

A Decidable Class of Groundable Formulas in the General Theory of Stable Models

Michael Bartholomew and Joohyung Lee

School of Computing, Informatics and Decision System Engineering
Arizona State University, Tempe, USA

Abstract

We present a decidable class of first-order formulas in the general theory of stable models that can be instantiated even in the presence of function constants. The notion of an *argument-restricted formula* presented here is a generalization of both the notion of an argument-restricted program and the notion of a semi-safe sentence that have been defined in different contexts. Based on this, for RASPL-1 programs and programs with arbitrary aggregates, we define the notion of safety, which ensures finite groundability of those programs even in the presence of function constants. We also show that under certain conditions, argument-restricted formulas can be turned into argument-restricted programs.

Introduction

Grounding is a widely used approach that allows us to identify the Herbrand¹ models of a first-order signature with the models of the corresponding propositional signature. For instance, assuming the signature $\{p/1, q/1, a, b\}$, grounding turns

$$p(a) \wedge q(b) \wedge \forall x(p(x) \rightarrow q(x)) \quad (1)$$

into

$$p(a) \wedge q(b) \wedge (p(a) \rightarrow q(a)) \wedge (p(b) \rightarrow q(b)) \quad (2)$$

by replacing the universal quantification with multiple conjunctions in which x ranges over the Herbrand universe $\{a, b\}$. The Herbrand models of (1) are the same as the Herbrand models of (2). Containing no variables, formula (2) can also be viewed as a propositional formula of the signature $\{p(a), p(b), q(a), q(b)\}$ (i.e., “ $p(a)$ ” is an atom in propositional logic). The (propositional) models of the propositional signature can be identified with the Herbrand models of the first-order signature.

However, this approach does not work in the presence of arbitrary function constants. For instance, one cannot turn

$$p(a) \wedge q(b) \wedge \forall x(p(x) \rightarrow q(f(x))) \quad (3)$$

into multiple conjunctions over the Herbrand universe because the universe is infinite:

¹Recall that an *Herbrand interpretation* of a signature σ (containing at least one object constant) is an interpretation of σ such that its universe is the set of all ground terms of σ , and every ground term represents itself. An Herbrand interpretation can be identified with the set of ground atoms (not containing equality) to which it assigns the value *true*.

$$\{a, b, f(a), f(b), f(f(a)), f(f(b)), \dots\}.$$

On the other hand, in answer set programming, where rules are used instead of formulas, function constants are allowed to some degree. In order to ensure finite grounding, answer set solvers impose syntactic conditions on the input languages, such as ω -restricted (Syrjänen 2001; Syrjänen 2004), λ -restricted (Gebser *et al.* 2007), finite ground, or finite domain programs (Calimeri *et al.* 2008). Recently Lierler and Lifschitz [2009] defined the most general decidable class of finitely groundable programs known so far, called argument-restricted programs.

In this paper, we extend the notion of an argument-restricted program to the syntax of an arbitrary formula, whose stable models are defined in (Ferraris *et al.* 2007; Ferraris *et al.* 2010). Unlike in first-order logic, argument-restricted formulas under the stable model semantics can be grounded even in the presence of function constants. For instance, formula (3) is argument-restricted, and the results of this paper tells that the Herbrand *stable* models of formula (3) can be identified with the Herbrand *stable* models of

$$p(a) \wedge q(b) \wedge (p(a) \rightarrow q(f(a))) \wedge (p(b) \rightarrow q(f(b))).$$

It turns out that the notion of an argument-restricted formula can also be viewed as an extension of semi-safety, which is defined in (Lee *et al.* 2009; Cabalar *et al.* 2009) for formulas containing no function constants other than object constants.

We apply the notion of an argument-restricted formula to ensure finite groundability of programs that contain aggregates, whose semantics are as given in (Lee *et al.* 2008) and (Faber *et al.* 2004). Lee *et al.* [2008] showed that the meaning of choice rules and counting aggregates in answer set programming can be understood by identifying these constructs as shorthand for first-order formulas under the stable model semantics. For instance, the “FOL-representation” of the RASPL-1 rule

$$s \leftarrow \{x : p(x)\} 1 \quad (4)$$

is the formula

$$\neg \exists xy(p(x) \wedge p(y) \wedge x \neq y) \rightarrow s. \quad (5)$$

The counterpart of (4) in the language defined in (Faber *et al.* 2004) is

$$s \leftarrow \#count\{x : p(x)\} \leq 1. \quad (6)$$

In that language, the meaning of $\#count\{x : p(x)\} \leq 1$ is understood by instantiating the expression $\{x : p(x)\}$ over the Herbrand universe, and then applying an interpretation on the ground set. Obviously this process results in an infinite ground set once the program contains function constants of positive arity as in the following:

$$\begin{aligned} p(f(x)) &\leftarrow q(x) \\ q(x) &\leftarrow p(x), r(x) \\ p(a) &\quad r(a) \quad r(f(a)) \\ s &\leftarrow \#count\{x : p(x)\} \leq 1 \end{aligned} \quad (7)$$

This program cannot be handled by DLV, which implements a large subset of the language defined in (Faber *et al.* 2004). On the other hand, the recent system DLV-COMPLEX, an extension of DLV with function constants, returns one answer set $\{p(a), p(f(a)), p(f(f(a))), q(a), q(f(a)), r(a), r(f(a))\}$ for the program above. While the system is based on the theory from (Calimeri *et al.* 2008) that ensures finite groundability of an answer set program, that paper does not consider aggregates. To the best of our knowledge, no justification has been provided regarding finite groundability of programs like (7).

We also show that under a certain syntactic condition, argument-restricted formulas can be turned into argument-restricted programs. This may be of interest in applications of system F2LP (Lee and Palla 2009), which computes answer sets of arbitrary formulas by reducing formulas to logic programs.

Preliminaries

Review: Argument-Restricted Programs

We distinguish between atoms and atomic formulas as follows: an *atom* of a signature σ is an n -ary predicate constant followed by a list of n terms; *atomic formulas* of σ are atoms of σ , equalities between terms of σ , and the 0-place connective \perp .

A *program* considered in (Lierler and Lifschitz 2009) is a finite list of *rules* of the form

$$A_1; \dots; A_k \leftarrow A_{k+1}, \dots, A_m, \text{not } A_{m+1}, \dots, \text{not } A_n, \quad (8)$$

($0 \leq k \leq m \leq n$), where each A_i is an atom, possibly containing function constants of positive arity. A program is called *safe* if every variable occurring in a rule of Π occurs also in the positive body of that rule.

We follow the terminology and the notations used in (Lierler and Lifschitz 2009). Given an atom $p(t_1, \dots, t_n)$, expression $p(t_1, \dots, t_n)^0$ represents its predicate constant p , and $p(t_1, \dots, t_n)^i$, where $1 \leq i \leq n$, represents its argument term t_i . As in (Calimeri *et al.* 2008), an *argument* is an expression of the form $p[i]$, where i is one of the argument positions $1, \dots, n$. The *depth* of a variable x in a term t that contains x , denoted by $d(x, t)$, is defined recursively as follows:

$$d(x, t) = \begin{cases} 0, & \text{if } t \text{ is } x; \\ 1 + \max(\{d(x, t_i) \mid t_i \text{ contains } x\}), & \text{if } t \text{ is } f(t_1, \dots, t_n). \end{cases}$$

An *argument ranking* for Π is a function α from arguments to integers such that, for every rule R of Π , every

atom A occurring in the head of R , and every variable x occurring in an argument term A^i , the positive body of R contains an atom B such that x occurs in an argument term B^j satisfying

$$\alpha(A^0[i]) - d(x, A^i) \geq \alpha(B^0[j]) - d(x, B^j). \quad (9)$$

A program is called *argument-restricted* if it is safe and has an argument ranking.

Review: Answer Sets of First-Order Formulas

This review follows (Ferraris *et al.* 2010), a journal version of (Ferraris *et al.* 2007), which is based on the *stable model operator with the intensional predicates* \mathbf{p} , denoted by $\text{SM}_{\mathbf{p}}$. Let \mathbf{p} be a list of distinct predicate constants p_1, \dots, p_n , and let \mathbf{u} be a list of distinct predicate variables u_1, \dots, u_n of the same length as \mathbf{p} . By $\mathbf{u} = \mathbf{p}$ we denote the conjunction of the formulas $\forall \mathbf{x}(u_i(\mathbf{x}) \leftrightarrow p_i(\mathbf{x}))$, where \mathbf{x} is a list of distinct object variables of the same arity as the length of p_i , for all $i = 1, \dots, n$. By $\mathbf{u} \leq \mathbf{p}$ we denote the conjunction of the formulas $\forall \mathbf{x}(u_i(\mathbf{x}) \rightarrow p_i(\mathbf{x}))$ for all $i = 1, \dots, n$, and $\mathbf{u} < \mathbf{p}$ stands for $(\mathbf{u} \leq \mathbf{p}) \wedge \neg(\mathbf{u} = \mathbf{p})$.

For any first-order formula F , $\text{SM}_{\mathbf{p}}[F]$ stands for the second-order formula

$$F \wedge \neg \exists \mathbf{u}((\mathbf{u} < \mathbf{p}) \wedge F^*(\mathbf{u})), \quad (10)$$

where $F^*(\mathbf{u})$ is defined recursively:²

- $p_i(\mathbf{t})^* = u_i(\mathbf{t})$ for any tuple \mathbf{t} of terms;
- $F^* = F$ for any atomic formula F that does not contain members of \mathbf{p} ;³
- $(F \wedge G)^* = F^* \wedge G^*$; $(F \vee G)^* = F^* \vee G^*$;
- $(F \rightarrow G)^* = (F^* \rightarrow G^*) \wedge (F \rightarrow G)$;
- $(\forall x F)^* = \forall x F^*$; $(\exists x F)^* = \exists x F^*$.

For any sentence F , a \mathbf{p} -*stable model* of F is an interpretation of the underlying signature that satisfies $\text{SM}_{\mathbf{p}}[F]$. In this paper we will often omit the subscript \mathbf{p} when it is the list of all predicate constants occurring in F .

The answer sets are defined as follows. By $\sigma(F)$ we denote the signature consisting of object, function and predicate constants occurring in F . For any sentence F that contains at least one object constant, we define the answer sets of F as the Herbrand interpretations of $\sigma(F)$ that satisfy $\text{SM}[F]$. Ferraris *et al.* [2010] shows that this definition of an answer set coincides with the usual definition of an answer set, which is given in terms of grounding and reduct when F can be identified with a logic program.

Semi-Safety and Safety in the Absence of Function Symbols

The concept of safety was extended to first-order formulas in (Lee *et al.* 2008) and was further generalized in (Cabalar *et al.* 2009), but the work there does not consider function constants of positive arity. In this review, we assume that the signature contains no function constants of positive arity.

To every quantifier-free formula F , we assign a set $\text{RV}(F)$ of its *restricted variables* as follows:

- For an atomic formula F ,

²We treat $\neg F$ as shorthand for $F \rightarrow \perp$.

³This includes the case when F is \perp .

- if F is an equality between two variables, then $RV(F) = \emptyset$;
- otherwise, $RV(F)$ is the set of all variables occurring in F ;
- $RV(\perp) = \emptyset$;
- $RV(F \wedge G) = RV(F) \cup RV(G)$;
- $RV(F \vee G) = RV(F) \cap RV(G)$;
- $RV(F \rightarrow G) = \emptyset$.

Recall that the occurrence of one formula in another is called *positive* if the number of implications containing that occurrence in the antecedent is even, and *negative* otherwise. We say that an occurrence of a subformula or a variable in a formula F is *strictly positive* if that occurrence is not in the antecedent of any implication. For example, in (5), the occurrences of $p(x)$ and $p(y)$ are positive, but not strictly positive; the occurrence of s is strictly positive.

Consider a sentence F in prenex form:

$$Q_1 x_1 \cdots Q_n x_n M \quad (11)$$

(each Q_i is \forall or \exists ; x_1, \dots, x_n are distinct variables; the matrix M is quantifier-free). It is shown that every first-order formula under the stable model semantics can be turned into prenex form (Lee and Palla 2007, Theorem 2). We say that F is *semi-safe* if every strictly positive occurrence of every variable x_i belongs to a subformula $G \rightarrow H$ where x_i belongs to $RV(G)$. If a sentence has no strictly positive occurrence of a variable, it is clearly semi-safe. The FOL-representation of (8) is semi-safe if every variable occurring in the head of a rule occurs also in the positive body of that rule. Thus the notion of semi-safety, applied to a logic program, is weaker than the notion of safety.

Proposition 1 below shows that all stable models of a semi-safe sentence have the small predicate property: the relation represented by any of its predicate constants p_i can hold for a tuple of arguments only if each member of the tuple is represented by an object constant occurring in F . The small predicate property can be expressed by the conjunction of the sentences

$$\forall v_1 \dots v_n \left(p(v_1, \dots, v_n) \rightarrow \bigwedge_{i=1, \dots, n} \bigvee_{\substack{c \text{ is an object constant} \\ \text{in } \sigma(F)}} v_i = c \right)$$

for all predicate constants p occurring in F , where v_1, \dots, v_n are distinct variables. We denote this sentence by SPP_F .

Proposition 1 (Lee et al. 2009) *For any semi-safe sentence F , formula $SM[F]$ entails SPP_F .*

For instance, in application to the prenex form of (1), which is semi-safe, this proposition asserts that $SM[(1)]$ entails

$$\forall x (p(x) \rightarrow x = a \vee x = b) \wedge \forall x (q(x) \rightarrow x = a \vee x = b).$$

Cabalar et al. [2009] generalized the definition of safety from (Lee et al. 2008b) as follows. First they define the following transformations:

- $\neg \perp \mapsto \top$, $\neg \top \mapsto \perp$,
- $\perp \wedge F \mapsto \perp$, $F \wedge \perp \mapsto \perp$, $\top \wedge F \mapsto F$, $F \wedge \top \mapsto F$,
- $\perp \vee F \mapsto F$, $F \vee \perp \mapsto F$, $\top \vee F \mapsto \top$, $F \vee \top \mapsto \top$,

- $\perp \rightarrow F \mapsto \top$, $F \rightarrow \top \mapsto \top$, $\top \rightarrow F \mapsto F$.

We say that a variable x is *positively weakly restricted* in a formula G if the formula obtained from G by

- first replacing every atomic formula A in it such that x belongs to $RV(A)$ by \perp ,
- and then applying the transformations above

is \top . Similarly, we say that x is *negatively weakly restricted* in G if the formula obtained from G by the same procedure is \perp .

A semi-safe sentence (11) is called *safe* if, for every occurrence of a variable x_i in (11),

- (a) if Q_i is \forall , then the occurrence belongs to
 - a positive subformula of (11) in which x_i is positively weakly restricted, or
 - a negative subformula of (11) in which x_i is negatively weakly restricted;
- (b) if Q_i is \exists , then the occurrence belongs to
 - a negative subformula of (11) in which x_i is positively weakly restricted, or
 - a positive subformula of (11) in which x_i is negatively weakly restricted.

This definition generalizes the definition of safety for logic programs. The prenex form of (5)

$$\exists x y (\neg(p(x) \wedge p(y) \wedge x \neq y) \rightarrow s) \quad (12)$$

is safe because the antecedent of the implication is a negative subformula in which both x and y are positively weakly restricted in it. (Or $p(x) \wedge p(y) \wedge x \neq y$ can be taken as a positive subformula in which x and y are negatively weakly restricted.) Formula $\exists x \forall y ((p(x) \rightarrow q(y)) \rightarrow r)$ is safe because, for x , $p(x)$ can be taken as a positive subformula, and, for y , $q(y)$ can be taken as a negative subformula. Formula $\exists x (\neg p(x) \rightarrow q)$ is safe, while $\forall x (\neg p(x) \rightarrow q)$ is semi-safe, but not safe.

For any sentence F in prenex form and the set c of all object constants in $\sigma(F)$, the variable-free formula $\text{Ground}[F]$ is defined as follows. If F is quantifier-free then $\text{Ground}[F] = F$. Otherwise,

$$\text{Ground}[\forall x F(x)] = \bigwedge_{c \in c} \text{Ground}[F(c)],$$

$$\text{Ground}[\exists x F(x)] = \bigvee_{c \in c} \text{Ground}[F(c)].$$

Proposition 2 (Lee et al. 2009; Cabalar et al. 2009) *For any safe sentence F , $SM[F]$ is equivalent to $SM[\text{Ground}[F]]$.*

Argument-Restricted Formulas

We extend the notion of argument-restricted programs to arbitrary formulas as follows. We visualize any term t as a rooted tree, with object constants and variables as leaves, and function constants of positive arity as internal nodes. The *height* $h(t)$ of a term t is the height of the corresponding tree. Let F be a quantifier-free formula (possibly containing function constants of positive arity) and let α be a function from arguments to integers. Function $lb_\alpha(x, F)$ maps a variable and a formula to an integer or ω , defined as follows:

- $lb_\alpha(x, p(t_1, \dots, t_n)) = \min(\{\alpha(p[i]) - d(x, t_i) \mid t_i \text{ contains } x\} \cup \{\omega\})$;
- When F is $t_1 = t_2$, if t_1 is x and t_2 is a ground term, then $lb_\alpha(x, F)$ is $h(t_2)$; Similar if t_2 is x and t_1 is a ground term; otherwise $lb_\alpha(x, F)$ is ω ;
- $lb_\alpha(x, \perp) = \omega$;
- $lb_\alpha(x, G \wedge H) = \min(\{lb_\alpha(x, G), lb_\alpha(x, H)\})$;
- $lb_\alpha(x, G \vee H) = \max(\{lb_\alpha(x, G), lb_\alpha(x, H)\})$;
- $lb_\alpha(x, G \rightarrow H) = \omega$.

An *argument ranking* for F is a function α from arguments to integers such that, for every atom A occurring strictly positively in F and every variable x occurring in an argument term A^i , the occurrence of A is in a subformula $G \rightarrow H$ such that

$$\alpha(A^0[i]) - d(x, A^i) \geq lb_\alpha(x, G). \quad (13)$$

We say that the sentence (11) is *argument-restricted* if M has an argument ranking.

For example, $\forall x(p(a) \wedge (p(f(x)) \vee p(x) \rightarrow p(f(x))))$ is not argument-restricted because for any function α ,

$$\begin{aligned} \alpha(p[1]) - d(x, f(x)) &= \alpha(p[1]) - 1 \\ &< lb_\alpha(x, p(f(x)) \vee p(x)) = \alpha(p[1]). \end{aligned}$$

Indeed, the sentence has an infinite answer set: $\{p(a), p(f(a)), \dots\}$. On the other hand,

$$\forall x(p(a, f(a)) \wedge (p(x, f(x)) \vee p(f(x), x) \rightarrow p(x, f(x)))) \quad (14)$$

is argument-restricted. Take $\alpha(p[1]) = \alpha(p[2]) = 1$.

$$\forall xy(p(x) \wedge y = f(x) \rightarrow p(y)) \quad (15)$$

is not argument-restricted, while

$$\forall xy(p(x) \wedge q(y, f(x)) \rightarrow p(y)) \quad (16)$$

is argument-restricted.

The small predicate property can be extended to argument-restricted sentences as follows. By $gt_\sigma(m)$ we denote the set of ground terms of height up to m which can be constructed from object constants and function constants from σ . Let \mathbf{p} be the list of all predicate constants occurring in F . By $SPP_F(\alpha)$ (the “small predicate property”) we denote the conjunction of the sentences

$$\forall v_1 \dots v_n \left(p(v_1, \dots, v_n) \rightarrow \bigwedge_{i=1, \dots, n} \bigvee_{c \in gt_{\sigma(F)}(\alpha(p[i]))} (v_i = c) \right)$$

for all predicate constants p in \mathbf{p} .

We consider the following axiom set *relaxed unique name assumption* for signature σ .

$$\begin{aligned} \forall x_1 \dots x_n y_1 \dots y_n (f(x_1, \dots, x_n) = f(y_1, \dots, y_n) \\ \rightarrow (x_1 = y_1 \wedge \dots \wedge x_n = y_n)) \end{aligned} \quad (17)$$

for all function constants f of arity > 0 ;

$$\forall x_1 \dots x_m y_1 \dots y_n (f(x_1, \dots, x_m) \neq g(y_1, \dots, y_n)) \quad (18)$$

for all pairs of distinct function constants f, g except for the pair in which both f and g have arity 0. In other words, the assumption is slightly weaker than the standard unique

name assumption in that object constants are not assumed to be distinct from each other.

We say that an argument ranking α is *strict* if $\alpha(A^0[i]) \geq h(A^i)$ for every strictly positive occurrence of an atom A and any of its argument position i . The following theorem tells that, for any argument-restricted sentence, the extent of a predicate under a stable model is named by a finite set of ground terms.

Theorem 1 *For any argument-restricted sentence F and any strict argument ranking α for its matrix, under the relaxed unique name assumption, formula $SM[F]$ entails $SPP_F(\alpha)$.*⁴

Theorem 1 becomes incorrect if we drop the relaxed unique name assumption. For example, take

$$F = p(0) \wedge \forall x(p(f(x)) \rightarrow p(x)),$$

which is argument-restricted (take $\alpha(p[1]) = 1$). However, consider a model I of $SM[F]$ such that the universe is the set of integers, $f^I(m) = m + 1$, $p^I = \{0, -1, -2, \dots\}$. The model I satisfies $SM[F]$ but not $SPP_F(\alpha)$.

Note that the notion of an argument-restricted program in (Lierler and Lifschitz 2009) requires that the program be safe, while the notion of an argument-restricted formula above does not.

Proposition 3 *Let Π be a program and F a prenex form of the FOL representation of Π . Π is argument-restricted in the sense of (Lierler and Lifschitz 2009) iff F is argument-restricted (in our sense) and safe.*

The notion of an argument-restricted formula is an extension of semi-safety by allowing function constants.

Proposition 4 *For any sentence F in prenex form that contains no function constants of positive arity, F is semi-safe according to (Lee et al. 2009) iff it is argument-restricted.*

However, the small predicate property is not enough to ensure the domain independence of answer sets. For instance,

$$\forall x(\neg q(x) \rightarrow p) \wedge q(a)$$

is argument-restricted. The answer set is $\{q(a)\}$ if the Herbrand universe is $\{a\}$; the answer set is $\{p, q(a)\}$ if the Herbrand universe is any proper superset of $\{a\}$ (e.g., $\{a, b\}$). In order to define grounding we need to further assume that the formula is safe.

The safety condition defined earlier for formulas without function constants can be extended to allow function constants. The changes required are

- in the definitions of positively (negatively, resp.) weakly restricted variables, to replace “ x belongs to $RV(A)$ ” by “ $lb_\alpha(x, A) \neq \omega$ ”.
- replace “semi-safe sentence” with “argument-restricted sentence.”

For any safe sentence F (in prenex form; possibly containing function constants of positive arity), and any strict argument ranking α for F , the result of its grounding is

⁴The proof of the theorem is available in a longer version at <http://peace.eas.asu.edu/joolee/papers/dcfg.pdf>.

the ground formula $\text{Ground}_\alpha[F]$ defined as follows. If F is quantifier-free then $\text{Ground}_\alpha[F] = F$. Otherwise

$$\text{Ground}_\alpha[\forall x F(x)] = \bigwedge_{c \in \text{gt}_\sigma(F)(m)} \text{Ground}_\alpha[F(c)],$$

$$\text{Ground}_\alpha[\exists x F(x)] = \bigvee_{c \in \text{gt}_\sigma(F)(m)} \text{Ground}_\alpha[F(c)],$$

where m is

$$\max(\{lb_\alpha(x, A) \mid A \text{ is an atomic formula occurring in } F(x) \text{ such that } lb_\alpha(x, A) \neq \omega\} \cup \{0\}).$$

For example, for $\alpha(p[1]) = \alpha(p[2]) = 1$, $\text{Ground}_\alpha[(14)]$ is $p(a, f(a)) \wedge (p(a, f(a)) \vee p(f(a), a) \rightarrow p(a, f(a)))$.

Theorem 2 *Let F be a safe sentence that contains at least one object constant, let α be a strict argument ranking for F , and let σ be any signature obtained by adding any function constants (including object constants) to $\sigma(F)$. An Herbrand interpretation of σ satisfies $\text{SM}[F]$ iff it is an answer set of $\text{Ground}_\alpha[F]$.*

Deciding Argument-Restricted Formulas

The algorithm described in (Lierler and Lifschitz 2009) can be extended to argument-restricted formulas in a straightforward way. Let F be a formula (11). Similar to (Lierler and Lifschitz 2009), we define the operator Ω on the set U of functions from arguments to nonnegative integers by the formula

$$\Omega(\alpha)(p[i]) = \max(\{\min(\{d(x, A^i) + lb_\alpha(x, G) \mid G \rightarrow H\} \cup \{\omega\}) \mid A, x\} \cup \{0\})$$

where

- A and x range over atoms and variables such that A occurs strictly positively in F , $A^0 = p$, and x occurs in A^i ;
- given such A and x , formula $G \rightarrow H$ ranges over subformulas of F such that the occurrence of A is in $G \rightarrow H$.

It is easy to check that $\alpha \in U$ is an argument ranking for the matrix of F iff $\alpha \geq \Omega(\alpha)$.

The operator Ω is monotone, so that the same argument in (Lierler and Lifschitz 2009) applies to this extension. If F is argument-restricted, then the set of nonnegative argument rankings has the least element $\alpha_{\min} = \Omega^i(\mathbf{0})$ for the smallest i such that $\Omega^{i+1}(\mathbf{0}) = \Omega^i(\mathbf{0})$.⁵ On the other hand, for any argument-restricted formula F , all values of α_{\min} do not exceed the product M of two numbers: the total number of arguments and the largest of the numbers $d(x, t)$ for the terms t occurring strictly positively in F and for the variables x occurring in t .

We can use the polynomial time algorithm given in (Lierler and Lifschitz 2009) with the extended definition of Ω to determine whether F is argument-restricted by iterating on $\mathbf{0}$ until

- $\Omega^{i+1}(\mathbf{0}) = \Omega^i(\mathbf{0})$ (α_{\min} is found), or
- one of the values of $\Omega^i(\mathbf{0})$ exceeds M (F is not argument-restricted).

⁵ $\mathbf{0}$ denotes a function that maps every argument positions to 0.

Applying to RASPL-1 programs

(Lee *et al.* 2008) defines the semantics of a program with choice and counting in terms of reduction to the general theory of stable models. That paper assumes that the signature does not contain function constants of positive arity, but here we remove that restriction.

A *cardinality expression* is an expression of the form

$$b \{ \mathbf{x} : F(\mathbf{x}) \} \quad (19)$$

where b is a positive integer (“the bound”), \mathbf{x} is a list of variables (possibly empty), and $F(\mathbf{x})$ is a conjunction of literals (possibly containing variables other than \mathbf{x}). This expression reads: there are at least b values of \mathbf{x} such that $F(\mathbf{x})$.

A *RASPL-1 rule* is an expression of the form

$$A_1 ; \dots ; A_l \leftarrow E_1, \dots, E_m, \text{not } E_{m+1}, \dots, \text{not } E_n \quad (20)$$

($l \geq 0; n \geq m \geq 0$), where each A_i is an atom, and each E_i is a cardinality expression. A *RASPL-1 program* is a finite list of rules.

The semantics of a RASPL-1 program is defined by turning it into a first-order sentence (*FOL-representation*) and applying the definition of an answer set for a first-order sentence. The FOL-representation of cardinality expression (19) is the formula

$$\exists \mathbf{x}^1 \dots \mathbf{x}^b \left[\bigwedge_{1 \leq i \leq b} F(\mathbf{x}^i) \wedge \bigwedge_{1 \leq i < j \leq b} \neg(\mathbf{x}^i = \mathbf{x}^j) \right] \quad (21)$$

where $\mathbf{x}^1, \dots, \mathbf{x}^b$ are lists of new variables of the same length as \mathbf{x} . An expression $1\{A\}$, where A is an atom, can be identified with A .

The FOL-representation of a RASPL-1 rule $\text{Head} \leftarrow \text{Body}$ is the universal closure of the implication $\text{Body} \rightarrow \text{Head}$ with each cardinality expression in Body replaced by its FOL-representation. The FOL-representation of a RASPL-1 program is the conjunction of the FOL-representations of its rules.

For any RASPL-1 program Π containing at least one object constant, the *answer sets* of Π are defined as the answer sets of the FOL-representation of Π , as defined earlier.

A variable v is *free* in a rule (20) if

- v occurs in the head $A_1 ; \dots ; A_l$ of the rule, or
- the body $E_1, \dots, \text{not } E_n$ of the rule contains a cardinality expression $b\{\mathbf{x} : F(\mathbf{x})\}$ such that v occurs in F and does not belong to \mathbf{x} .

In the following we assume that every free variable is distinct from \mathbf{x} occurring in (19) by possibly renaming the variables.

An *argument ranking* for Π is a function α from arguments to integers such that, for every rule R of Π , every atom A occurring in the head of R , and every variable v occurring in an argument term A^i , the positive body of R contains a cardinality expression $b\{\mathbf{x} : F(\mathbf{x})\}$ such that

$$\alpha(A^0[i]) - d(v, A^i) \geq lb_\alpha(v, F(\mathbf{x})).$$

A program is called *argument-restricted* if it has an argument ranking.

The safety condition for RASPL-1 programs defined in (Lee *et al.* 2008) can be extended to allow function con-

stands as follows. Given an argument ranking α , we say that a cardinality expression $b\{x : F(x)\}$ is *allowed* if, for every member x of x , $lb_\alpha(x, F) \neq \omega$. For instance, $2\{x : p(f(x), y)\}$ is allowed; $2\{x : \text{not } p(f(x), y)\}$ is not allowed.

A RASPL-1 program is *safe* if it has an argument ranking α and, for each rule (20) in it,

- each cardinality expression in its body is allowed, and
- for each of its free variables v , there is a cardinality expression $b\{x : F(x)\}$ among E_1, \dots, E_m such that $lb_\alpha(v, F) \neq \omega$.

The following proposition follows from Proposition 2.

Proposition 5 *Let Π be a safe RASPL-1 program (allowing function constants of positive arity) containing at least one object constant, let F be its FOL-representation, and let α be a strict argument ranking for F . For any signature σ obtained by adding any function constants to $\sigma(F)$, an Herbrand interpretation of σ satisfies $SM[F]$ iff it is an answer set of $\text{Ground}_\alpha[F]$.*

For example, program (7) (assuming the last rule is written in the syntax of RASPL-1) is safe, and can be equivalently rewritten as a ground formula under the stable model semantics.

Applying to FLP-Aggregates

An FLP-aggregate expression is of the form

$$\text{OP}\{\{x : F(x)\}\} \succeq b \quad (22)$$

where

- OP is a symbol for an *aggregate function*, such as SUM, COUNT;
- x is a nonempty list of distinct object variables;
- $F(x)$ is a conjunction of literals;
- \succeq is a symbol for a binary relation over integers, such as $\leq, \geq, <, >, =, \neq$;
- b is an integer constant.

An FLP-rule is an expression of the form

$$A_1; \dots; A_l \leftarrow E_1, \dots, E_m, \text{not } E_{m+1}, \dots, \text{not } E_n \quad (23)$$

($l \geq 0$; $n \geq m \geq 0$), where each A_i is an atom and each E_i is an atom or an FLP-aggregate expression. An FLP-program is a finite list of rules.

The definition of argument ranking for an FLP program is same as that for a program whose rules have the form (8). The definitions of “allowed expressions” and “free variables” for FLP programs are defined similar to those for RASPL-1 programs.

According to (Faber *et al.* 2004), a rule (23) is *safe* if

- each FLP-aggregate expression in its body is allowed,
- for each of its free variables v , there is an atom E_i ($i = 1, \dots, m$) that contains v .

Again we assume that every free variable is distinct from x in (22) by possibly renaming the variables. For any argument-restricted FLP-program Π and any strict argument ranking α for Π , ground program $\text{Ground}_\alpha[\Pi]$ is defined as follows:

- replace every free occurrence of a variable x with a ground term from $gt_{\sigma(F)}(m)$ where m is

$$\max(\{lb_\alpha(x, A) \mid A \text{ is an atomic formula occurring in } \Pi \text{ such that } lb_\alpha(x, A) \neq \omega\} \cup \{0\});$$

- replace each occurrence of $\text{OP}\{x : F(x)\} \succeq b$ in the resulting program with $\text{OP}(S) \succeq b$ where S is the set of all ground instantiations of $F(x)$ w.r.t. $gt_{\sigma(F)}(m)$ where m is

$$\max(\{lb_\alpha(x, A) \mid x \text{ is in } x \text{ and } A \text{ is an atomic formula occurring in } F(x) \text{ such that } lb_\alpha(x, A) \neq \omega\} \cup \{0\}).$$

In the following proposition, the definition of an answer set refers to the one in (Faber *et al.* 2004). By $\sigma(\Pi)$ we denote the signature consisting of object, function and predicate constants occurring in Π .

Proposition 6 *Let Π be a safe FLP program containing at least one object constant, and let α be a strict argument ranking for Π . For any signature σ obtained by adding function constants to $\sigma(\Pi)$, a ground set of atoms is an answer set of Π iff it is an answer set of ground FLP program $\text{Ground}_\alpha[\Pi]$.*

For example, program (7) is safe, and grounding the last rule of (7) yields

$$s \leftarrow \#count\{p(a), p(f(a)), p(f(f(a))), q(a), q(f(a)), r(a), r(f(a))\} \leq 1$$

with $\alpha(p[1]) = 2, \alpha(q[1]) = 1, \alpha(r[1]) = 1$.

The following is another safe program in the syntax of the DLV-COMPLEX language.

```
string(cons(a, cons(b, cons(a, cons(c, nil))))) .
letter(a). letter(b). letter(c). letter(d).
tail(X) :- string(X).
tail(Y) :- tail(cons(X,Y)), letter(X).
containsTwo(X) :- #count{Y: tail(cons(X,Y))}>=2,
                  letter(X).
```

The program has only one answer set. For the string “[a,b,a,c]”, $\text{tail}(X)$ belongs to the answer set if X is the tail of the string. $\text{containsTwo}(X)$ belongs to the answer set if letter X occurs at least twice in the string. The answer set returned by DLV-COMPLEX contains $\text{containsTwo}(a)$.

Allowing Extensional Predicates

We defined the notion of an argument-restricted formula only in the case when the intensional predicates \mathbf{p} in $SM_{\mathbf{p}}[F]$ are all the predicates occurring in F . The notion can be extended to the general case, in which intensional predicates \mathbf{p} are any subset of the predicates in the signature, by only modifying the first clause of the definition of $lb_\alpha(x, F)$ as

$$lb_\alpha(x, p(t_1, \dots, t_n)) = \begin{cases} \min(\{\alpha(p[i]) - d(x, t_i) \mid t_i \text{ contains } x\} \cup \{\omega\}) & \text{if } p \text{ is intensional;} \\ \omega & \text{otherwise.} \end{cases}$$

For example, similar to (15), formula (16) becomes non-argument-restricted if q is regarded as extensional.

The definition of safety in the general case remains the same as in the case when all predicate constants are regarded intensional.

Theorems 1 and 2 can be extended to allow extensional predicates by allowing \mathbf{p} to be any subset of the predicate constants in the signature, and by replacing $\text{SM}[F]$ in the statements with $\text{SM}_{\mathbf{p}}[F]$.

Reducing Argument-Restricted Formulas to Argument-Restricted Programs

System F2LP (Lee and Palla 2009) turns a first-order formula into a program whose rules have the form

$$A_1 ; \dots ; A_k \leftarrow A_{k+1}, \dots, A_m, \text{not } A_{m+1}, \dots, \text{not } A_n, \text{not not } A_{n+1}, \dots, \text{not not } A_p \quad (24)$$

($0 \leq k \leq m \leq n \leq p$), where each A_i is an atomic formula. The definition of an argument-restricted program from (Lierler and Lifschitz 2009) can be extended to this syntax as follows: A program whose rules have the form (24) is called *argument-restricted* if its FOL-representation is argument-restricted.

Is it possible to turn argument-restricted formulas into such argument-restricted programs? This is an important question in applications of system F2LP. A part of the transformation that is implemented in F2LP follows the one from (Cabalar *et al.* 2005) in order to turn a quantifier-free formula into a set of rules. However, this transformation does not preserve argument-restrictedness: among the transformation rules defined in (Cabalar *et al.* 2005), only one rule (L5) does not preserve it. For instance, given

$$\begin{aligned} &\forall x((p(x) \rightarrow r(f(x))) \\ &\wedge ((p(x) \rightarrow q(x)) \wedge r(x) \rightarrow s(x))), \end{aligned} \quad (25)$$

F2LP turns it into the following program

$$\begin{aligned} r(f(x)) &\leftarrow p(x) \\ s(x) &\leftarrow \text{not } p(x), r(x) \\ s(x) &\leftarrow q(x), r(x) \\ p(x) ; s(x) &\leftarrow r(x), \text{not not } q(x), \end{aligned}$$

which is finitely groundable, but not argument-restricted.

Such reduction is possible under a certain syntactic condition as described in the following. We call a formula *singular* if every subformula $G \rightarrow H$ that occurs in the antecedent of another implication is contained in a formula that has no strictly positive occurrence of predicate constants. Formula (25) is not singular because of $p(x) \rightarrow q(x)$.

Proposition 7 *Every singular argument-restricted quantifier-free formula is strongly equivalent to an argument-restricted program.*

It remains an open question if the restriction that formula be singular can be removed from the statement.

Conclusion

Our contributions are as follows:

- We have presented a decidable class of the general language of stable models that can be instantiated even in the presence of function constants. The notion extends the notion of an argument-restricted program and the notion of a semi-safe sentence.

- Based on this, we presented a syntactic condition under which RASPL-1 programs and FLP programs can be finitely instantiated.
- We showed that under a certain condition, any argument-restricted quantifier-free formula can be turned into an argument-restricted program.

References

- Pedro Cabalar, David Pearce, and Agustin Valverde. Reducing propositional theories in equilibrium logic to logic programs. In *Proceedings of Portuguese Conference on Artificial Intelligence (EPIA)*, pages 4–17, 2005.
- Pedro Cabalar, David Pearce, and Agustin Valverde. A revised concept of safety for general answer set programs. In *Proc. LPNMR*, pages 58–70, 2009.
- Francesco Calimeri, Susanna Cozza, Giovambattista Ianni, and Nicola Leone. Computable functions in ASP: theory and implementation. In *Proc. ICLP*, pages 407–424, 2008.
- Wolfgang Faber, Nicola Leone, and Gerald Pfeifer. Recursive aggregates in disjunctive logic programs: Semantics and complexity. In *Proc. JELIA*, 2004.
- Paolo Ferraris, Joohyung Lee, and Vladimir Lifschitz. A new perspective on stable models. In *Proc. IJCAI*, pages 372–379, 2007.
- Paolo Ferraris, Joohyung Lee, and Vladimir Lifschitz. Stable models and circumscription.⁶ *Artificial Intelligence*, 2010. To appear.
- M. Gebser, T. Schaub, and S. Thiele. Gringo: A new grounder for answer set programming. In *Proc. LPNMR*, pages 266–271, 2007.
- Joohyung Lee and Ravi Palla. Yet another proof of the strong equivalence between propositional theories and logic programs. In *Working Notes of the Workshop on Correspondence and Equivalence for Nonmonotonic Theories*, 2007.
- Joohyung Lee and Ravi Palla. System F2LP – computing answer sets of first-order formulas. In *Proc. LPNMR*, pages 515–521, 2009.
- Joohyung Lee, Vladimir Lifschitz, and Ravi Palla. A reductive semantics for counting and choice in answer set programming. In *Proc. AAAI*, pages 472–479, 2008.
- Joohyung Lee, Vladimir Lifschitz, and Ravi Palla. Safe formulas in the general theory of stable models (preliminary report). In *Proc. ICLP*, pages 672–676, 2008.
- Joohyung Lee, Vladimir Lifschitz, and Ravi Palla. Safe formulas in the general theory of stable models. Unpublished Draft, 2009.
- Yuliya Lierler and Vladimir Lifschitz. One more decidable class of finitely ground programs. In *Proc. ICLP*, pages 489–493, 2009.
- Tommi Syrjänen. Omega-restricted logic programs. In *Proc. LPNMR*, pages 267–279, 2001.
- Tommi Syrjänen. Cardinality constraint programs. In *Proc. JELIA*, pages 187–199, 2004.

⁶<http://peace.eas.asu.edu/joolee/papers/smcirc.pdf>.

Selected Proofs

Proof of Theorem 1

Given an argument ranking α for formula F , \mathbf{e}_α denotes the list of predicate expressions consisting of

$$\lambda v_1, \dots, v_n \left(p(v_1, \dots, v_n) \wedge \bigwedge_{v_i \in \mathbf{v}} \bigvee_{c \in C(\alpha(p[i]))} v_i = c \right)$$

for each p that occurs in F .

In the following three lemmas, let F be a quantifier-free formula and let α be a function from arguments to integers such that $\alpha(A^0[i]) \geq h(A^i)$ for every strictly positive occurrence of a variable x in A^i where A is an atom.

Let F_α be the formula obtained from $F^*(\mathbf{u})$ by replacing all occurrences of $u_i(t_1, \dots, t_n)$ with

$$p_i(t_1, \dots, t_n) \wedge \bigwedge_{t_j \text{ contains } x} \bigvee_{c \in C(\alpha(p_i[j]) - d(x, t_j))} x = c.$$

Lemma 1 *Under the relaxed unique name assumption, F_α is equivalent to $F^*(\mathbf{e}_\alpha)$.*

Proof. By induction. ■

Lemma 2 *Under the relaxed unique name assumption,*

$$F^*(\mathbf{e}_\alpha) \rightarrow \bigwedge_{x: lb_\alpha(x, F) \neq \omega} \bigvee_{c \in C(lb_\alpha(x, F))} x = c$$

is logically valid.

Proof. By induction using Lemma 1. ■

Lemma 3 (Ferraris et al. 2010, Lemma 5) *The formula*

$$(\mathbf{u} \leq \mathbf{p}) \wedge F^*(\mathbf{u}) \rightarrow F$$

is logically valid.

Proof: by induction on F . ■

Given a quantifier-free formula F , we say that an occurrence of a variable x in F is *unsafe* if the occurrence is in an atom that occurs strictly positively in F but not contained in any subformula $G \rightarrow H$ of F such that $lb_\alpha(x, G) \neq \omega$.

Lemma 4 *Assume that every non-unsafe strictly positive occurrence of every variable x is contained in a subformula $G \rightarrow H$ of F that satisfies (13) where i is the argument position of A containing the occurrence. Under the relaxed unique name assumption,*

$$F \wedge \bigwedge_{\substack{\text{the occurrence } \tilde{x} \text{ of } x \text{ is} \\ \text{unsafe in } F \text{ and } \tilde{x} \text{ is in } A^i}} \bigvee_{c \in C(\alpha(A^0[i]) - d(x, A^i))} x = c \rightarrow F^*(\mathbf{e}_\alpha)$$

is logically valid.

Proof. We consider the case when F is $G \rightarrow H$. Assume $G \rightarrow H$,

$$\bigwedge_{\substack{\text{the occurrence } \tilde{x} \text{ of } x \text{ is unsafe in } G \rightarrow H, \\ \tilde{x} \text{ is in } A^i}} \bigvee_{c \in C(\alpha(A^0[i]) - d(x, A^i))} x = c \quad (26)$$

and $G^*(\mathbf{e}_\alpha)$. We get G by Lemma 3. Consequently, H ,

Also from $G^*(\mathbf{e}_\alpha)$, by Lemma 2,

$$\bigwedge_{x: lb_\alpha(x, G) \neq \omega} \bigvee_{c \in C(lb_\alpha(x, G))} x = c. \quad (27)$$

Let \tilde{x} be an unsafe occurrence of x in H . Either \tilde{x} is unsafe or not unsafe in $G \rightarrow H$.

Case 1: \tilde{x} is unsafe in $G \rightarrow H$. By (26)

$$\bigvee_{c \in C(\alpha(A^0[i]) - d(x, A^i))} x = c$$

where \tilde{x} is in t_i of $p(t_1, \dots, t_n)$.

Case 2: \tilde{x} is not unsafe in $G \rightarrow H$, which means that x is in $\{x : lb_\alpha(x, G) \neq \omega\}$. By (27),

$$\bigvee_{c \in C(lb_\alpha(x, G))} x = c.$$

By the condition on α , it follows that

$$\bigvee_{c \in C(\alpha(A^0[i]) - d(x, A^i))} x = c.$$

Consequently,

$$\bigwedge_{\substack{\text{the occurrence } \tilde{x} \text{ of } x \text{ is unsafe in } H, \\ \tilde{x} \text{ is in } A^i}} \bigvee_{c \in C(\alpha(A^0[i]) - d(x, A^i))} x = c.$$

By I.H. on H , we conclude $H^*(\mathbf{e}_\alpha)$. ■

Lemma 5 *For any argument-restricted sentence F with a strict argument ranking α , F entails $F^*(\mathbf{e}_\alpha)$.*

Proof. Immediate from Lemma 4.

Proof of Theorem 1 Let α be an argument ranking for F . We will show that $\text{SM}[F]$ entails U_α . It is sufficient to show that $F \wedge \neg U_\alpha$ entails

$$\exists \mathbf{u} (\mathbf{u} < \mathbf{p} \wedge F^*(\mathbf{u})).$$

Assume F and $\neg U_\alpha$. It is sufficient to show that

$$\mathbf{e}_\alpha < \mathbf{p} \wedge F^*(\mathbf{e}_\alpha). \quad (28)$$

Note that

$$\begin{aligned} \mathbf{e}_\alpha < \mathbf{p} \equiv & \bigwedge_{p \in \mathbf{p}} \left(\forall \mathbf{v} \left((p(\mathbf{v}) \wedge \bigwedge_{v_i \in \mathbf{v}} \bigvee_{c \in C(\alpha(p[i]))} v_i = c) \rightarrow p(\mathbf{v}) \right) \right) \\ & \wedge \neg \bigwedge_{p \in \mathbf{p}} \forall \mathbf{v} \left(p(\mathbf{v}) \rightarrow (p(\mathbf{v}) \wedge \bigwedge_{v_i \in \mathbf{v}} \bigvee_{c \in C(\alpha(p[i]))} v_i = c) \right) \end{aligned} \quad (29)$$

Formula

$$\bigwedge_{p \in \mathbf{p}} \left(\forall \mathbf{v} \left((p(\mathbf{v}) \wedge \bigwedge_{v_i \in \mathbf{v}} \bigvee_{c \in C(\alpha(p[i]))} v_i = c) \rightarrow p(\mathbf{v}) \right) \right)$$

is logically valid. Assume for the sake of contradiction,

$$\bigwedge_{p \in \mathbf{p}} \forall \mathbf{v} \left(p(\mathbf{v}) \rightarrow (p(\mathbf{v}) \wedge \bigwedge_{v_i \in \mathbf{v}} \bigvee_{c \in C(\alpha(p[i]))} v_i = c) \right). \quad (30)$$

Formula (30) entails U_α , which contradicts the assumption $\neg U_\alpha$. Consequently, we conclude $e_\alpha < p$. By Lemma 5, the second conjunctive term of (28) follows from the assumption about F . ■

Proof of Theorem 2

Given the small predicate property ensured by the argument-restricted condition, the proof is almost same as the proof of Proposition 2 from (Lee *et al.* 2009).

Proof of Proposition 6

The proof uses the following lemma.

Lemma 6 *Let F be a quantifier-free formula, α a function that maps arguments to nonnegative integers, and let \mathbf{A} be the set of atomic formulas occurring in F such that $lb_\alpha(x, A) \neq \omega$. If $lb_\alpha(x, F) \neq \omega$, then*

$$F \rightarrow \bigvee_{A \in \mathbf{A}} A$$

is logically valid.

Proof. By induction. ■