Evaluation Techniques and Systems for Answer Set Programming: A Survey

Martin Gebser,[1] **Nicola Leone,**[2] **Marco Maratea,**[3] **Simona Perri,**[2]
Francesco Ricca[2] and **Torsten Schaub**[1]
[1] University of Potsdam, Germany
[2] University of Calabria, Italy
[3] University of Genova, Italy

Abstract

Answer set programming (ASP) is a prominent knowledge representation and reasoning paradigm that found both industrial and scientific applications. The success of ASP is due to the combination of two factors: a rich modeling language and the availability of efficient ASP implementations. In this paper we trace the history of ASP systems, describing the key evaluation techniques and their implementation in actual tools.

1 Introduction

Answer Set Programming (ASP) [**?**] is a well-known approach to knowledge representation and reasoning, with roots in the areas of logic programming and non-monotonic reasoning [**?**]. ASP is in close relationships to other formalisms such as Propositional Satisfiability (SAT), Satisfiability Modulo Theories (SMT), Quantified Boolean Formulas (QBF), Constraint Programming (CP), Planning and Scheduling, and many others.

The language of ASP allows for defining solutions to complex problems in a purely declarative way. The main construct is a rule, i.e., an expression of the form $Head \leftarrow Body$, where $Body$ is a logic conjunction possibly involving negation, and $Head$ is, in the basic formulation, either an atomic formula or a logic disjunction. ASP programs, i.e., collection of rules, are first order theories that can model uniformly solutions for a given problem over varying instances. The semantics of ASP programs is given in terms of answer sets (or stable models) [**?**]. The basic language introduced in [**?**] has been extended over the years by introducing many new constructs [**?**; **?**], that further simplify the modeling of complex problems. A standardized syntax for ASP has been then introduced in [**?**].

The readability and expressiveness of the language combined with the availability of efficient implementations made ASP one of the major declarative paradigms for logic-based Artificial Intelligence (AI), suitably employed for realizing many AI applications [**?**].

The research of efficient evaluation techniques for the language of ASP started during the 90-ties [**?**], after a first period in which complexity arguments discouraged the implementation of systems. Indeed, the first serious attempts appeared about 10 years after the introduction of the stable model semantics. The subsequent introduction of new evaluation techniques, often fruit of cross-fertilization with the neighboring communities of SAT and CP, resulted in the design of more and more efficient ASP systems.

The advancement of the state of the art in ASP solving was consistent and continuous in the last few years, as witnessed by the results of the biannual ASP Competitions series (see [**?**; **?**] for the last events). This paper traces this history surveying the major contributions to the development of evaluation techniques and solvers for ASP that made it one of the most attractive paradigms of logic-based AI.

2 The History of ASP Evaluation

Almost all ASP implementations follow a two-step evaluation process [**?**]. The first step, called *grounding*, takes an ASP program as input and produces an equivalent variable-free (or *ground*) program that has the same answer sets as the input program. Such a ground program is the input of the second step of the process, called *solving*, having the role of searching for one or more answer sets, which coincide with the ones of the non-ground program.

In the following we first present the evolution of grounders, followed by the history of solvers; then, we mention the approaches blending existing techniques in portfolios, those capable of exploiting multi-processing, and, finally, we review some (relatively new) attempts deviating from the customary two-steps evaluation, called *grounding-less* systems, where grounding and solving are somehow interleaved.

2.1 Grounders

Grounders perform a complex task that may have a big impact on the performance of the whole system, as their output is the input for the solving phase: if input non-ground programs can be assumed to be fixed (data complexity), grounding is polynomial; however, as soon as variable programs are given in input the produced ground program is potentially of exponential size with respect to the input program. Thus, grounders employ smart procedures that are geared toward efficiently producing a ground program that preserves the semantics but is sensibly smaller than the one that could be obtained by a simple replacement of the variables with all the constants of the program.

One of the first released grounders was LPARSE [?], a front-end grounder system, whose output encoded in a suitable numeric format, is intended to be given as input to a separated solver. LPARSE accepts logic programs respecting its ω-restrictedness condition, i.e. it enforces each variable in a rule to occur in a positive body literal, called domain literal, whose predicate (i) is not mutually recursive with the head, and (ii) is neither unstratified nor dependent (also transitively) on an unstratified predicate. A similar approach is used in the earlier versions of GRINGO [?], that bind non-global variables by domain predicates to enforce a λ-restrictedness condition, an extension of ω-restrictedness. Essentially, these restrictions are used in order to treat positive recursion among predicates, and guarantee to have a finite grounding.

The grounder of the DLV system [?], and the more recent versions of GRINGO (starting from 3.0) instead imposes the less restrictive condition of *safety*, requiring that each variable in a rule appears in some positive body literal. These grounders are based on semi-naive database evaluation techniques [?] for avoiding duplicate work during grounding. Grounding is seen as an iterative bottom-up process guided by the successive expansion of a program's term base, that is, the set of variable-free terms constructible from the signature of the program at hand.

Recently a brand new version of the DLV grounder has been released, as the stand-alone grounder I-DLV [?]. The new grounder relies on the theoretical foundations of its predecessor but has been completely redesigned, integrating new optimizations, and usability features such as, annotations, python interface, interoperability mechanisms.

2.2 Solvers

The search for (optimal) answer sets (i.e., ASP solving) basically amounts to Boolean constraint solving. This task was pioneered in the area of SAT by the classical Davis-Putnam-Logemann-Loveland (DPLL) procedure [?]; more recently, it has been solved by resorting to the Conflict-Driven Clause Learning (CDCL) [?] algorithm. Similarly, the so-called *native* answer set solvers are based on the same backtracking procedures used in SAT solving, yet refined to the semantics and additional constructs of ASP programs. Alternatively, *translation-based* solvers can be obtained by applying transformations to ground programs to obtain instances of other formalisms featuring efficient solvers, and exploiting these solvers for computing answer sets.

Native Approaches. The first developed ASP solvers, namely DLV [?] and SMODELS [?], were pioneered in the late '90s and pursued "native" approaches based on the classical DPLL procedure. While DPLL augments basic backtracking search with unit propagation on clauses, DLV and SMODELS adjust such techniques to ASP programs. In particular, this includes dedicated inference mechanisms to detect and falsify so-called unfounded sets [?; ?], which particularly address positive recursion in case of non-tight programs. Later on, DLV was extended with backjumping techniques [?], in place of basic backtracking, and look-back heuristics [?] that take advantage of the backjumping process [?]. Similarly, the SMODELSCC solver [?] extended the algorithm of SMODELS with backjumping, while further adding mechanisms for conflict-driven clause learning, as pioneered by CDCL in the area of SAT.

The second generation of native ASP solvers, including CLASP [?] and WASP [?], integrates CDCL-style search with propagation principles dedicated to ASP programs. Implementation features shared with modern SAT solvers include, e.g., watched literals, activity-based heuristics, and rapid restarts [?]. Such basic features are accompanied by techniques for dealing with unfounded sets, aggregates, and optimization in order to cover the range of modeling concepts and computational tasks available in ASP [?]. While the IDP system [?] has been conceived as a model generator for first-order theories extended with inductive definitions, it has much in common with the aforementioned ASP grounders and solvers. That is, it includes a grounder, GIDL [?], a solver, MINISATID [?], and handles (positive) recursion among atoms in inductive definitions.

For solving a program P which is not-Head-Cycle-Free (non-HCF), i.e. where checking whether a model of P is $\subseteq$-minimal w.r.t. P^I is in general coNP-complete, native ASP solvers employ a two-level architecture in which DPLL- or CDCL-style search is used for (i) generating candidate models of P and (ii) checking for the existence of smaller counter-models of P^I. To this end, the propagation principles of the first respective solver, DLV, are capable of handling disjunctive rules [?], while the check for counter-models is delegated to a SAT solver. The GNT system [?] pursues a corresponding approach by casting the tasks of generating and checking candidate models to ASP programs processed with separate instances of SMODELS. Similarly, CLASP (or its nowadays deprecated sibling CLASPD) and WASP couple complementary instances of their CDCL-style search engines to perform model generation or checking, respectively. Both CLASP and WASP encode the checking task for arbitrary candidate models [?; ?], and perform checking via assumption-based reasoning [?].

Translation-Based Approaches. From ? (?), we know that answer sets of a tight program [?] P coincide with propositional models of P's completion the answer sets of tight programs can be computed by running SAT solvers. By relying on this result, the first version of the SAT-based solver CMODELS [?] was based on this correspondence. As a generalization to the non-tight case, [?] proposed loop formulas whose addition to a program's completion establishes correspondence between propositional models and answer sets. Since the number of required loop formulas can be exponential [?], the SAT-based solvers ASSAT [?] and CMODELS, from its second version on, add loop formulas incrementally to eliminate models that are no answer sets. In fact, loop formulas deny unfounded sets [?], which are also handled by native systems, so that there is a close proximity between native and SAT-based solvers utilizing loop formulas, and both kinds of systems are based on similar search procedures. This also carries forward to non-HCF programs [?], where the third version [?] of CMODELS utilizes SAT solvers also for stability checking.

The translation by LP2SAT [?; ?], instead, is based on so-called level rankings [?] to check $\subseteq$-minimality w.r.t. the

reduct of an HCF program in the non-tight case. Such level rankings are encoded a priori, rather than incrementally, and expressing them in SAT requires sub-quadratic instead of exponential space. Technically, the tool LP2ACYC [?] instruments an ASP program such that propositional models of its completion subject to an acyclicity condition match the answer sets of the program. The required acyclicity can then be established via level rankings, where linear representations are feasible in several target formalisms, including ASP, Pseudo-Boolean Constraints/Optimization, or SAT Modulo Acyclicity [?; ?], SMT with Difference or Bit-Vector Logic [?; ?], and Mixed Integer Programming [?]. In fact, the translation-based systems participating in the Sixth ASP Competition [?] are based on this infrastructure, while SAT-based solvers utilizing loop formulas have come out of fashion.

2.3 Portfolio Approaches

Automated algorithm selection techniques [?] aim at robustness across a range of heterogeneous inputs. Inspired by SATZILLA in the area of SAT, the CLASPFOLIO system [?; ?] uses support vector regression to learn scoring functions approximating the performance of several CLASP variants in a training phase. Given an instance, CLASPFOLIO then extracts features and evaluates such functions in order to pick the most promising CLASP variant for solving the instance. This algorithm selection approach was particularly successful in the Third ASP Competition [?], held in 2011, where CLASPFOLIO won the first place in the NP category and the second place overall (without participating in the Beyond-NP category). The ME-ASP system [?; ?] goes beyond the solver-specific setting of CLASPFOLIO and chooses among different grounders as well as solvers. Grounder selection traces back to [?], and similar to the QBF solver AQME [?], ME-ASP uses a classification method for performance prediction. Notably, "bad" classifications can be treated by adding respective instances to the training set of ME-ASP [?], which enables an adjustment to new problems or instances thereof. In the Seventh ASP Competition [?], the winning system was I-DLV+S that utilizes I-DLV for grounding and automatically selects back-ends for solving through classification between CLASP and WASP.

Going beyond the selection of a single solving strategy from a portfolio, the ASPEED system [?] indeed runs different solvers, sequentially or in parallel, as successfully performed by PPFOLIO in the 2011 SAT Competition. Given a benchmark set, a fixed time limit per instance, and performance results for candidate solvers, the idea of ASPEED is to assign time budgets to the solvers such that a maximum number of instances can be completed within the allotted time. In other words, the goal is to divide the total runtime per computing core among solvers such that the number of instances on which at least one solver successfully completes its run is maximized. The portfolio then consists of all solvers assigned a non-zero time budget along with a schedule which solvers to run on the same computing core. Calculating such an optimal portfolio for a benchmark set is an Optimization problem addressed with ASP in ASPEED.

2.4 Multi-processing in ASP Systems

We below describe approaches taking advantage of multi-core/processor support during both grounding and/or solving.

Concerning grounding, parallel grounding techniques were developed as extensions of LPARSE [?] and the DLV grounder [?]. The former approach was designed for distributing LPARSE incarnations, working in local memory, on Beowulf clusters, while the latter aims at shared-memory parallelism on multi-core/processor machines. In particular, the parallel version of the DLV grounder allows for a concurrent instantiation at several levels: it allows for instantiating in parallel subprograms of the program in input, rules within a given subprogram, and also for parallelizing the evaluation of a single rule. This is accompanied by techniques for granularity control and dynamic load balancing to achieve an efficient parallelization.

Concerning ASP solvers, these have been also extended to exploit multi-core/multi-processor machines by introducing parallel evaluation methods. The first approaches in this direction [?; ?; ?; ?] were based on SMODELS and divided its DPLL-style search among cluster machines or multiple threads, primarily using guiding paths [?], pioneered in the area of SAT, to process separate subproblems in parallel. Cluster and multi-threaded versions of the CDCL-based solver CLASP [?; ?], however, turned out to be particularly successful when applying a parallel portfolio of CLASP variants to a common problem. Recent versions of CLASP [?] further extend the multi-threaded search infrastructure to non-HCF programs [?]. The aforementioned ASPEED system allows for scheduling solver runs on multiple computing cores. Also, translation-based systems may readily exploit parallelism provided by respective back-end solvers, as done in the 2013 and 2014 editions of the ASP Competition. Finally, [?] provide an approach to parallel CDCL-style ASP solving utilizing GPUs.

2.5 Grounding-less ASP Systems

In contraposition with the traditional ground&solve approach, systems such as *Gasp* [?], *Asperix* [?], and *Omiga* [?] perform a *lazy grounding* technique in which grounding and solving steps are interleaved, and rules are grounded on-demand during solving. These systems try to overcome the so called *grounding bottleneck*, that occurs on problems for which the instantiation is inherently so huge that the traditional approach is not suitable, and this occurs also when problems can be solved in polynomial space [?]. However, in general, they suffer from poor search performance since they do not perform conflict-driven learning and backjumping for reducing the search space. Recently, a new ASP system, namely *Alpha* [?], has been developed aiming at blending lazy-grounding and learning of non-ground clauses. Finally, [?] proposed *lazy model expansion* within the FO(ID) formalism supporting partially lazy instantiation of constraints.

3 System Interfaces

Given that the implementations of ASP systems are nowadays highly optimized and sophisticated, modifying a system to support a richer language or dedicated reasoning techniques

is non-obvious. In order to facilitate making such extensions, modern ASP systems like CLINGO [**?**; **?**] and WASP [**?**; **?**] provide APIs.

The interface of CLINGO offers customized control about grounding and solving, where specific routines can be developed in *C* as well as the scripting languages *lua* and *python*. One major target application is *incremental solving* [**?**], where parts of an ASP program are successively (re)grounded and search is repeatedly performed w.r.t. the growing ground instantiation. Corresponding techniques have been successfully utilized in areas like automated planning [**?**] and finite model finding [**?**], where the horizon needed for a problem solution is a priori unknown or its theoretical bound prohibitively large, respectively. The second main application of CLINGO's interface consists of its extension by *external propagators* [**?**], implementing customized reasonings on top of the basic CDCL-style search.

Along the same lines, the interface of WASP allows for equipping its search procedure with external propagators, which can be implemented in *C++*, *perl* and *python*, for extending the basic reasoning capabilities. As demonstrated in [**?**], the incorporation of propagators can overcome the *grounding bottleneck* on several combinatorial problems. Moreover, WASP's interface supports the integration of custom *search heuristics* [**?**], and respective methods were successfully applied to the industrial Partner Units [**?**] and Combined Configuration [**?**] problems. A first attempt of combining I-DLV and WASP in a monolithic design has been described in [**?**]. The resulting project is called DLV2, and combines the python interface of WASP with the one offered by I-DLV with the aim of easier the development of propagators or heuristics.

4 Systems for ASP Extensions

Advanced frameworks that extend the common syntax and semantics of ASP programs include Constraint ASP (CASP) [**?**; **?**; **?**], ASP Modulo Theories (ASPMT) [**?**], Bound-Founded ASP (BFASP) [**?**], and Higher-order EXternal (HEX) programs [**?**]. The idea of CASP systems like ACSOLVER, CLINGCON, EZCSP, IDP, INCA, and MINGO, whose key features are analyzed and compared in [**?**], is to augment the Boolean problem representations of ASP with constraints over multi-valued variables in order to compactly formulate quantitative conditions. Corresponding implementation techniques are inspired by SAT Modulo Theories (SMT) solvers, and ASPMT systems like ASPMT2SMT and EZSMT furnish translations to exploit SMT solvers as search back-ends. The BFASP system CHUFFED extends a Constraint Programming (CP) solver with a propagator for bound-founded integer variables that default to either the smallest or largest value available in their domains, thus generalizing the minimality condition on answer sets from Boolean to multi-valued variables. HEX programs allow for equipping ASP programs with external sources of knowledge and/or computation, and the DLVHEX system features techniques to integrate the evaluation of external sources into the search of ASP solvers. Application domains making use of ASP extensions as furnished by the CASP, ASPMT, BFASP, and HEX frameworks include, e.g., hybrid task and motion planning in robotics, time scheduling, as well as ontological reasoning.

5 Conclusion

The research of techniques for implementing the stable model semantics started about twenty years ago. This paper retraces this history surveying the major contributions that made ASP one of the most attractive paradigms of logic-based AI.