

Complexity and compilation of GZ-aggregates in answer set programming

MARIO ALVIANO and NICOLA LEONE

Department of Mathematics and Computer Science, University of Calabria, Italy

submitted 29 April 2015; revised 3 July 2015; accepted 15 July 2015

Abstract

Gelfond and Zhang recently proposed a new stable model semantics based on Vicious Circle Principle in order to improve the interpretation of logic programs with aggregates. The paper focuses on this proposal, and analyzes the complexity of both coherence testing and cautious reasoning under the new semantics. Some surprising results highlight similarities and differences versus mainstream stable model semantics for aggregates. Moreover, the paper reports on the design of compilation techniques for implementing the new semantics on top of existing ASP solvers, which eventually lead to realize a prototype system that allows for experimenting with Gelfond-Zhang's aggregates.

KEYWORDS: answer set programming; aggregates; complexity; compilation.

1 Introduction

Answer set programming (ASP) is a declarative language for knowledge representation and reasoning (Brewka *et al.* 2011). ASP specifications are sets of logic rules, possibly using disjunction and default negation, interpreted according to the stable model semantics (Gelfond and Lifschitz 1988; Gelfond and Lifschitz 1991). The basic language is extended by several constructs to ease the representation of practical knowledge. Aggregate functions are among these extensions (Simons *et al.* 2002; Liu *et al.* 2010; Faber *et al.* 2011; Bartholomew *et al.* 2011), and allow to express properties on sets of atoms declaratively. For example, aggregate functions are often used to enforce *functional dependencies*; a rule of the following form:

$$\perp \leftarrow R'(\bar{X}), \text{COUNT}[\bar{Y} : R(\bar{X}, \bar{Y}, \bar{Z})] \leq 1$$

constrains relation R to satisfy the functional dependency $\bar{X} \rightarrow \bar{Y}$, where $\bar{X} \cup \bar{Y} \cup \bar{Z}$ is the set of attributes of R , and R' is the projection of R on \bar{X} . Aggregate functions are also commonly used in ASP to constrain a nondeterministic guess. For example, in the *knapsack problem* the total weight of the selected items must not exceed a given limit, which can be modeled by the following rule aggregating over a multiset:

$$\perp \leftarrow \text{SUM}[W, O : \text{object}(O, W, C), \text{in}(O)] \leq \text{limit}.$$

Aggregate functions may also ease the representation of logic circuits made of gates of unbounded fan-in (Gelfond and Zhang 2014); the following rule models that the output of an XOR gate is 1 if an odd number of its inputs have value 1:

$$\text{value}(O, 1) \leftarrow \text{xor}(G), \text{output}(G, O), \text{ODD}[I : \text{input}(G, I), \text{value}(I, 1)].$$

Several semantics were proposed for ASP programs with aggregates. Two of them (Ferraris 2011; Faber *et al.* 2011) are implemented in popular ASP solvers (Gebser *et al.* 2012; Faber *et al.* 2008). These two semantics agree for programs without negated aggregates, and are referred in this paper as F-stable model semantics. An alternative semantics, presented at the 30th International Conference on Logic Programming (ICLP'14) by (Gelfond and Zhang 2014) and here referred to as GZ- or G-stable model semantics, is based on the notion of *vicious circle principle*, which essentially asserts that the truth of an atom must be inferred by means of a definition not referring, directly or indirectly, to the truth of the atom itself.

The present paper explores this new semantics, reporting a detailed complexity analysis of *coherence testing* and *cautious reasoning* (Eiter and Gottlob 1995), two of the main computational tasks in ASP. In a nutshell, coherence testing amounts to check the existence of a stable model of an input program, while cautious reasoning consists in checking whether a given atom is true in all stable models of a program.

Concerning coherence testing, membership in Σ_2^P was proved in (Gelfond and Zhang 2014), and Σ_2^P -hardness is proved here already for negation-free programs with a very limited form of aggregate functions, referred to as *monotone* aggregates in the literature. This result is in contrast with F-stable model semantics, for which coherence of negation-free programs with monotone aggregates is guaranteed. Whether this must be considered a strength or a weakness of G-stable models is not the focus of this paper, but we remark here that the increase in complexity also comes with a higher expressive power in this case: aggregates referred to as monotone in the literature allow to simulate integrity constraints and possibly default negation when interpreted according to the semantics by (Gelfond and Zhang 2014). Moreover, there are also many cases in which G-stable models actually decrease the complexity of the reasoning tasks. In fact, while for F-stable model semantics coherence testing is Σ_2^P -hard already for disjunction-free programs, this computational task is proved to be NP-complete for these programs under G-stable model semantics. Finally, P-completeness is proved for programs with monotone aggregates if disjunction and negation are not used, a result compatible with F-stable model semantics. However, also in this case G-stable models allow to simulate integrity constraints, which is not possible with F-stable models.

As for the complexity of cautious reasoning, membership and hardness in the complementary complexity classes are proved for all the analyzed fragments of the language. These complexity results also implicitly characterize the computational complexity of *brave reasoning*, another common reasoning task in ASP which consists in checking whether a given propositional atom is true in some stable model of an input program. In fact, brave reasoning has the same complexity of coherence testing under G-stable model semantics, while this is not necessarily the case for F-stable models. Again, the reason for this discrepancy is the power of G-stable models to

simulate integrity constraints, as this is the additional construct that is commonly used for reducing brave reasoning to coherence testing.

Further results in the paper are two rewriting techniques for compiling programs interpreted according to G-stable semantics into programs interpreted according to F-stable semantics. The first rewriting is simpler and introduces fewer auxiliary symbols, while the second has the advantage of producing programs with *non recursive* aggregates only. Both rewritings are *polynomial, faithful and modular* translation functions (Janhunen 2006), and are implemented in a system prototype. It is publicly available (<http://alviano.net/software/g-stable-models/>) and allows for experimenting with this newly proposed semantics.

2 Background

After defining the syntax of logic programs with aggregates, two semantics are introduced, referred to as F- (Ferraris 2011; Faber *et al.* 2011) and G-stable models (Gelfond and Zhang 2014). It is remarked here, and clarified in Section 5, that the original definitions are properly adapted to better fit the results in this paper.

Syntax. Let **T**,**F** denote the Boolean truth values true and false, respectively. Let \mathcal{U} be a finite set of propositional atoms. An aggregate atom A is a Boolean function whose domain, denoted $dom(A)$, is a subset of \mathcal{U} . A literal is a propositional atom, or a propositional atom preceded by (one or more occurrences of) the negation as failure symbol \sim , or an aggregate atom. A rule r is of the following form:

$$p_1 \vee \cdots \vee p_m \leftarrow l_1, \dots, l_n \quad (1)$$

where p_1, \dots, p_m are propositional atoms, l_1, \dots, l_n are literals, $m \geq 1$ and $n \geq 0$. Set $\{p_1, \dots, p_m\}$ is the head of r , denoted $H(r)$, and set $\{l_1, \dots, l_n\}$ is the body of r , denoted $B(r)$. A program Π is a finite set of rules of the form (1). The set of propositional atoms occurring in Π is denoted $At(\Pi)$.

Semantic notions. An interpretation I is a subset of \mathcal{U} . Let S, S' be sets of interpretations, and C be a set of propositional atoms. Sets S and S' are equivalent in the context C , denoted $S \equiv_C S'$, if $|S| = |S'|$ and $\{I \cap C \mid I \in S\} = \{I \cap C \mid I \in S'\}$. Aggregates are usually classified in three groups (Liu and Truszczyński 2006): an aggregate A is *monotone* if $A(I) = \mathbf{T}$ implies $A(J) = \mathbf{T}$, for all $I \subseteq J \subseteq \mathcal{U}$; an aggregate A is *convex* if $A(I) = A(K) = \mathbf{T}$ implies $A(J) = \mathbf{T}$, for all $I \subseteq J \subseteq K \subseteq \mathcal{U}$; the remaining aggregates are called *non-convex*. Note that monotone aggregates are convex, and the inclusion is strict. Relation \models is inductively defined as follows: for a propositional atom $p \in \mathcal{U}$, $I \models p$ if $p \in I$; for a negated literal $\sim l$, $I \models \sim l$ if $I \not\models l$; for an aggregate atom A , $I \models A$ if $A(I \cap dom(A)) = \mathbf{T}$; for a set or conjunction C , $I \models C$ if $I \models p$ holds for each $p \in C$; for a rule r , $I \models r$ if $H(r) \cap I \neq \emptyset$ whenever $I \models B(r)$. I is a *model* of a program Π if $I \models \Pi$, i.e., if $I \models r$ for all $r \in \Pi$.

Example 1

Let I be an interpretation, and $k \geq 1$. An aggregate A such that $A(I)$ equals $|dom(A) \cap I| \geq k$ is monotone. An aggregate A such that $A(I)$ equals $|dom(A) \cap I| = k$ is convex. An aggregate A such that $A(I)$ equals $|dom(A) \cap I| \neq k$ is non-convex.

Let A_1 be an aggregate such that $dom(A_1) = \{a, b\}$ and $A_1(I)$ equals $|\{a, b\} \cap I| \geq 1$, for all interpretations I . A program using A_1 is $\Pi_1 = \{a \leftarrow \sim a; b \vee c \leftarrow A_1\}$. The models of Π_1 , restricted to the atoms occurring in the program, are the following: \emptyset , $\{a, b\}$, $\{a, c\}$, $\{a, b, c\}$, $\{b\}$, $\{b, c\}$, and $\{c\}$.

F-stable models. Let Π be a program and I an interpretation. The F-reduct of Π with respect to I is defined as follows: $F(\Pi, I) = \{F(r, I) \mid r \in \Pi, I \models B(r)\}$, where $F(r, I) = p_1 \vee \dots \vee p_m \leftarrow F(l_1, I), \dots, F(l_n, I)$ for r being of the form (1), $F(l, I) = l$ if l is a propositional atom or an aggregate atom A , and $F(l, I) = \emptyset$ if l is a negative literal. I is an F-stable model of Π if $I \models \Pi$ and there is no $J \subset I$ such that $J \models F(\Pi, I)$. The set of F-stable models of Π is denoted $FSM(\Pi)$.

G-stable models. Let Π be a program and I an interpretation. The G-reduct of Π with respect to I is defined as follows: $G(\Pi, I) = \{G(r, I) \mid r \in \Pi, I \models B(r)\}$, where $G(r, I) = p_1 \vee \dots \vee p_m \leftarrow G(l_1, I), \dots, G(l_n, I)$ for r being of the form (1), $G(l, I) = p$ if l is a propositional atom p , $G(l, I) = I \cap dom(A)$ if l is an aggregate atom A , and $G(l, I) = \emptyset$ if l is a negative literal. I is a G-stable model of Π if $I \models \Pi$ and there is no $J \subset I$ such that $J \models G(\Pi, I)$. The set of G-stable models of Π is denoted $GSM(\Pi)$.

Example 2

The F-stable models of Π_1 in Example 1 are the following: \emptyset , $\{a, b\}$, and $\{a, c\}$. Indeed, note that $F(\Pi_1, \emptyset) = \emptyset$, $F(\Pi_1, \{a, b\}) = F(\Pi_1, \{a, c\}) = \{a \leftarrow; b \vee c \leftarrow A_1\}$, and each model is minimal for its reduct. On the other hand, $\{b\}$ is not an F-stable model because \emptyset is a model of $F(\Pi_1, \{b\}) = \{b \vee c \leftarrow A_1\}$. The G-stable models of Π_1 are the following: \emptyset and $\{a, c\}$. Indeed, $G(\Pi_1, \emptyset) = \emptyset$ and $G(\Pi_1, \{a, c\}) = \{a \leftarrow; b \vee c \leftarrow a\}$. Note that A_1 is replaced by a in the last rule of $G(\Pi_1, \{a, c\})$ because $\{a, c\} \cap dom(A_1) = \{a\}$. Also observe that $\{a, b\}$ is not a G-stable model because $G(\Pi_1, \{a, b\}) = \{a \leftarrow; b \vee c \leftarrow a, b\}$, and $\{a\}$ is a model of this reduct.

Computational problems. Let $X \in \{F, G\}$. A program Π is *X-coherent* if Π has at least one *X-stable* model; otherwise, Π is *X-incoherent*. *X-coherence testing* is the computational problem of checking whether an input program Π is *X-coherent*. A propositional atom p is an *X-cautious consequence* of Π if p belongs to all *X-stable* models of Π . *X-cautious reasoning* is the computational problem of checking whether a given atom p is an *X-cautious consequence* of an input program Π .

3 Complexity

Complexity of F-cautious reasoning, and implicitly also of F-coherence testing, was analyzed in (Faber *et al.* 2011). A similar analysis is reported here for G-stable semantics, and in particular the combination of monotone (M), convex (C) and non-convex (N) aggregates with negation (\sim) and disjunction (\vee) is analyzed. A summary of results is shown in Table 1, where all complexity bounds are tight. Note that in some cases the existence of a stable model is guaranteed, and hence

Table 1. Complexity of G-coherence testing and G-cautious reasoning. An \uparrow denotes an increase in complexity with respect to F-stable model semantics, where the considered complexity classes are $\mathbf{K} \subseteq P \subseteq NP \subseteq \Sigma_2^P$, and $\mathbf{K} \subseteq P \subseteq co-NP \subseteq \Pi_2^P$. Similarly, \downarrow denotes a decrease in complexity

	COHERENCE TESTING				CAUTIOUS REASONING			
	{}	{~}	{V}	{~, V}	{}	{~}	{V}	{~, V}
—	K	NP	K	Σ_2^P	P	co-NP	co-NP	Π_2^P
M	P \uparrow	NP	Σ_2^P $\uparrow\uparrow\uparrow$	Σ_2^P	P	co-NP	Π_2^P \uparrow	Π_2^P
C	NP	NP	Σ_2^P	Σ_2^P	co-NP	co-NP	Π_2^P	Π_2^P
N	NP \downarrow	NP \downarrow	Σ_2^P	Σ_2^P	co-NP \downarrow	co-NP \downarrow	Π_2^P	Π_2^P

constant complexity **K** is reported. Throughout this section, aggregates are assumed to be polynomial-time computable functions, and $ASP(X)$ will denote the class of programs using constructs in the list X . For example, $ASP(\neg, C)$ is the class of programs possibly using negation and convex aggregates, while $ASP(-)$ is the class of programs not using negation, disjunction or aggregates.

3.1 Complexity of coherence testing

Complexity of coherence testing for programs without aggregates, reported in the first row of Table 1, is well-known (see for example (Eiter and Gottlob 1995)). Membership in Σ_2^P is implicit in (Gelfond and Zhang 2014) for the general case. For the other membership results, the immediate consequence operator is used.

Definition 1

Let Π be a program and I an interpretation. The immediate consequence operator T_Π is defined as follows: $T_\Pi(I) = \{p \in H(r) \mid r \in \Pi, I \models B(r)\}$.

For an $ASP(M)$ program Π , T_Π is monotone and therefore has a least fixpoint, and this fixpoint is computable in polynomial-time because a single application requires linear-time, and at most $|At(\Pi)|$ applications are required to reach the fixpoint. Moreover, G-stable models of $ASP(M)$ programs can be characterized in terms of T_Π , from which P-membership follows.

Lemma 1

Let Π be in $ASP(M)$. The least fixpoint of $T_\Pi(I)$ exists and is polytime computable. Let I be the least fixpoint of T_Π , and J be the least fixpoint of $T_{G(\Pi, I)}$. If $I \neq J$ then Π is G-incoherent, otherwise $GSM(\Pi) = \{I\}$.

Theorem 1

G-coherence testing is in P for $ASP(M)$.

To obtain NP-membership in the disjunction-free case, the following algorithm is used: Guess a model I of Π and check that I is a minimal model of $G(\Pi, I)$. Checking that I is a model of Π and that I is minimal for $G(\Pi, I)$ is polynomial-time doable (note that $G(\Pi, I)$ is in $ASP(-)$ and hence Lemma 1 can be used).

Theorem 2

G-coherence testing is in NP for programs in $\text{ASP}(\neg, M, C, N)$.

As for the hardness, it is known that coherence testing is NP-hard if negation is present (Dantsin *et al.* 2001), while adding also disjunction increases the hardness to Σ_2^P (Eiter and Gottlob 1995). These results propagate top-down in Table 1. The missing results for programs with only convex or only non-convex aggregates can be obtained by the following transformations from aggregate-free programs with negation to negation-free programs with aggregates.

Definition 2

Let Π be in $\text{ASP}(\neg, \vee)$. Let $C(\Pi)$ be the program obtained from Π by replacing every occurrence of a negative literal $\sim p$ with an aggregate A such that $\text{dom}(A) = \{p\}$ and $A(I) = |\{p\} \cap I| \leq 0$, for all $I \subseteq \mathcal{U}$. Let $N(\Pi)$ be the program obtained from Π by replacing every occurrence of a negative literal $\sim p$ with an aggregate A such that $\text{dom}(A) = \{p, \perp\}$ and $A(I) = |\{p, \perp\} \cap I| \neq 1$, for all $I \subseteq \mathcal{U}$, where \perp is a fixed atom not occurring in Π .

Lemma 2

Let Π be in $\text{ASP}(\neg, \vee)$. Then, $\text{GSM}(\Pi) \equiv_{\text{At}(\Pi)} \text{GSM}(C(\Pi)) \equiv_{\text{At}(\Pi)} \text{GSM}(N(\Pi))$.

Since $C(\Pi)$ and $N(\Pi)$ can be obtained in polynomial-time and only comprise convex and non-convex aggregates, respectively, the hardness results are obtained.

Theorem 3

G-coherence testing is Σ_2^P -hard for both $\text{ASP}(\vee, C)$ and $\text{ASP}(\vee, N)$. It is NP-hard for both $\text{ASP}(C)$ and $\text{ASP}(N)$.

The only missing cases are now for programs with monotone aggregates. If disjunction and negation are not allowed, G-coherence testing is P-hard because of the following reduction: Let Π be in $\text{ASP}(-)$, and p be a propositional atom. Checking whether p is a cautious consequence of Π is equivalent to test coherence of $\Pi \cup \{p \leftarrow A\}$, where $\text{dom}(A) = \{p\}$ and $A(I) = |\{p\} \cap I| \geq 0$, for all $I \subseteq \mathcal{U}$. Since cautious reasoning is P-hard for $\text{ASP}(-)$ (i.e., checking if a propositional atom belongs to the unique model of a Datalog program), and program $\Pi \cup \{p \leftarrow A\}$ can be built using constant space, the complexity result is obtained.

Theorem 4

G-coherence testing is P-hard for $\text{ASP}(M)$.

For the disjunctive case, instead, the following transformation is used.

Definition 3

Let Π be in $\text{ASP}(\neg, \vee)$. Let $M(\Pi)$ be the program obtained from Π by replacing every occurrence of a negative literal $\sim p$ with p^F , where p^F is a fresh propositional atom associated with p , and by adding rule $p \vee p^F \leftarrow A$, where $\text{dom}(A) = \{p\}$ and $A(I) = |\{p\} \cap I| \geq 0$, for all $I \subseteq \mathcal{U}$.

Lemma 3

Let Π be in $\text{ASP}(\neg, \vee)$. The following relation holds: $\text{GSM}(\Pi) \equiv_{\text{At}(\Pi)} \text{GSM}(M(\Pi))$.

Since $M(\Pi)$ is polynomial-time constructible and only comprises monotone aggregates, Σ_2^P -hardness follows.

Theorem 5

G-coherence testing is Σ_2^P -hard for $\text{ASP}(\vee, M)$.

3.2 Complexity of cautious reasoning

As in the previous section, the first row of Table 1 reports well-known results concerning complexity of cautious reasoning for programs without aggregates. Moreover, membership in Π_2^P is proved in (Gelfond and Zhang 2014) for the general case. For a program Π in $\text{ASP}(M)$, membership in P is obtained by noting that the unique G-stable model candidate of Π can be computed in polynomial-time, as shown in the previous section.

Theorem 6

G-cautious reasoning is in P for $\text{ASP}(M)$.

For a disjunction-free program Π and a propositional atom p , the complementary problem can be solved by guessing an interpretation I such that $p \notin I$, and checking that I is a G-stable model of Π . It is a polytime check because $G(\Pi, I)$ is $\text{ASP}(-)$.

Theorem 7

G-cautious reasoning is in co-NP for programs in $\text{ASP}(\neg, M, C, N)$.

As for the hardness, first of all observe that P-hardness for the simplest case provides P-hardness for any case. Moreover, coherence testing can be reduced to (the complement of) cautious reasoning in general. In more detail, a program Π is G-coherent if and only if \perp is not a G-cautious consequence of Π , where $\perp \in \mathcal{U} \setminus \text{At}(\Pi)$ is an atom not occurring in Π . In fact, if Π is G-coherent then its G-stable models cannot contain \perp , and therefore \perp is not a G-cautious consequence of Π . Otherwise, \perp is a G-cautious consequence because Π has no stable models. Since G-coherence of Π can be equivalently checked on $M(\Pi)$, $C(\Pi)$, or $N(\Pi)$, all other hardness results are obtained from Theorems 3–5.

Theorem 8

G-cautious consequence is Π_2^P -hard for $\text{ASP}(\vee, M)$, $\text{ASP}(\vee, C)$ and $\text{ASP}(\vee, N)$. It is co-NP-hard for $\text{ASP}(C)$ and $\text{ASP}(N)$.

4 Compilation

G-stable models of a logic program can be computed by compiling into F-stable model semantics, for which efficient implementations are available. (Translations into other frameworks, for example Ferraris's propositional theories (Ferraris 2005), are also possible, but not the focus of this paper.) Two different rewritings are presented

in this section. The first rewriting is more compact, in the sense that it introduces fewer auxiliary atoms. The second rewriting instead requires more auxiliary atoms, but has the advantage that the output program only comprises stratified aggregates (essentially, in these programs no recursive definition involves an aggregate; see (Faber *et al.* 2011) for a formal definition).

Definition 4 (Rewriting 1)

Let Π be a program. Let $rew(\Pi)$ be the program obtained from Π by performing the following operations:

1. For each $p \in At(\Pi)$, a fresh propositional atom p' and the following rules are introduced: $p' \leftarrow \sim p$, and $p' \leftarrow p$.
2. For each aggregate A occurring in a rule r of Π and such that $dom(A) = \{p_1, \dots, p_n\}$ (for some $n \geq 0$), literals p'_1, \dots, p'_n are added to the body of r .

Example 3

Consider again program Π_1 from Example 1, whose G-stable models are \emptyset and $\{a, c\}$. Program $rew(\Pi_1)$ is the following: $\{a \leftarrow \sim \sim a; b \vee c \leftarrow A_1, a', b'\} \cup \{a' \leftarrow \sim a; a' \leftarrow a; b' \leftarrow \sim b; b' \leftarrow b; c' \leftarrow \sim c; c' \leftarrow c\}$. Its F-stable models are the following: $\emptyset \cup X$ and $\{a, c\} \cup X$, where $X = \{a', b', c'\}$. In fact, a', b', c' are necessarily true because of the rules introduced in item 1 of Definition 4. Moreover, note that if a is false in some model I then a' is necessarily true in any model of the reduct $F(\Pi_1, I)$. On the other hand, if a is true in I then a' can be possibly assumed false in a model of $F(\Pi_1, I)$. Similarly for b and b' , and for c and c' .

Intuitively, as also observed in the above example, all auxiliary propositional atoms are true in any model I of $rew(\Pi)$ because of the rules introduced in item 1 of Definition 4. Moreover, if J is a minimal model of $F(rew(\Pi), I)$ then the following properties are satisfied: (i) if $p \notin I$, then $p' \in J$; (ii) if $p \in I$, then $p' \in J$ if and only if $p \in J$. Correctness of the first compilation is thus established.

Theorem 9

Let Π be a program. The following relation holds: $GSM(\Pi) \equiv_{At(\Pi)} FSM(rew(\Pi))$.

A drawback of this first compilation is that the evaluation of the resulting program may be on a higher complexity class than the evaluation of the original program. For example, G-coherence testing of disjunction-free programs is NP-complete in general, while a Σ_2^P procedure will be used to test F-coherence of the rewritten program. Such a drawback motivates the introduction of a second compilation. To ease the presentation, and to provide a better analysis later, the syntax of the language is extended by allowing the use of *integrity constraints*, that is, rules of the form (1) with empty heads. Note that the semantics provided in Section 2 can already cope with such an extension.

Definition 5 (Rewriting 2)

Let Π be a program. Let $str(\Pi)$ be the (stratified) program obtained from Π by performing the following operations:

1. For each $p \in At(\Pi)$, two fresh propositional atoms p', p'' and the following three groups of rules are introduced: (i) $p' \leftarrow \sim p$, and $p' \leftarrow p$; (ii) $p'' \leftarrow \sim \sim p''$; (iii) $\leftarrow \sim p'', p$, and $\leftarrow p'', \sim p$.
2. For each aggregate A occurring in a rule r of Π and such that $dom(A) = \{p_1, \dots, p_n\}$ ($n \geq 0$), literals p'_1, \dots, p'_n are added to $B(r)$, and A is replaced by a new aggregate A' such that $dom(A') = \{p'_1, \dots, p'_n\}$ and $A'(I) = A(\{p \in \mathcal{U} \mid p'' \in I\})$, for all $I \subseteq \mathcal{U}$.

Example 4

Resorting again to Π_1 from Example 1, $str(\Pi_1)$ is the following program: $\{a \leftarrow \sim \sim a; b \vee c \leftarrow A'_1, a', b'\} \cup \{a' \leftarrow \sim a; b' \leftarrow \sim b; c' \leftarrow \sim c; a' \leftarrow a; b' \leftarrow b; c' \leftarrow c\} \cup \{a'' \leftarrow \sim \sim a''; b'' \leftarrow \sim \sim b''; c'' \leftarrow \sim \sim c''\} \cup \{\leftarrow \sim a'', a; \leftarrow \sim b'', b; \leftarrow \sim c'', c; \leftarrow a'', \sim a; \leftarrow b'', \sim b; \leftarrow c'', \sim c\}$. where $dom(A'_1) = \{a'', b''\}$ and $A'_1(I) = |\{a'', b''\} \cap I| \geq 1$, for all $I \subseteq \mathcal{U}$. The F-stable models of $str(\Pi_1)$ are the following: $\emptyset \cup X$ and $\{a, c\} \cup X \cup \{a'', c''\}$, where $X = \{a', b', c'\}$. In fact, for atoms a', b', c' , comments in Example 3 apply. Atom a'' instead is forced to have the same truth value of a because of rules of the group (iii). Similarly for b'' and b , and for c'' and c . Moreover, atoms a'', b'', c'' fix the interpretation of A'_1 in the reduct thanks to rules of the group (ii).

Rules of the group (i) are as in the first compilation, and therefore the already discussed properties on atoms of the form p' hold for $str(\Pi)$. Rules of the group (ii), instead, are used to guess an interpretation for atoms of the form p'' . Actually, they also force the interpretation of atoms of the form p because of rules of the group (iii). However, while the interpretation of an atom p'' is fixed also in the reduct, the interpretation of an atom p can be changed. Also the interpretation of aggregates is fixed in the reduct because their domains only contain atoms of the form p'' . Correctness of the second compilation is finally established.

Theorem 10

Let Π be a program. The following relation holds: $GSM(\Pi) \equiv_{At(\Pi)} FSM(str(\Pi))$.

4.1 Properties

The rewritings introduced in the previous section are *polynomial, faithful and modular* translation functions (Janhunen 2006), i.e., they are polynomial-time computable, preserve stable models (if auxiliary atoms are ignored), and can be computed independently on parts of the input program. In fact, faithfulness is preserved because of Theorems 9–10, and modularity can be easily proved by assuming that different auxiliary atoms are used for different parts of the program.

Theorem 11

Let Π, Π' be programs such that $\Pi \cap \Pi' = \emptyset$. For $tr \in \{rew, str\}$, the following conditions are satisfied: $tr(\Pi \cup \Pi') = tr(\Pi) \cup tr(\Pi')$, and $tr(\Pi) \cap tr(\Pi') = \emptyset$.

It is also possible to show that the rewritings are polynomial-time computable and have linear size with respect to the original program. For this purpose, the size of a program Π , denoted $\|\Pi\|$, is defined as the number of symbols occurring in

Π . In more detail, every occurrence of a propositional atom or of a negated literal is considered one symbol, while every occurrence of an aggregate A is counted as $|dom(A)|$ symbols. (No other symbol is considered in the size of Π .) The rewriting in Definition 4 introduces $|At(\Pi)|$ new propositional atoms, and $2 \cdot |At(\Pi)|$ new rules of size 2. Moreover, for each rule r in Π , program $rew(\Pi)$ contains a rule of size at most $2 \cdot \|r\|$ (because for each aggregate A in r , $|dom(A)|$ propositional atoms are added to the body of r). The rewriting in Definition 5, instead, introduces $2 \cdot |At(\Pi)|$ new propositional atoms, and $5 \cdot |At(\Pi)|$ new rules of size 2. The other rules in $str(\Pi)$ are obtained from rules in Π and have the same size of the corresponding rules in $rew(\Pi)$.

Theorem 12

Let Π be a program. The programs $rew(\Pi)$ and $str(\Pi)$ are polynomial-time constructible, and the following relations holds: (i) $\|rew(\Pi)\| \leq 4 \cdot |At(\Pi)| + 2 \cdot \|\Pi\|$; (ii) $\|str(\Pi)\| \leq 10 \cdot |At(\Pi)| + 2 \cdot \|\Pi\|$.

There are a few additional observations concerning the rewritings presented in the previous section, which also positively affect their sizes. The first observation is that fresh atoms could be added just for propositional atoms belonging to the domain of some aggregate occurring in Π . In fact, note that atoms c', c'' are not required in the rewritings reported in Examples 3–4. Such atoms are included in Definitions 4–5 to simplify the presentation of the rewritings. The second observation is more technical and concerns the implementations of current ASP solvers, which are essentially based on F -stable model semantics. ASP solvers use two modules, called *model generator* and *model checker*. The first module produces a model I of the input program Π , while the second module tests the stability of I , i.e., it checks whether no strict subset of I is a model of $F(\Pi, I)$. In both rewritings, atoms of the form p' are irrelevant for the model generator, in the sense that they are immediately derived true. Hence, the search space of the model generator is not increased at all when $rew(\Pi)$ is processed. A similar observation also applies to $str(\Pi)$. Indeed, atoms of the form p'' are constrained to have the same truth values of the corresponding atoms of the form p because of rules of the group (3). In addition, atoms of the form p'' are irrelevant for the model checker because their interpretation is fixed in this module by rules of the group (2). As a consequence, also the interpretation of all aggregates in $str(\Pi)$ is fixed in the model checker because their domains only comprise atoms of the form p'' .

Theorem 13

Let Π be a program, and I be an interpretation. If $I \models rew(\Pi)$ or $I \models str(\Pi)$ then $\{p' \mid p \in At(\Pi)\} \subseteq I$. Moreover, for each $J \subseteq I$ such that $J \models F(str(\Pi), I)$, it holds that $\{p'' \mid p \in I\} \subseteq J$.

A final observation, which is eventually linked to the previous, is that current ASP solvers only rely on the model generator to process disjunction-free programs. More specifically, this is the case if additionally non-convex aggregates are stratified. As already observed, the rewriting in Definition 5 is such that all aggregates in

the rewritten program are stratified. Moreover, note that $str(\Pi)$ does not introduce disjunction in Π (this is also true for $rew(\Pi)$).

Theorem 14

Let Π be a program. All aggregates in $str(\Pi)$ are stratified, and if Π has no disjunction then both $rew(\Pi)$ and $str(\Pi)$ have no disjunction.

Therefore, checking G-coherence of Π by means of checking F-coherence of $str(\Pi)$ is an appropriate technique from the complexity point of view, with the only corner case of ASP(M), i.e., programs without negation and disjunction, and whose aggregates are all monotone. A similar comment applies to performing G-cautious reasoning on Π by means of F-cautious reasoning on $str(\Pi)$. In fact, in the disjunction-free case, the rewriting in Definition 5 provides alternative proofs for NP-membership of G-coherence testing and co-NP-membership of G-cautious reasoning.

5 Related work

The challenge of extending stable model semantics with aggregate constructs has been investigated quite intensively in the previous decade. Among the many proposals, F-stable model semantics (Ferraris 2011; Faber *et al.* 2011) is of particular interest as many ASP solvers are currently based on this semantics (Gebser *et al.* 2012; Faber *et al.* 2008). Actually, the definition provided in Section 2 is slightly different than those in (Ferraris 2011; Faber *et al.* 2011). In particular, the language considered in (Ferraris 2011) has a broader syntax allowing for arbitrary nesting of propositional formulas. The language considered in (Faber *et al.* 2011), instead, does not allow explicitly the use of double negation, which however can be simulated by means of auxiliary atoms. For example, in (Faber *et al.* 2011) a rule $p \leftarrow \sim\sim p$ must be modeled by using a fresh atom p^F and the following subprogram: $\{p \leftarrow \sim p^F; p^F \leftarrow \sim p\}$. On the other hand, negated aggregates are permitted in (Faber *et al.* 2011), while they are forbidden in this paper. Actually, programs with negated aggregates are those for which (Ferraris 2011) and (Faber *et al.* 2011) disagree. As a final remark, the reduct of (Faber *et al.* 2011) does not remove negated literals from bodies, which however are necessarily true in all counter-models because double negation is not allowed in the syntax considered by (Faber *et al.* 2011).

Other relevant stable model semantics for logic programs with aggregates are reported in (Pelov *et al.* 2007; Son and Pontelli 2007) for disjunction-free programs, recently extended to the disjunctive case in (Shen *et al.* 2014). In these semantics the stability check is not given in terms of minimality of the model for the program reduct but obtained by means of a fixpoint operator similar to immediate consequence, and the following relation holds in general: stable models of (Shen *et al.* 2014) are a selection of F-stable models, and they coincide up to ASP(\neg, M, C), which is also the complexity boundary between the first and second level of the polynomial hierarchy for F-stable model semantics (Alviano and Faber 2013). Finally, a more recent proposal is G-stable model semantics (Gelfond and Zhang 2014), whose relation with other semantics has been highlighted by (Alviano and Faber 2015) in

the disjunction-free case: G-stable models are F-stable models, but the converse is not always true.

A detailed complexity analysis for F-stable models is reported in (Faber *et al.* 2011) and summarized in Table 1. Complexity of reasoning under stable models by (Pelov *et al.* 2007; Son and Pontelli 2007), instead, is analyzed in (Pelov 2004), where in particular Σ_2^P -completeness of coherence testing is proved for disjunction-free programs with aggregates. Concerning G-stable models, the general case was studied in (Gelfond and Zhang 2014), and a more detailed analysis is provided by this paper. In particular, for disjunction-free programs, the main reasoning tasks are in the first level of the polynomial hierarchy in general when G-stable models are used. On the other hand, coherence testing jumps from \mathbf{K} to Σ_2^P when F-stable models are replaced by G-stable models in programs with monotone aggregates only. Indeed, in contrast to previous semantics, monotone aggregates are enough to simulate integrity constraints and negation when G-stable models are used.

Techniques to rewrite logic programs with aggregates into equivalent aggregate-free programs were also investigated in the literature. For example, a rewriting into aggregate-free programs is presented by (Ferraris 2011) for F-stable model semantics. However, it must be noted that the rewriting of (Ferraris 2011) produces nested expressions in general, and current mainstream ASP systems cannot process directly such constructs, but instead require additional translations such as those by (Lee and Palla 2009). Other relevant rewriting techniques were proposed in (Bomanson and Janhunen 2013; Bomanson *et al.* 2014), also proved to be quite efficient in practice. However, these rewritings preserve F-stable models only in the stratified case, or if recursion is limited to convex aggregates.

Aggregate functions are also semantically similar to DL (Eiter *et al.* 2008) and HEX atoms (Eiter *et al.* 2014), extensions of ASP for interacting with external knowledge bases, possibly expressed in different languages.

6 Conclusion

G-stable models are a recent proposal for interpreting logic programs with aggregates. A detailed complexity analysis of the main reasoning tasks for this new semantics was reported in Section 3, highlighting similarities and differences versus mainstream ASP semantics, here referred to as F-stable models. In more detail, G-coherence testing is NP-complete for disjunction-free programs, in contrast to Σ_2^P -completeness of F-coherence testing. An even more surprising result was shown for negation-free programs with monotone aggregates: Such programs are guaranteed to be F-coherent, while G-coherence testing was shown to be Σ_2^P -hard because negation can be simulated by means of disjunction and monotone aggregates in the new semantics. Similar results were shown for G-cautious reasoning.

A further link between G- and F-stable models is provided by the rewritings in Section 4: G-stable models of an input program can be obtained by computing F-stable models of a rewritten program, where the size of the rewritten program is linear with respect to the size of the original program. In particular, two different rewritings are presented and analyzed. Moreover, one of these rewritings outputs

programs with stratified aggregates only, which are handled efficiently by modern ASP solvers. A prototype system supporting common aggregation functions such as COUNT, SUM, AVG, MIN, MAX, ODD, and EVEN is thus implemented by means of this rewriting, and using the ASP solver WASP (Alviano *et al.* 2013; Alviano *et al.* 2014) to obtain G-stable models of the original program.

Acknowledgement

This work was partially supported by MIUR within project “SI-LAB BA2KNOW – Business Analytics to Know”, and by Regione Calabria, POR Calabria FESR 2007-2013, within project “ITravel PLUS” and project “KnowRex”. Mario Alviano was partly supported by the National Group for Scientific Computation (GNCS-INDAM), and by Finanziamento Giovani Ricercatori UNICAL.

References

- ALVIANO, M., DODARO, C., FABER, W., LEONE, N. AND RICCA, F. 2013. WASP: A native ASP solver based on constraint learning. In *Logic Programming and Nonmonotonic Reasoning, 12th International Conference, LPNMR 2013, Corunna, Spain, September 15-19, 2013. Proceedings*, P. Cabalar and T. C. Son, Eds. Lecture Notes in Computer Science, vol. 8148. Springer, 54–66.
- ALVIANO, M., DODARO, C. AND RICCA, F. 2014. Anytime computation of cautious consequences in answer set programming. *Theory and Practice of Logic Programming* 14, 4-5, 755–770.
- ALVIANO, M. AND FABER, W. 2013. The complexity boundary of answer set programming with generalized atoms under the FLP semantics. In *Logic Programming and Nonmonotonic Reasoning, 12th International Conference, LPNMR 2013, Corunna, Spain, September 15-19, 2013. Proceedings*, P. Cabalar and T. C. Son, Eds. Lecture Notes in Computer Science, vol. 8148. Springer, 67–72.
- ALVIANO, M. AND FABER, W. 2015. Stable model semantics of abstract dialectical frameworks revisited: A logic programming perspective. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence*. IJCAI Organization, Buenos Aires, Argentina, To appear.
- BARTHOLOMEW, M., LEE, J. AND MENG, Y. 2011. First-order semantics of aggregates in answer set programming via modified circumscription. In *Logical Formalizations of Commonsense Reasoning, Papers from the 2011 AAAI Spring Symposium, Technical Report SS-11-06, Stanford, California, USA, March 21-23, 2011*. AAAI.
- BOMANSON, J., GEBSER, M. AND JANHUNEN, T. 2014. Improving the normalization of weight rules in answer set programs. In *Logics in Artificial Intelligence - 14th European Conference, JELIA 2014, Funchal, Madeira, Portugal, September 24-26, 2014. Proceedings*, E. Fermé and J. Leite, Eds. Lecture Notes in Computer Science, vol. 8761. Springer, 166–180.
- BOMANSON, J. AND JANHUNEN, T. 2013. Normalizing cardinality rules using merging and sorting constructions. In *Logic Programming and Nonmonotonic Reasoning, 12th International Conference, LPNMR 2013, Corunna, Spain, September 15-19, 2013. Proceedings*, P. Cabalar and T. C. Son, Eds. Lecture Notes in Computer Science, vol. 8148. Springer, 187–199.
- BREWKA, G., EITER, T. AND TRUSZCZYNSKI, M. 2011. Answer set programming at a glance. *Commun. ACM* 54, 12, 92–103.
- DANTSIN, E., EITER, T., GOTTLÖB, G. AND VORONKOV, A. 2001. Complexity and expressive power of logic programming. *ACM Comput. Surv.* 33, 3, 374–425.

- EITER, T., FINK, M., KRENNWALLNER, T., REDL, C. AND SCHÜLLER, P. 2014. Efficient hex-program evaluation based on unfounded sets. *J. Artif. Intell. Res. (JAIR)* 49, 269–321.
- EITER, T. AND GOTTLÖB, G. 1995. On the computational cost of disjunctive logic programming: Propositional case. *Ann. Math. Artif. Intell.* 15, 3-4, 289–323.
- EITER, T., IANNI, G., LUKASIEWICZ, T., SCHINDLAUER, R. AND TOMPITS, H. 2008. Combining answer set programming with description logics for the semantic web. *Artif. Intell.* 172, 12-13, 1495–1539.
- FABER, W., PFEIFER, G. AND LEONE, N. 2011. Semantics and complexity of recursive aggregates in answer set programming. *Artif. Intell.* 175, 1, 278–298.
- FABER, W., PFEIFER, G., LEONE, N., DELL'ARMI, T. AND IELPA, G. 2008. Design and implementation of aggregate functions in the DLV system. *Theory and Practice of Logic Programming* 8, 5-6, 545–580.
- FERRARIS, P. 2005. Answer sets for propositional theories. In *Logic Programming and Nonmonotonic Reasoning, 8th International Conference, LPNMR 2005, Diamante, Italy, September 5-8, 2005, Proceedings*, C. Baral, G. Greco, N. Leone, and G. Terracina, Eds. Lecture Notes in Computer Science, vol. 3662. Springer, 119–131.
- FERRARIS, P. 2011. Logic programs with propositional connectives and aggregates. *ACM Trans. Comput. Log.* 12, 4, 25.
- GEBSER, M., KAUFMANN, B. AND SCHAUB, T. 2012. Conflict-driven answer set solving: From theory to practice. *Artif. Intell.* 187, 52–89.
- GELFOND, M. AND LIFSCHITZ, V. 1988. The stable model semantics for logic programming. In *Logic Programming, Proceedings of the Fifth International Conference and Symposium, Seattle, Washington, August 15-19, 1988 (2 Volumes)*, R. A. Kowalski and K. A. Bowen, Eds. MIT Press, 1070–1080.
- GELFOND, M. AND LIFSCHITZ, V. 1991. Classical negation in logic programs and disjunctive databases. *New Generation Comput.* 9, 3/4, 365–386.
- GELFOND, M. AND ZHANG, Y. 2014. Vicious circle principle and logic programs with aggregates. *Theory and Practice of Logic Programming* 14, 4-5, 587–601.
- JANHUNEN, T. 2006. Some (in)translatability results for normal logic programs and propositional theories. *Journal of Applied Non-Classical Logics* 16, 1-2, 35–86.
- LEE, J. AND PALLA, R. 2009. System f2lp - computing answer sets of first-order formulas. In *Logic Programming and Nonmonotonic Reasoning, 10th International Conference, LPNMR 2009, Potsdam, Germany, September 14-18, 2009. Proceedings*, E. Erdem, F. Lin, and T. Schaub, Eds. Lecture Notes in Computer Science, vol. 5753. Springer, 515–521.
- LIU, L., PONTELLI, E., SON, T. C. AND TRUSZCZYNSKI, M. 2010. Logic programs with abstract constraint atoms: The role of computations. *Artif. Intell.* 174, 3-4, 295–315.
- LIU, L. AND TRUSZCZYNSKI, M. 2006. Properties and applications of programs with monotone and convex constraints. *J. Artif. Intell. Res. (JAIR)* 27, 299–334.
- PELOV, N. 2004. Semantics of Logic Programs with Aggregates. Ph.D. thesis, Katholieke Universiteit Leuven, Leuven, Belgium.
- PELOV, N., DENECKER, M. AND BRUYNOOGHE, M. 2007. Well-founded and stable semantics of logic programs with aggregates. *Theory and Practice of Logic Programming* 7, 3, 301–353.
- SHEN, Y., WANG, K., EITER, T., FINK, M., REDL, C., KRENNWALLNER, T. AND DENG, J. 2014. FLP answer set semantics without circular justifications for general logic programs. *Artif. Intell.* 213, 1–41.
- SIMONS, P., NIEMELÄ, I. AND SOININEN, T. 2002. Extending and implementing the stable model semantics. *Artif. Intell.* 138, 1-2, 181–234.
- SON, T. C. AND PONTELLI, E. 2007. A constructive semantic characterization of aggregates in answer set programming. *Theory and Practice of Logic Programming* 7, 3, 355–375.