &. \end{aligned}$$

Consider now the Herbrand interpretation  $M = \{p(a), q(a), s(p), r(q)\}$ . One can easily check that  $M$  is a model of the ground instantiation of the program. However, the above model is not extensional: since  $p$  and  $q$  are extensionally equal, the atoms  $s(q)$  and  $r(p)$  should also belong to  $M$  in order to ensure extensionality. It remains to show that  $M$  is also a stable model. Consider the reduct of the above program based on  $M$ :

$$\begin{aligned} r(q) &. \\ s(p) &. \\ q(a) &. \\ p(a) &. \end{aligned}$$

Obviously, the least model of the reduct is the interpretation  $M$ , and therefore  $M$  is a stable model of the initial program. In other words, we have found a program with a non-extensional stable model.  $\square$

Continuing the discussion on the above example, one can easily verify that the above program also has two extensional stable models, namely  $M_1 = \{p(a), q(a), s(p), s(q)\}$  and  $M_2 = \{p(a), q(a), r(p), r(q)\}$ . This creates the hope that we could somehow adapt the standard stable model construction procedure in order to produce only extensional stable models. The following example makes this hope vanish.

**Example 8.3.** Consider the program:

$$\begin{aligned} r(Q) &: \sim s(Q), \sim r(p). \\ s(Q) &: \sim r(Q), \sim s(q). \\ q(a) &. \\ p(a) &. \end{aligned}$$

where, in the first two clauses,  $Q$  is of type  $\iota \rightarrow o$ . The ground instantiation of the program is the following:

$$\begin{aligned} r(p) &: \sim s(p), \sim r(p). \\ r(q) &: \sim s(q), \sim r(p). \\ s(p) &: \sim r(p), \sim s(q). \\ s(q) &: \sim r(q), \sim s(q). \\ q(a) &. \\ p(a) &. \end{aligned}$$

This program has the non-extensional stable model  $M = \{p(a), q(a), s(p), r(q)\}$ . However, it has *no* extensional stable models: there are four possible extensional interpretations that are potential candidates, namely  $M_1 = \{p(a), q(a)\}$ ,  $M_2 = \{p(a), q(a), r(p), r(q)\}$ ,  $M_3 = \{p(a), q(a), s(p), s(q)\}$ , and  $M_4 = \{p(a), q(a), s(p), s(q), r(p), r(q)\}$ ; one can easily verify that none of these interpretations is a stable model of the ground instantiation

of the program. The conclusion is that there exist higher-order logic programs with negation which have only non-extensional stable models!  $\square$

The above examples seem to suggest that the extensional approach of [2, 3] is incompatible with the stable model semantics. Possibly this behaviour can be explained by an inherent characteristic of the stable model semantics, which appears even in the case of classical logic programs with negation. Consider for example the simple propositional program:

$$\begin{aligned} p &: \sim \sim q. \\ q &: \sim \sim p. \end{aligned}$$

The above program has two stable models, namely  $\{p\}$  and  $\{q\}$ . In the first stable model,  $p$  is true and  $q$  is false despite the fact that the program is completely symmetric and there is no apparent reason to prefer  $p$  over  $q$ ; a similar remark applies to the second stable model. It is possible that it is this characteristic of stable models (i.e., the resolution of negative circularities by making specific choices that are not necessarily symmetric) that leads to their non-extensionality.

## 9. CONNECTIONS TO THE RESEARCH OF ZOLTÁN ÉSIK

The work reported in this paper is closely connected with research conducted by Zoltán Ésik in collaboration with the first author (Panos Rondogiannis). In this section we briefly describe the roots of this joint collaboration which unfortunately was abruptly interrupted by the untimely loss of Zoltán.

In 2005, the first author together with Bill Wadge proposed [20] the infinite-valued semantics for logic programs with negation (an overview of this work is given in Section 3 of the present paper). In 2013, the first author together with Zoltán Ésik started a collaboration supported by a Greek-Hungarian Scientific Collaboration Program with title “Extensions and Applications of Fixed Point Theory for Non-Monotonic Formalisms”. The purpose of the program was to create an abstract fixed point theory based on the infinite-valued approach, namely a theory that would not only be applicable to logic programs but also to other non-monotonic formalisms. This abstract theory was successfully developed and is described in detail in [11]. As an application of these results, the first extensional semantics for higher-order logic programs with negation was developed in [6]. Another application of this new theory to the area of non-monotonic formal grammars was proposed in [12]. Moreover, Zoltán himself further investigated the foundations and the properties of the infinite-valued approach [13], highlighting some of its desirable characteristics.

The work reported in the present paper is an alternative approach to the semantics of higher-order logic programs with negation developed in collaboration with Zoltán in [6]. The two approaches both use the infinite-valued approach but are radically different otherwise. In particular, the present approach heavily relies on the ground instantiation of the source program, while the approach of [6] operates directly on the source program and extensively uses domain theoretic constructions. It is easy to find a program where our approach gives a different denotation from that of [6]. Actually, the two approaches differ even for positive programs. The example given below is borrowed from [8] where it was used to demonstrate the differences between the approach of [2, 3] versus that of [7] (which applies to positive programs but has similarities to [6] because it is also based on domain theory).

**Example 9.1.** Consider the following program:

$$p(\mathbf{a}) : \neg Q(\mathbf{a}).$$

where the predicate variable  $Q$  is of type  $\iota \rightarrow o$ . Under the semantics developed in this paper (which generalizes that of [2, 3]), the above program intuitively states that  $p$  is true of  $\mathbf{a}$  if there exists a *predicate that is defined in the program* that is true of  $\mathbf{a}$ . If we take the ground instantiation of the above program:

$$p(\mathbf{a}) : \neg p(\mathbf{a}).$$

and compute the least model of the instantiation, we find the least model of the program which assigns to the atom  $p(\mathbf{a})$  the truth value  $F_0$ .

Under the semantics of [6] the atom  $p(\mathbf{a})$  is true in the minimum Herbrand model of the initial program. This is because in this semantics, the initial program reads (intuitively speaking) as follows: “ $p(\mathbf{a})$  is true if there exists a *relation* that is true of  $\mathbf{a}$ ”; actually, there exists one such relation, namely one in which the constant  $\mathbf{a}$  has the value  $T_0$ . This discrepancy between the semantics of [2, 3] and that of [6] is due to the fact that the latter is based on (infinite-valued) *sets* and not on the syntactic entities that appear in the program.  $\square$

Despite the above difference, there certainly exists common ground between the two techniques. As it is demonstrated in [8], there exists a large and useful class of positive programs for which the approach of [2, 3] coincides with that of [7]. Intuitively speaking, this class contains all positive programs that do not contain *existential* variables in the bodies of clauses (like  $Q$  in the above example). We believe that such a result also holds for programs with negation. More specifically, we conjecture that there exists a large fragment of  $\mathcal{H}$  for which the semantics of [6] coincides with the one given in the present paper. Establishing such a relationship between the two semantics, will lead to a better understanding of extensional higher-order logic programming.

## REFERENCES

- [1] Krzysztof R. Apt, Howard A. Blair, and Adrian Walker. Towards a theory of declarative knowledge. In Jack Minker, editor, *Foundations of Deductive Databases and Logic Programming*, pages 89–148. Morgan Kaufmann, 1988.
- [2] Marc Bezem. Extensionality of simply typed logic programs. In Danny De Schreye, editor, *Logic Programming: The 1999 International Conference, Las Cruces, New Mexico, USA, November 29 - December 4, 1999*, pages 395–410. MIT Press, 1999.
- [3] Marc Bezem. An improved extensionality criterion for higher-order logic programs. In Laurent Fribourg, editor, *Computer Science Logic, 15th International Workshop, CSL 2001. 10th Annual Conference of the EACSL, Paris, France, September 10-13, 2001, Proceedings*, volume 2142 of *Lecture Notes in Computer Science*, pages 203–216. Springer, 2001.
- [4] Stephen L. Bloom and Zoltán Ésik. *Iteration Theories - The Equational Logic of Iterative Processes*. EATCS Monographs on Theoretical Computer Science. Springer, 1993.
- [5] Arnaud Carayol and Zoltán Ésik. An analysis of the equational properties of the well-founded fixed point. In Chitta Baral, James P. Delgrande, and Frank Wolter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016.*, pages 533–536. AAAI Press, 2016.
- [6] Angelos Charalambidis, Zoltán Ésik, and Panos Rondogiannis. Minimum model semantics for extensional higher-order logic programming with negation. *TPLP*, 14(4-5):725–737, 2014.
- [7] Angelos Charalambidis, Konstantinos Handjopoulos, Panos Rondogiannis, and William W. Wadge. Extensional higher-order logic programming. *ACM Trans. Comput. Log.*, 14(3):21, 2013.

- [8] Angelos Charalambidis, Panos Rondogiannis, and Ioanna Symeonidou. Equivalence of two fixed-point semantics for definitional higher-order logic programs. In Ralph Matthes and Matteo Mio, editors, *Proceedings Tenth International Workshop on Fixed Points in Computer Science, FICS 2015, Berlin, Germany, September 11-12, 2015.*, volume 191 of *EPTCS*, pages 18–32, 2015 (to appear in extended form in *Theoretical Computer Science*, 2017).
- [9] Weidong Chen, Michael Kifer, and David Scott Warren. HILOG: A foundation for higher-order logic programming. *Journal of Logic Programming*, 15(3):187–230, 1993.
- [10] Peter Cholak and Howard A. Blair. The complexity of local stratification. *Fundam. Inform.*, 21(4):333–344, 1994.
- [11] Zoltán Ésik, and Panos Rondogiannis. A fixed point theorem for non-monotonic functions. *Theoretical Computer Science*, 574:18–38, 2015.
- [12] Zoltán Ésik, and Panos Rondogiannis. Theorems on Pre-fixed Points of Non-Monotonic Functions with Applications in Logic Programming and Formal Grammars. In Ulrich Kohlenbach, Pablo Barceló, and Ruy J. G. B. de Queiroz, editors, *Logic, Language, Information, and Computation - 21st International Workshop, WoLLIC 2014, Valparaíso, Chile, September 1-4, 2014, Proceedings*, volume 8652 of *Lecture Notes in Computer Science*, pages 166–180. Springer, 2014.
- [13] Zoltán Ésik. Equational properties of stratified least fixed points (extended abstract). In Valeria de Paiva, Ruy J. G. B. de Queiroz, Lawrence S. Moss, Daniel Leivant, and Anjolina Grisi de Oliveira, editors, *Logic, Language, Information, and Computation - 22nd International Workshop, WoLLIC 2015, Bloomington, IN, USA, July 20-23, 2015, Proceedings*, volume 9160 of *Lecture Notes in Computer Science*, pages 174–188. Springer, 2015.
- [14] Allen Van Gelder, Kenneth A. Ross, and John S. Schlipf. The well-founded semantics for general logic programs. *J. ACM*, 38(3):620–650, 1991.
- [15] Michael Gelfond, and Vladimir Lifschitz. The Stable Model Semantics for Logic Programming. In *Proceedings of the Fifth Logic Programming Symposium*. MIT Press, 1070–1080, 1988.
- [16] Vassilis Kountouriotis, Panos Rondogiannis, and William W. Wadge. Extensional higher-order datalog. In *Short Paper Proceeding of the 12th International Conference on Logic for Programming, Artificial Intelligence and Reasoning (LPAR)*, pages 1–5, December 2005.
- [17] Vladimir Lifschitz. Twelve Definitions of a Stable Model. In *Proceedings of the 24th International Conference on Logic Programming (ICLP)*, pages 37–51, 2008.
- [18] John W. Lloyd. *Foundations of Logic Programming*. Springer Verlag, 1987.
- [19] Dale Miller and Gopalan Nadathur. *Programming with Higher-Order Logic*. Cambridge University Press, New York, NY, USA, 1st edition, 2012.
- [20] Panos Rondogiannis and William W. Wadge. Minimum model semantics for logic programs with negation-as-failure. *ACM Trans. Comput. Log.*, 6(2):441–467, 2005.
- [21] Panos Rondogiannis and Ioanna Symeonidou. Extensional Semantics for Higher-Order Logic Programs with Negation. In *Proceedings of the 15th European Conference on Logics in Artificial Intelligence (JELIA)*, pages 447–462, 2016.
- [22] Panos Rondogiannis and Ioanna Symeonidou. The intricacies of three-valued extensional semantics for higher-order logic programs. *TPLP*, 17(5-6):974–991, 2017.
- [23] William W. Wadge. Higher-order horn logic programming. In Vijay A. Saraswat and Kazunori Ueda, editors, *Logic Programming, Proceedings of the 1991 International Symposium, San Diego, California, USA, Oct. 28 - Nov 1, 1991*, pages 289–303. MIT Press, 1991.