

Towards Dynamic Consistency Checking in Goal-directed Predicate Answer Set Programming ^{*}

Joaquín Arias¹[0000-0003-4148-311X], Manuel Carro^{2,3}[0000-0001-5199-3135], and Gopal Gupta⁴[0000-0001-9727-0362]

¹ CETINIA, Universidad Rey Juan Carlos, Madrid, Spain

² Universidad Politécnica de Madrid, Spain

³ IMDEA Software Institute, Madrid, Spain

⁴ University of Texas at Dallas, Richardson, USA

joaquin.arias@urjc.es, manuel.carro@{upm.es, imdea.org}, gupta@utdallas.edu

Abstract. Goal-directed evaluation of Answer Set Programs is gaining traction thanks to its amenability to create AI systems that can, due to the evaluation mechanism used, generate explanations and justifications. s(CASP) is one of these systems and has been already used to write reasoning systems in several fields. It provides enhanced expressiveness w.r.t. other ASP systems due to its ability to use constraints, data structures, and unbound variables natively. However, the performance of existing s(CASP) implementations is not on par with other ASP systems: model consistency is checked once models have been generated, in keeping with the generate-and-test paradigm. In this work, we present a variation of the top-down evaluation strategy, termed *Dynamic Consistency Checking*, which interleaves model generation and consistency checking. This makes it possible to determine when a literal is not compatible with the *denials* associated to the global constraints in the program, prune the current execution branch, and choose a different alternative. This strategy is specially (but not exclusively) relevant in problems with a high combinatorial component. We have experimentally observed speedups of up to 90× w.r.t. the standard versions of s(CASP).

Keywords: Dynamic Consistency Checking · Goal-Directed Evaluation · Constraints · Answer Set Programming

1 Introduction

s(CASP) [4] is a novel non-monotonic reasoner that evaluates Constraint Answer Set Programs without a grounding phase, either before or during execution. s(CASP) supports predicates and thus retains logical variables (and constraints)

^{*} Work partially supported by EIT Digital, MICINN projects RTI2018-095390-B-C33 InEDGEMobility (MCIU/AEI/FEDER, UE), PID2019-108528RB-C21 ProCode, Comunidad de Madrid project S2018/TCS-4339 BLOQUES-CM co-funded by EIE Funds of the European Union, US NSF Grants IIS 1718945, IIS 1910131, IIP 1916206.

both during the execution and in the answer sets. The operational semantics of s(CASP) relies on backward chaining, which is intuitive to follow and lends itself to generating explanations that can be translated into natural language [2]. The execution of an s(CASP) program returns partial stable models: the subsets of the stable models [10] which include only the (negated) literals necessary to support the initial query. To the best of our knowledge, s(CASP) is the only system that exhibits the property of relevance [19]. s(CASP) has been already applied in relevant fields related to the representation of commonsense reasoning:

- An automated reasoner that uses Event Calculus (EC) [1] (<http://bit.ly/EventCalculus>). s(CASP) can make deductive reasoning tasks in domains featuring constraints involving dense time and continuous properties. It is being used to model real-world avionics systems, to verify (timed) properties, and to identify gaps with respect to system requirements [11].
- The s(CASP) justification framework has been used to bring Explainable Artificial Intelligence (XAI) principles to rule-based systems capturing expert knowledge [2,8].
- It is at the core of two natural language understanding systems [6]: SQuARE, a Semantic-based Question Answering and Reasoning Engine, and Sta-CAACK, Stateful Conversational Agent using Commonsense Knowledge. They use the s(CASP) engine to “truly understand” and perform reasoning while generating natural language explanations for their responses. Building on these systems, s(CASP) was used to develop one of the nine systems selected to participate in the Amazon Alexa Socialbot Grand Challenge 4 [7],⁵ and is being used to develop a conversational AI chatbot.
- It has been used in ILP systems that generate ASP programs [20] and concurrent imperative programs from behavioral, observable specifications [22].
- A legal expert system [18], developed at the SMU Centre for Computational Law at Singapore, coded rule 34 of Singapore’s Legal Profession.⁶ Its front-end is a web interface that collects user information, runs queries on s(CASP), and displays the results with explanations in natural language.
- s(LAW), an administrative and judicial discretion reasoner [5], which allows modeling legal rules involving ambiguity and infers conclusions, providing (natural language) justifications based on them.

However, in the standard implementation of s(CASP), the global constraints in a program are checked when a tentative but complete model is computed. This strategy takes a large toll on the performance of programs that generate many tentative models and use global constraints to discard those that do not satisfy the specifications of the problem.

In this work, we propose a technique termed *Dynamic Consistency Checking* (DCC) that anticipates the evaluation of global constraints to discard inconsistent models as early as possible. Before adding a literal to the tentative model, DCC checks if any global constraint is violated. If so, this literal is discarded and

⁵ <https://cs.utdallas.edu/computer-scientists-enhance-alexas-small-talk-skills/>

⁶ https://github.com/smucclaw/r34_sCASP

the evaluation backtracks to look for other alternatives. We show, through several examples, that using this preliminary implementation, s(CASP) with DCC is up to 90× faster. Section 2 contains an overview of the syntax, operational semantics, and implementation of s(CASP). Section 3 explains the motivation behind DCC with examples, and describes its design and implementation. Section 4 presents the evaluating results of several benchmarks using s(CASP) with DCC enabled or not. Finally, in Section 5 we draw conclusions and propose future work. All the program and files used or mentioned in this paper are available at <http://platon.etsii.urjc.es/~jarias/papers/dcc-pad121/>.

2 Background: s(CASP)

An s(CASP) program is a set of clauses of the following form:

$$a :- c_a, b_1, \dots, b_m, \text{not } b_{m+1}, \dots, \text{not } b_n.$$

where a and b_1, \dots, b_n are atoms. An atom is either a propositional variable or the expression $p(t_1, \dots, t_n)$ if p is an n -ary predicate symbol and t_1, \dots, t_n are terms. A term is either a variable x_i or a function symbol f of arity n , denoted as f/n , applied to n terms, e.g., $f(t_1, t_2, \dots, t_n)$, where each t_i is in turn a term. A function symbol of arity 0 is called a constant. Therefore, s(CASP) accepts terms with the same conventions as Prolog: $f(a, b)$ is a term, and so are $f(g(X), Y)$ and $[f(a)|\text{Rest}]$ (to denote a list with head $f(a)$ and tail Rest). Program variables are usually written starting with an uppercase letter,⁷ while function and predicate symbols start with a lowercase letter. Numerical constants are written solely with digits.

The term c_a is a simple constraint or a conjunction of constraints: an expression establishing relations among variables in some constraint system [15]. Similar to CLP, s(CASP) is parametrized w.r.t. the constraint system, from which it inherits its semantics. Since the execution of an s(CASP) program needs negating constraints, soundness requires that this can be done in the chosen constraint system by means of a finite disjunction of basic constraints [9,21].

At least one of a , b_i , or $\text{not } b_i$ must be present. When the head a is not present, it is supposed to be substituted by the head *false*. Headless rules have then the form

$$:- c_a, b_1, \dots, b_m, \text{not } b_{m+1}, \dots, \text{not } b_n.$$

and their interpretation is that the conjunction of the constraints and goals has to be false, so at least one constraint or goal has to be false. For example, the rule $:- p, q$, expresses that the conjunction of atoms $p \wedge q$ cannot be true: either p , q , or both, have to be false in any stable model. ASP literature often uses the term *constraint* to denote these constructions. To avoid the ambiguity that may arise from using the same name for constraints appearing among (free) variables during program execution and in the final models and for rules without heads, we will refer to headless rules as *denials* [16].

⁷ There are additional syntactical conventions to distinguish variables and non-variables that are of no interest here.

The execution of an s(CASP) program starts with a *query* of the form

$$?- c_a, b_1, \dots, b_m, \text{not } b_{m+1}, \dots, \text{not } b_n.$$

The s(CASP) answers to a query are *partial* stable models where each one is a subset of a stable model that satisfies the constraints, makes non-negated atoms true, makes the negated atoms non-provable, and, in addition, includes only atoms that are relevant to support the query. Additionally, for each partial stable model s(CASP) returns on backtracking *partial answer sets* with the justification tree and the bindings for the free variables of the query that correspond to the most general unifier (*mgu*) of a successful top-down derivation.

While in regular ASP systems, such as clingo [?], a preliminary grounding phase is required, s(CASP) retains logical variables both during execution and in the answer sets. As a consequence of the grounding phase, when constraints are included in regular ASP systems, it causes a loss of communication from elimination of variables and the execution methods for CASP systems are complex, e.g., EZSMT[?] needs explicit hooks to communicate the ASP solver and the constraint solver used, and CLINGCON [?] generates and propagates new constraints during the search and they are checked for consistency using external solvers.

2.1 Execution Procedure of s(CASP)

Let us present an abridged description of the top-down strategy of s(CASP):

1. Rules expressing the constructive negation of the predicates in the original ASP program are synthesized. We call this the *dual program*. Its mission is to provide a means to constructively determine the conditions and constraints under which calls to non-propositional predicates featuring variables would have failed: if we want to know when a rule such as $p(X, Y) :- q(X), \text{not } r(Y)$ succeeds, the dual program computes the constraints on Y under which the call $r(Y)$ would fail. This is an extension of the usual ASP semantics that is compatible with the case of programs that can be finitely grounded.⁸ A description of the construction of the dual program can be found in [3,4,14].
2. The original program is checked for loops of the form $r :- q, \text{not } r.$ and introduces additional denials to ensure that the models satisfy $\neg q \vee r$, even if the atoms r or q are not needed to solve the query. This is done by building a dependency graph of the program and detecting the paths where this may happen, including call paths across calls. Therefore, for the program:

```
1 p :- not q.           2 q :- not p.           3 r :- not r.
```

s(CASP) will determine that there are no stable models, regardless of the initial query. For the propositional case, such an analysis can be precise. For the non-propositional case, an over-approximation is calculated. In both

⁸ Note that, in the presence of function symbols and constraints on dense domains, this is in general not the case for s(CASP) programs.

cases, denials that are not used during program evaluation can be generated. These may impose a penalty in execution time, but are safe.

3. The denials generated in point 2, together with any denials present in the original program, are collected in predicates synthesized by the compiler that are evaluated by adding an auxiliary goal, `nmr_check/0`, at the end of the top-level query.
4. The union of the original program, the dual program, and the denials is handled by a top-down algorithm that implements the stable model semantics.

Point number 4 is specially relevant. The dual program (point 1) is synthesized by means of program transformations drawing from classical logic. However, its intended meaning differs from that of first-order logic. That is so because it is to be executed by a metainterpreter that does not implement the inference mechanisms of first-order logic, as it is designed to ensure that the semantics of answer set programs is respected (see Section 2.4). In particular, it treats specifically cyclic dependencies involving negation.

2.2 Unsafe Variables and Uninterpreted Function Symbols

The following code, from [3, Pag. 9] has variables that would be termed as *unsafe* in regular ASP systems: variables that appear in negated atoms in the body of a clause, but that do not appear in any positive literal in the same body.

```

1 p(X):- q(X, Z), not r(X).           3 q(X, a):- X #> 5.
2 p(Z):- not q(X, Z), r(X).         4 r(X):- X #< 1.

```

Since s(CASP) synthesizes explicit constructive goals for these negated goals, the aforementioned code can be run as-is in s(CASP). The query `?- p(A) .` generates three different answer sets, one for each binding:

```

{ p(A| {A #> 5}), q(A| {A #> 5}, a), not r(A| {A #> 5}) }
  A #> 5
{ p(A| {A \= a}), not q(B| {B #< 1}, A| {A \= a}), r(B| {B #< 1}) }
  A \= a
{ p(a), not q(B| {B #< 1}, a), r(B| {B #< 1}) }
  A = a

```

where the notation $V| \{C\}$ for a variable V is intended to mean that V is subject to the constraints in $\{C\}$. The constraints $A = 5$, $A \neq a$, and $A = a$ correspond to the bindings of variable A that make the atom from the query `?- p(A) .` belong to the stable model.

Another very relevant point where s(CASP) differs from ASP is in the possibility of using arbitrary uninterpreted function symbols to build, for example, data structures. While in mainstream ASP implementations these could give rise to an infinite grounded program (i.e., if the program does not have the *bound-term-depth* property), the s(CASP) execution model can deal with them similarly to Prolog, with the added power of the use of constructive negation in the execution and in the returned models.

Example 1. The predicate `member/2` below, from [3, Pag. 11], models the membership to a list as it is usual in (classical) logic programming. The query is intended to derive the conditions for one argument not to belong to a given list.

```

1 member(X, [X|Xs]).                3 list([1,2,3,4,5]).
2 member(X, [_|Xs]):- member(X, Xs). 4 ?- list(A), not member(B, A).

```

This program and query return in `s(CASP)` the following model and bindings:

```

{ list([1,2,3,4,5]),
  not member(B| {B \= 1,B \= 2,B \= 3,B \= 4,B \= 5}, [1,2,3,4,5]),
  not member(B| {B \= 1,B \= 2,B \= 3,B \= 4,B \= 5}, [2,3,4,5]),
  not member(B| {B \= 1,B \= 2,B \= 3,B \= 4,B \= 5}, [3,4,5]),
  not member(B| {B \= 1,B \= 2,B \= 3,B \= 4,B \= 5}, [4,5]),
  not member(B| {B \= 1,B \= 2,B \= 3,B \= 4,B \= 5}, [5]),
  not member(B| {B \= 1,B \= 2,B \= 3,B \= 4,B \= 5}, []) }
A = [1,2,3,4,5], B \= 1, B \= 2, B \= 3, B \= 4, B \= 5

```

I.e., for variable `B` not to be a member of the list `[1,2,3,4,5]` it has to be different from each of its elements.

2.3 `s(CASP)` as a Conservative Extension of ASP

The behavior of `s(CASP)` and ASP is the same for propositional programs. They differ in programs with unsafe variables (not legal in mainstream ASP systems), programs that could create unbound data structures, or whose variable ranges are defined in infinite domains (either unbound or bound but dense). Such programs are outside the standard domain of ASP systems as they cannot be finitely grounded. For them, `s(CASP)` extends ASP consistently.

In addition, the domain of the variables is implicitly expanded to include a domain which can be potentially infinite. Let us use the following example, from [3, Pag. 12], where we are interested in knowing whether `p(X)` (for some `X`) is or not part of a stable model:

```

1 d(1).                2 p(X) :- not d(X).

```

The only constant in the program is `1`, which is the only possible domain for `X` in the second clause. That clause is not legal for ASP, as `X` is an unsafe variable (Section 2.2). Adding a domain predicate call for it (i.e., adding `d(X)` to the body of the second clause), makes its model be `{d(1)}` (`not p(1)` is implicit).

That second clause is however legal in `s(CASP)`. Making the query `?- p(X)` returns the *partial* model `{p(X| {X \= 1}), not d(X|{X \= 1})}` stating that `p(X)` and `not d(X)` are true when `X \= 1`, which is consistent with, but more general than, the model given by ASP. As the model is partial, only the atoms (perhaps negated) involved in the proof for `?- p(X)` appear in that model.

```

1  scasp(Query) :-
2      solve(Query, [], Mid),
3      solve_goal(nmr_check, Mid, Model),
4      print_just_model(Model).
5
6  solve([], In, ['$success' | In]).
7  solve([Goal | Gs], In, Out) :-
8      solve_goal(Goal, In, Mid),
9      solve(Gs, Mid, Out).
10
11 solve_goal(Goal, In, Out) :-
12     user_defined(Goal), !,
13     check_loops(Goal, In, Out).
14 solve_goal(Goal, In, Out) :-
15     Goal=forall(Var, G), !,
16     c_forall(V, G, In, Out).
17 solve_goal(Goal, In, Out) :-
18     call(Goal),
19     Out=['$success', Goal | In].
20
21 check_loops(Goal, In, Out) :-
22     loop_type(Goal, In, Loop),
23     s_loop(Loop, Goal, In, Out).
24
25 s_loop(odd, _, _, _) :- fail.
26 s_loop(even, G, In, [chs(G) | In]).
27 s_loop(no_loop, G, In, Out) :-
28     pr_rule(G, Body),
29     solve(Body, [G | In], Out).
30 s_loop(proved, G, In, [proved(G) | In]).
31 s_loop(positive, _, _, _) :- fail.

```

Fig. 1. Outline of the s(CASP) interpreter’s code implemented in Ciao Prolog.

2.4 The s(CASP) Interpreter

Queries to the original program extended with the dual rules are evaluated by a runtime environment. This is currently a metainterpreter (see Fig. 1) in Prolog that executes an algorithm [12] that has similarities with SLD resolution. But it takes into account specific characteristics of ASP and the dual programs, such as the denials, the different kinds of loops, and the introduction of universal quantifiers in the body of the clauses:

1. The query `Query` is evaluated invoking `solve/3` in line 2 starting with an empty model represented as the empty list `[]`.
2. After the query evaluation, and to ensure that the returned model, `Mid`, is consistent with the denials, `nmr_check` (item 3 in page 7) is evaluated in line 3.
3. In line 4, the models that are consistent (and their justifications), `Model`, are output by `print_just_model/1`.
4. The predicate `solve/3` receives a list with the literals in the query (or in the body of some rule) and evaluates them, one by one, invoking `solve_goal/3`.
5. When the literal is a user defined predicate (line 12), the interpreter checks if there is a loop invoking `check_loops/3`. Three main cases are distinguished by `type_loop/3`:

Odd loop When a call eventually invokes itself and there is an odd number of intervening negations (as in, e.g., `p:- q. q:- not r. r:- p.`), the evaluation fails in line 25, to avoid contradictions of the form $p \wedge \neg p$, and backtracking takes place.

Even loop When there is an even number of intervening negations (as in `p:- not q. q:- r. r:- not p.`), the metainterpreter succeeds in line 26 to generate several stable models, such as `{p, not q, not r}` and `{q, r, not p}`.

No loop If no loops are detected, in line 27 the interpreter invokes `pr_rule/2` to retrieve the rule that unifies with the goal `G` and continues the evaluation by invoking `solve/3` with the literals of the rule.

6. The construction `forall(Var, G)` in line 15 is the dual of the existential quantifications in the body of the clauses. To evaluate them the runtime environment invokes the predicate `c_forall/4` in line 16, which determines if `G` holds for all the values of `Var` — see [4] for implementation details.
7. Finally, operations involving constraints and/or builtins are natively handled by invoking `call/1` in line 18.

3 Dynamic Consistency Checking in s(CASP)

The Dynamic Consistency Checking proposal of [13] is designed for propositional programs, while our proposal can also take care of predicate ASP programs. It is based on anticipating the evaluation of denials to fail as early as possible

3.1 Motivation

As we mentioned before, a denial such as $\neg p, q$, expresses that the conjunction of atoms $p \wedge q$ cannot be true: either p , q , or both, have to be false in any stable model. In predicate ASP the atoms have variables and a denial such as $\neg p(X), q(X, Y)$ means that:

$$false \leftarrow \exists x, y (p(x) \wedge q(x, y))$$

i.e., $p(X)$ and $q(X, Y)$ can not be simultaneously true for any possible values of X and Y in any stable model. To ensure that the tentative partial model is consistent with this denial, the compiler generates a rule of the form

$$\forall x, y (chk_i \leftrightarrow \neg(p(x) \wedge q(x, y)))$$

and to ensure that each sub-check (chk_i) is satisfied, they are included in the rule $nmr_check \leftarrow chk_1 \wedge \dots \wedge chk_k \wedge \dots$, which is transparently called after the program query by the s(CASP) interpreter (see Fig. 1, line 3).

However, this generate-and-test strategy has a high impact on the performance of programs that create many tentative models and use denials to discard those that do not satisfy the constraints of the problem.

Example 2 (Hamiltonian path problem).

Consider the Hamiltonian path problem, in which for a given graph we search for a cyclic path that visits each node of the graph only once. The standard ASP code for this problem, available at [hamiltonian.pl](#), is in Fig. 2. The conditions of the problem are captured (i) in line 6 to discard tentative paths that do not visit all the nodes, and (ii) in lines 7-8 to discard paths that have edges violating the properties of the Hamiltonian path. For the query in line 11, using the graph in Fig. 3 there are three stable models, one for each Hamiltonian cycle:


```

1 reachable(V) :- chosen(a, V).
2 reachable(V) :- chosen(U, V), reachable(U).
3 chosen(U, V) :- edge(U, V), not other(U, V).      % Choose or not an
4 other(U, V) :- edge(U, V), not chosen(U, V).      % edge of the graph.
5
6 :- vertex(U), not reachable(U).                  % Every vertex must be reachable.
7 :- chosen(U, W), U \= V, chosen(V, W).          % Do not choose edges to or
8 :- chosen(W, U), U \= V, chosen(W, V).          % from the same vertex.
9 #show chosen/2.
10
11 ?- reachable(a).                                % Is there a path from a to a?

```

Fig. 2. Code of the Hamiltonian problem à la ASP, available at [hamiltonian.pl](#).

```

1 % Graph      6 edge(b, a).   11 edge(c, d).
2 vertex(a).  7 edge(b, d).   12 edge(d, a).
3 vertex(b).  8 edge(a, c).   13 edge(c, a).
4 vertex(c).  9 edge(a, b).   14 edge(a, d).
5 vertex(d). 10 edge(b, c).   15 edge(d, b).

```

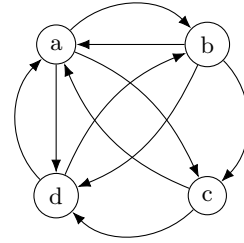


Fig. 3. Graph with 4 nodes available at [graph.4.pl](#).

```

{ chosen(a,c), chosen(c,d), chosen(d,b), chosen(b,a),... }
{ chosen(a,b), chosen(b,c), chosen(c,d), chosen(d,a),... }
{ chosen(a,d), chosen(d,b), chosen(b,c), chosen(c,a),... }

```

As mentioned before, the standard s(CASP) execution follows a generate-and-test scheme, choosing a cycle that reaches node a from node a and then discards any cycle in which:

- Not all vertices are reached (line 6), e.g., $\{\text{chosen}(a,b), \text{chosen}(b,a)\}$.
- Two chosen edges reach / leave the same vertex (line 7), e.g., $\{\text{chosen}(a,b), \text{chosen}(d,b), \text{chosen}(b,a)\}$ or $\{\text{chosen}(a,b), \text{chosen}(b,d), \text{chosen}(b,a)\}$.

As a consequence, if the evaluation chooses an edge that breaks any of these conditions, trying combinations with the rest of the edges would be misused effort.

3.2 Outline of the DCC Approach

The main idea behind our proposal is to anticipate the evaluation of the denials to fail and backtrack as early as possible. When an atom involved in a denial is a candidate to be added to a tentative model, the denial is checked to ensure that it is not already violated. By contrast, the classical implementation of s(CASP)

checked the tentative model as a whole once it had been completely built. The latter is akin to the generate-and-test approach to combinatorial problems, while the former tries to cut the search early.

In the most general case, DCC can take the form of a constraint-and-test instance. While detecting some denial violations can be performed by just checking the current state of the (partial) candidate model, better, more powerful algorithms can impose additional constraints on the rest of the atoms belonging to the partial model and also on the parts of the denials that remain to be checked.

These constraints can propagate additional conditions through the candidate model to ensure that it is consistent with the denials. These conditions will also remain active for the rest of the construction of the model, so that they can be carried forward, further reducing the search space. Note that since $s(\text{CASP})$ includes constraints *à la* CLP, the effect is very similar to the constraint propagation mechanisms that take place in constraint satisfaction systems, therefore making a full $s(\text{CASP}) + \text{DCC}$ system an instance of a constraint-and-generate evaluation engine.

The current implementation, which we describe in this paper, is a *proof of concept* that only checks grounds literals and does not anticipate the consistency check of constrained literals. As we will see later, this does not have a negative impact on the soundness of the system.

3.3 Implementation of DCC in $s(\text{CASP})$

The implementation of $s(\text{CASP}) + \text{DCC}$ is available as part of the source code of $s(\text{CASP})$ at <https://gitlab.software.imdea.org/ciao-lang/scasp>.

Compilation of the Denials As we mentioned before, the denials are compiled in such a way that the interpreter checks consistency by proving that for all possible values, the negation of the denial is satisfied. For example, for the denial $:- p(X), q(X,Y)$, the compiler generates the rule:

```

1  chk1 :- forall(X, forall(Y, not chk_body(X,Y))).
2  not chk_body(X,Y) :- not p(X).
3  not chk_body(X,Y) :- p(X), not q(X,Y).

```

The last clause includes a call to $p(X)$ to avoid duplicated solutions provided by the two clauses. If the interpreter is able to prove that for all the possible values of X and Y the tentative partial model is consistent with the predicate $\text{not chk_body}(X,Y)$, then it means that the tentative partial model is a stable partial model.

The approach followed by the DCC proposal is to detect when a rule like the above determines that a model candidate is inconsistent. If that is the case, $s(\text{CASP})$ fails and provokes backtracking to explore the generation of a different model. In the example above, if a model being generated is consistent with $p(X) \wedge q(X,Y)$, then it should be discarded.

Since it is only necessary to check for violation of denials when adding a goal involved in one of the denials, the compiler creates a series of rules that state

```

1  % The only clauses that changes      13 holds_dcc([],_).
2  s_loop(even,G,In,[chs(G)|In]) :-    14 holds_dcc([F_A|F_As],In) :-
3      eval_dcc(G,In). % New call      15     holds_dcc_(F_A,In),
4  s_loop(no_loop,G,In,Out) :-         16     holds_dcc(F_As,In).
5      pr_rule(G, Body),               17
6      solve(Body,[G|In],Out),         18
7      eval_dcc(G,In). % New call      19 holds_dcc_(F_A,In) :-
8                                          20     user_defined(F_A), !,
9  eval_dcc(G,In):-                    21     member(F_A,In).
10     \+ ( ground(G),                 22 holds_dcc_(F_A,In) :-
11         pr_dcc_rule(G,F_Atoms),      23     call(F_A).
12         holds_dcc(F_Atoms,In) ).

```

Fig. 4. Outline of the changes to the s(CASP) interpreter extended with DCC.

what has to be checked for in case a goal involved in the denial is generated. For the case above, if $p(X)$ is added, then $q(X,Y)$ has to be checked to ensure it does not hold, and the other way around. This is represented as:

```

1  dcc(p(X), [q(X,Y)]).
2  dcc(q(X,Y), [p(X)]).

```

which can be understood as “if $p(X)$ is present, check that $q(X,Y)$ is not present”.

In general, given a user defined denial⁹ of the form $:- c_a, b_1, \dots, b_n$ for each (negated) literal b_k , of a user defined predicate, the compiler generates a DCC rule $dcc(b_k, [c_a, \dots, b_{k-1}, b_{k+1}, \dots])$ for each $k, 1 \leq k \leq n$.

Extending the s(CASP) interpreter with DCC Fig. 4 shows the relevant fragment of the s(CASP) interpreter extended with Dynamic Consistency Checking. The basic intuition is that as soon as an atom that is involved in a denial is a candidate to be added to a model, DCC checks whether the candidate model is consistent with the rest of the atoms (including builtins) in that denial. Depending on the result of this check, the candidate atom is added or not. The modified interpreter performs the following steps:

1. The DCC check starts when the interpreter proves that a goal G holds, by invoking the predicate `eval_dcc/2`. If it succeeds, G is added to the model and the evaluation continues. Otherwise, an inconsistency has been detected and backtracking takes place.
2. The current implementation of `eval_dcc/2` only evaluates fully instantiated goals. Therefore, if the G is not ground (line 10), `eval_dcc/2` succeeds and the evaluation continues. This is not a source of unsoundness, as in any case the whole set of denials are checked before finally returning a model.
3. Otherwise, `pr_dcc_rule(G,F_Atoms)` (line 11) retrieves the DCC rules that involve goal G . If there are no DCC rules, `eval_dcc/2` succeeds.

⁹ The current implementation does not check denials introduced due to olon loops.

4. Since one atom in the rule (G) is to be added to the model, we need to ensure that not all the rest of the atoms in the rule (F_Atoms) appear in the candidate model In. In order not to instantiate the model, this is done by checking with `holds_dcc(F_Atoms, In)` whether *all* the atoms appear in In¹⁰ and then negating (by failure) the calling predicate.

Let us consider a program including the denial `:- p(X), q(X,Y)`. When the evaluation of `p(1)` succeeds, and before it is added to the tentative model, the interpreter calls `eval_dcc(p(1), [...])`, where `[...]` is the current tentative model. Since `p(1)` is ground, `pr_dcc_rule(p(1), F_Atoms)` retrieves in `F_Atoms` a list of the atoms that cannot be true in the model — in this case, `[q(1,Y)]`. Then, the interpreter checks if the literals in `[q(1,Y)]` are true in the current (tentative) model In. If that is the case, `holds_dcc([q(1,Y)], [...])` succeeds, `eval_dcc/2` fails, and the interpreter backtracks because the denial has been violated. Otherwise, the evaluation continues.

It is easy to see that this implementation of dynamic consistency checking is *complete*, i.e., we do not lose answers: since only ground goals are checked, there is no risk of instantiating free variables which could restrict degrees of freedom of the tentative model and therefore potentially removing solutions. Furthermore, to ensure *correctness*, we keep the non-monotonic rule checking that is performed once the tentative model is found. Note that non-ground goals are at the moment not subject to DCC rules, but they may be involved in denials, and denials of atoms not needed to support the query must be checked.

DCC is also used during the execution of the `nmr_check` predicate. As we mentioned before (item 3 in page 5), denials are compiled into a synthesized goal, `nmr_check`, that is executed after a model has been generated. During its execution, DCC rules are actively used to look for atoms that are introduced and when an inconsistency is flagged, execution fails and backtracks.

Example 3 (Cont. Example 2).

Let us consider the Hamiltonian program in Fig. 2. As explained above, the compiler generates the DCC rules below. For this example each denial is translated into two specialized rules.

```

1 dcc(vertex(U), [not_reachable(U)]).
2 dcc(not_reachable(U), [vertex(U)]).
3 dcc(chosen(U,W), [U \= V, chosen(V,W)]).
4 dcc(chosen(V,W), [chosen(U,W), U \= V]).
5 dcc(chosen(W,U), [U \= V, chosen(W,V)]).
6 dcc(chosen(W,V), [chosen(W,U), U \= V]).

```

By invoking `scasp --dcc11 hamiltonian.pl graph_4.pl, s(CASP)` evaluates the query `?- reachable(a)` following a goal-directed strategy. Let us refer to the code in Fig. 2 and the graph in Fig. 3 to explain how the evaluation with DCC takes place:

¹⁰ For builtins: checking that they succeed if evaluated in the environment of the model.

¹¹ The `--dcc` flag imply the use of `--prev_forall`, valid for non-constrained programs.

1. The query unifies with the clause in line 1 but the goal `chosen(a,a)` fails because `edge(a,a)` does not exist.
2. From the clause in line 2, `chosen(b,a)` is added to the tentative model, because no DCC rule succeeds. The goal `reachable(b)` is then called.
3. The goal `reachable(b)` unifies with the clause in line 1 and `chosen(a,b)` is added, because it is consistent with `chosen(b,a)`.
4. As the query succeeds for the model `{chosen(b,a), chosen(a,b), reachable(a), reachable(b), ...}`, `s(CASP)` invokes `nmr_check`.
5. `nmr_check` executes checks for all the denials. The code corresponding to line 6 is:

```

1  chk1 :- forall(U, not chk1_1(U)).
2  not chk1_1(U) :- not vertex(U).
3  not chk1_1(U) :- vertex(U), reachable(U).

```

i.e., all vertices (`vertex(U)`) must be reachable (`reachable(U)`). For vertices $U = a$ and $U = b$, `reachable(a)` and `reachable(b)` are already in the model, so there is nothing to check. But for vertex $U = c$, `reachable(c)` is not in the model and therefore `reachable(c)` has to be invoked while checking the denials.

6. From the clause in line 1, `chosen(a,c)` is selected to be added to the model, but it is discarded by DCC, because of the DCC rule `dcc(chosen(V,W), [chosen(U,W), U\=V])`, corresponding to the denial in line 7. Note that this DCC rule is instantiated to `dcc(chosen(a,c), [c\=b, chosen(a,b)])` and the literal `chosen(a,b)` is already in the model.
7. The evaluation backtracks and continues the search using another edge.

The denial in line 6 makes the interpreter to select edges to reach all vertices. The interleaving of the dynamic consistency checking prunes the search, which, as shown in Section 4, improves performance.

4 Evaluation

In this section we compare the performance of `s(CASP)` with and without DCC using a macOS 11.5.2 Intel Core i7 at 2.6GHz. We use `s(CASP)` version 0.21.10.09 available at <https://gitlab.software.imdea.org/ciao-lang/scasp> and, as mentioned before, all the benchmarks used or mentioned in this paper are available at <http://platon.etsii.urjc.es/~jarias/papers/dcc-pad121>.

Table 1 shows the results of the performance comparison (in seconds), and the speedup of `s(CASP)` with DCC w.r.t. `s(CASP)` – note that DCC activates the use of the *forall* predicate optimized for non-constrained programs, therefore, for a fair comparison, `s(CASP)` w.o. DCC should be executed including the `--prev_forall` flag.

First, we evaluate the Hamiltonian path problem (Example 2) using the encoding in Fig. 2 (available at [hamiltonian.pl](#)) and the graph with 4 vertices in Fig. 3 (available at [graph_4.pl](#)). We see that `s(CASP)` with DCC obtains a

Table 1. Performance comparison: s(CASP) and s(CASP)+DCC.

	Speedup	s(CASP)	s(CASP)+DCC
Hamiltonian (4 vertices)	10.0	11.985	1.196
Hamiltonian (7 vertices)	41.1	134.460	3.191
n_queens (n=4)	4.3	8.147	1.910
n_queens (n=5)	4.9	92.756	18.786
n_queens (n=6)	90.8	1362.840	15.001
n_queens_attack (n=6)	1.0	77.039	76.827

speedup of **10.0** \times . When the size of the graph is increased by adding three vertices¹² we obtain a speedup of **41.1** \times .

We also evaluated the performance of s(CASP) with DCC with the well-known n -queens problem, using two different versions. These examples are especially interesting, as they have no finite grounding and thus cannot be run by other implementations of the stable model semantics. In particular, the size of the board is not fixed by the programs and therefore the same code can be used to find solutions to several board sizes.

The first version, `n_queens` (available at [n_queens.pl](#)), uses denials to discard solutions where two queens attack each other. The speedup obtained by s(CASP) with DCC ranges from **4.3** \times (for $n = 4$) to **90.8** \times (for $n = 6$).

The second version is the one presented in [12, Pag. 37] (available at [n_queens_attack.pl](#)). In this case, the predicate `attack/3` is used to check whether a candidate queen attacks any previously selected queen. This version does not have denials, as the check is done as part of the computation, and therefore DCC checks are not useful.

Two interesting conclusions can be drawn from the numbers in the table:

- The execution time of `n_queens_attack` does not significantly change using or not DCC, which supports our assumption that the overhead of using DCC checks when they are not needed is negligible.
- On the other hand, the DCC-enabled execution for the version with denials (`n_queens`, column “s(CASP)+DCC”) is faster than the version without denials (`n_queens_attack`, column “s(CASP)”) for a factor of **4.1** \times . This can be attributed to the runtime being sophisticated enough to perform checks earlier and more efficiently than the hand-crafted code, even with the current, preliminary implementation.

Table 2 sheds some additional light on the effectiveness of DCC. It contains, for the same benchmarks as Table 1, how many models were returned, how many candidate models were discarded by `nqr_check` after they were generated (column “#models discarded – s(CASP)”), how many (partial) candidate models

¹² The graph with seven vertices is available at [graph-7.pl](#).

Table 2. Models generated and/or discarded: s(CASP) vs. s(CASP)+DCC.

	#models returned	#models discarded		#DCC detected
		s(CASP)	s(CASP)+DCC	
Hamiltonian (4 vertices)	3	7	7	52
Hamiltonian (7 vertices)	1	13	13	34
n_queens (n=4)	2	253	0	44
n_queens (n=5)	10	3116	0	167
n_queens (n=6)	4	46652	0	742

were discarded using DCCs (column “#models discarded – s(CASP) + DCC”) and how many times the dynamic consistency checking detects an inconsistency and backtracks (column “#DCC detected”).

Let us first focus on the n_queens benchmark.¹³ As the size of the board grows, the number of models that are completely generated and discarded by s(CASP) without using DCCs grows exponentially; this is of course the reason why its execution time also increases very quickly. If we look at the column “#models discarded – s(CASP) + DCC” we see that, when DCC is active, none of final models is discarded by `nmr_check`. That means that all models not consistent with the denials have been removed early, while they were being built. We also see that the number of times that DCC rules were activated is much smaller than the number of times that `nmr_check` was executed — early pruning made it possible to avoid attempting to generate many other candidate models. On top of that, executing DCC rules is done directly in Prolog, while `nmr_check` is executed using the s(CASP) metainterpreter, and the former is considerably faster than the latter. This adds to the advantage of early pruning.

The Hamiltonian path benchmark is different and very interesting. The number of models discarded by `nmr_check` is the same regardless of whether DCC is activated or not. That means that the DCC could not detect inconsistencies in candidate models. In this case the advantage of using DCC comes from applying it when `nmr_check` is invoked to ensure that the final model is consistent with the denials. `nmr_check` is executed as a piece of (synthesized) code by the metainterpreter. The denials in the Hamiltonian path not only check, but also generate new atoms which are checked by the DCC. This accelerates the execution of `nmr_check`, making it fail earlier, and it is the cause of the speedup of the Hamiltonian path benchmark.

5 Conclusions

In this paper, we have reported on a preliminary design and implementation of Dynamic Consistency Checking (DCC), a technique that anticipates the con-

¹³ n_queens_attack does not have denials, hence we do not include it in this table.

sistency evaluation of tentative models in s(CASP), a goal-directed (predicate) Constraint Answer Set Programming. This technique translates the denials to check them as early as possible rather than when a full model is found for a given query. Its ability to detect inconsistencies before a literal is added to the tentative model greatly increases the performance. with respect with executions without DCC. Early denial checking can also beat programs that use auxiliary predicates explicitly called from the user code to check for inconsistencies.

The current DCC implementation can still be improved, in particular to properly handle constraints by reducing the domain of constrained variables: checking denials using literals with constrained variables has to keep track of the domains of the variables, while avoiding introducing non-determinism when conflicting values are removed from the domain of the variables.

References

1. Arias, J., Chen, Z., Carro, M., Gupta, G.: Modeling and Reasoning in Event Calculus Using Goal-Directed Constraint Answer Set Programming. In: Pre-Proc. of the 29th Int'l. Symposium on Logic-based Program Synthesis and Transformation (September 2019)
2. Arias, J., Carro, M., Chen, Z., Gupta, G.: Justifications for goal-directed constraint answer set programming. In: Proceedings 36th International Conference on Logic Programming (Technical Communications). EPTCS, vol. 325, pp. 59–72. Open Publishing Association (2020). <https://doi.org/10.4204/EPTCS.325.12>
3. Arias, J., Carro, M., Chen, Z., Gupta, G.: Modeling and Reasoning in Event Calculus using Goal-Directed Constraint Answer Set Programming. CoRR **abs/2106.14566** (2021), <https://arxiv.org/abs/2106.14566>
4. Arias, J., Carro, M., Salazar, E., Marple, K., Gupta, G.: Constraint Answer Set Programming without Grounding. Theory and Practice of Logic Programming **18**(3-4), 337–354 (2018). <https://doi.org/10.1017/S1471068418000285>
5. Arias, J., Moreno-Rebato, M., Rodríguez-García, J.A., Ossowski, S.: Modeling Administrative Discretion Using Goal-Directed Answer Set Programming. In: Advances in Artificial Intelligence, CAEPIA 20/21. pp. 258–267. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-85713-4_25
6. Basu, K., Varanasi, S., Shakerin, F., Arias, J., Gupta, G.: Knowledge-driven Natural Language Understanding of English Text and its Applications. AAAI'21 (2021)
7. Basu, K., Wang, H., Dominguez, N., Li, X., Li, F., Varanasi, S.C., Gupta, G.: CASPR: A Commonsense Reasoning-based Conversational Socialbot. In: 4th Proceedings of Alexa Prize (Alexa Prize 2021). (2021)
8. Chen, Z., Marple, K., Salazar, E., Gupta, G., Tamil, L.: A Physician Advisory System for Chronic Heart Failure Management Based on Knowledge Patterns. Theory and Practice of Logic Programming **16**(5-6), 604–618 (2016). <https://doi.org/10.1017/S1471068416000429>
9. Dovier, A., Pontelli, E., Rossi, G.: A Necessary Condition for Constructive Negation in Constraint Logic Programming. Information Processing Letters **74**(3-4), 147–156 (2000)
10. Gelfond, M., Lifschitz, V.: The Stable Model Semantics for Logic Programming. In: 5th International Conference on Logic Programming. pp. 1070–1080 (1988), <http://www.cse.unsw.edu.au/~cs4415/2010/resources/stable.pdf>

11. Hall, B., Varanasi, S.C., Fiedor, J., Arias, J., Basu, K., Li, F., Bhatt, D., Driscoll, K., Salazar, E., Gupta, G.: Knowledge-Assisted Reasoning of Model-Augmented System Requirements with Event Calculus and Goal-Directed Answer Set Programming. In: Proc. 8th Workshop on Horn Clause Verification and Synthesis (2021)
12. Marple, K., Bansal, A., Min, R., Gupta, G.: Goal-Directed Execution of Answer Set Programs. In: Schreye, D.D., Janssens, G., King, A. (eds.) Principles and Practice of Declarative Programming, PPDP'12, Leuven, Belgium - September 19 - 21, 2012. pp. 35–44. ACM (2012). <https://doi.org/10.1145/2370776.2370782>
13. Marple, K., Gupta, G.: Dynamic Consistency Checking in Goal-Directed Answer Set Programming. *Theory and Practice of Logging Programming* **14**(4-5), 415–427 (2014). <https://doi.org/10.1017/S1471068414000118>
14. Marple, K., Salazar, E., Gupta, G.: Computing Stable Models of Normal Logic Programs Without Grounding. arXiv **1709.00501** (2017), <http://arxiv.org/abs/1709.00501>
15. Marriott, K., Stuckey, P.J.: Programming with Constraints: an Introduction. MIT Press (1998)
16. Mellarkod, V.S., Gelfond, M., Zhang, Y.: Integrating Answer Set Programming and Constraint Logic Programming. Tech. rep., Texas Tech University (October 2008), this is a long version of [17], available at <https://www.depts.ttu.edu/cs/research/documents/46.pdf>
17. Mellarkod, V.S., Gelfond, M., Zhang, Y.: Integrating Answer Set Programming and Constraint Logic Programming. *Annals of Mathematics and Artificial Intelligence* **53**(1-4), 251–287 (2008)
18. Morris, J.: Constraint answer set programming as a tool to improve legislative drafting: a rules as code experiment. In: ICAIL. pp. 262–263. ACM (2021)
19. Pereira, L.M., Aparício, J.N.: Relevant counterfactuals. In: EPIA 89, 4th Portuguese Conference on Artificial Intelligence, Lisbon, Portugal, September 26-29, 1989, Proceedings. pp. 107–118 (1989). https://doi.org/10.1007/3-540-51665-4_78
20. Shakerin, F., Gupta, G.: Induction of Non-Monotonic Logic Programs to Explain Boosted Tree Models Using LIME. In: AAAI 2019. pp. 3052–3059 (January - February 2019). <https://doi.org/10.1609/aaai.v33i01.33013052>
21. Stuckey, P.: Constructive negation for constraint logic programming. In: Proc. LICS'91. pp. 328–339 (1991)
22. Varanasi, S.C., Salazar, E., Mittal, N., Gupta, G.: Synthesizing Imperative Code from Answer Set Programming Specifications. In: LOPSTR. *Lecture Notes in Computer Science*, vol. 12042, pp. 75–89. Springer (2019). https://doi.org/10.1007/978-3-030-45260-5_5